



ICCM Proceedings

**Proceedings
of the International Conference
on Computational Methods**

(Vol.5, 2018)

9th ICCM, 6th-10th August 2018, Rome, Italy

Editors: G. R. Liu, Patrizia Trovalusci

ICCM2018

Proceedings of the International Conference on
Computational Methods (Vol.5, 2018)

9th ICCM, 6th-10th August 2018, Rome, Italy

Edited by

G. R. Liu

University of Cincinnati, USA

Patrizia Trovalusci

Sapienza University of Rome, Italy

Proceedings of the International Conference on Computational Methods, Vol.5, 2018

This volume contains full papers accepted by the 9th ICCM, 6th-10th August 2018, Rome, Italy

First Edition, August 2018

International Standard Serial Number: ISSN 2374-3948 (online)

Papers in this Proceedings may be identically cited in the following manner: Author names, *Paper title, Proceedings at the 9th ICCM2018, 6th-10th August 2018, Rome, Italy*, Eds: G.R. Liu, Patrizia Trovalusci, ScienTech Publisher.

Note: The papers/data included in this volume are directly from the authors. The editors are not responsible of the inaccuracy, error, etc. Please discuss with the authors directly, if you have any questions.

Published by
Scienteck Publisher LLC, USA
<http://www.sci-en-tech.com/>

PREFACE

On behalf of the organizing committee and the co-chairs, we would like to welcome you to the 9th *International Conference on Computational Methods* (ICCM2018) at the Auditorium Antonianum, Rome, Italy, between August 6th and 10th, 2018. The conference aims at providing an international forum for scholars, researchers, industry practitioners, engineers and graduate and undergraduate students to promote exchange and disseminate recent findings on both contemporary and traditional subjects in computational methods, numerical modeling and simulation, and their applications in science and engineering. It accommodates presentations on a wide range of topics to facilitate interdisciplinary exchange of ideas in science, engineering and allied disciplines, and helps to foster collaborations.

Computational Modelling and Simulation are fundamental subjects in engineering and sciences. They can be applied to many of the primary engineering disciplines, including Aerospace, Bio-medical, Civil, Chemical, Mechanical, and Materials Engineering among others. Computational Modeling and Simulation covers a broad range of research areas, from conventional structural and mechanical designs, failure analysis, dynamic and vibration analysis, and fluid mechanics up to cutting-edge computational mechanics, nano-micro mechanics, multiscale mechanics, coupled multi-physics problems and novel materials. This is reflected in the variety of fields featured in the conference topics.

The genesis of the ICCM series dates back to 2004, when the first ICCM2004 conference was held in Singapore founded and chaired by Professor Gui-Rong Liu, followed by ICCM2007 in Hiroshima, Japan, ICCM2010 in Zhangjiajie, China, ICCM2012 in Gold Coast, Australia, ICCM2014 in Cambridge, UK, and ICCM2015, Auckland, New Zealand, ICCM2016, Berkeley, California, USA, ICCM2017, Guilin, Guangxi, China. The present ICCM conference in Rome, Italy encompasses about 330 oral presentations organized in 64 Mini-Symposia and general sessions, including 3 Plenary Lectures, 14 Thematic Plenary Lectures, and several Keynotes.

The ICCM conference is unique in the sense that it showcases the current developments and trends in the general topic of Computational Methods and their relationship to global priorities in science and engineering. We would like to express our gratitude for the contributions of all ICCM2018 participants and presenters at this international event. We gratefully acknowledge the contributions from the International Scientific Committee, Mini-Symposium Organizers, and expert reviewers for their efforts and assistance in the organization. Special thanks go to Dr. Nicholas Fantuzzi for his efficient assistance to the scientific organization of the Conference and his patient handling of bureaucratic issues. We thank also the volunteers and the local staff in helping out in the organizing and running this important event.

Finally, we would like to warmly thank you for the contribution of our authors and participants in making ICCM2018 in Rome a very prominent scientific event. We believe the ICCM2018 has become a special event that widens the bridge between West and East in our worldwide community for computational methods.

We are looking forward to your participation and continued engagement for the future ICCM conferences, and contribute further in the development of computational methods.

Professor Patrizia Trovalusci
Conference Chairman, ICCM2018
Sapienza, University of Rome, Italy

Professor Gui-Rong Liu
Honorary Chairman, ICCM2018
University of Cincinnati, USA

ORGANIZATION COMMITTEES

Chairman: Professor Patrizia Trovalusci, Sapienza University of Rome, Italy.

Honorary Chairman: Professor Gui-Rong Liu, University of Cincinnati, USA

International Co-Chairs

Raj Das (RMIT University Melbourne, Australia)

Nasr Ghoniem (University of California at Los Angeles, USA)

Qing Li (University of Sydney, Australia)

Xikui Li (Dalian University of Technology, China)

Martin Ostojic-Starzewski (University of Illinois at Urbana Champaign, USA)

Tomasz Sadowski (University of Lublin, Poland)

Local Co-Chairmen

Ferdinando Auricchio (University of Pavia, Italy)

Fernando Fraternali (University of Salerno, Italy)

Luigi Gambarotta (University of Genova, Italy)

Stefano Lenci (Marche Polytechnic University, Italy)

Walter Lacarbonara (Sapienza University of Rome, Italy)

Raimondo Luciano (University of Cassino, Italy)

Anna Pandolfi (Politecnico di Milano, Italy)

Antonina Pirrotta (University of Palermo, Italy)

Alessandro Reali (University of Pavia, Italy)

Nicola Luigi Rizzi (University of Roma Tre, Italy)

Giuseppe Rega (Sapienza University of Rome, Italy)

Elio Sacco (University of Naples Federico II, Italy)

Giorgio Zavarise (Politecnico di Torino, Italy)

International Organizing Committee

Daniela Addressi (Sapienza University of Rome, Italy)

Elena Benvenuti (University of Ferrara, Italy)

Carlo Callari (University of Molise, Italy)

Rossana Dimitri (University of Salento, Italy)

Vincenzo Gattulli (Sapienza University of Rome, Italy)

Leon Mishnaevsky (Technical University of Denmark, Denmark)

Mahmood Jabareen (Technion - Israel Institute of Technology, Israel)

Francesco Tornabene (University of Bologna, Italy)

Giuseppe Vairo (Tor Vergata University of Rome, Italy)

Secretary General

Nicholas Fantuzzi (University of Bologna, Italy)

Local Organizing Committee

Patrizia Trovalusci, Nicholas Fantuzzi, Emanuele Reccia, Marco Pingaro, Marco Pepe, Ada Amendola

Local Staff

Roberto Panei, Alessandro Tinelli, Anna Irene Del Monaco

Ada Amendola, Maria Laura De Bellis, Mariella De Piano, Giorgia Di Gangi, Paolo Di Re, Cristina Gatta, Lorenzo Leonetti, Mariacarla Nocera, Marco Pepe, Marco Pingaro, Emanuele Reccia, Marialuigia Sangirardi, Valeria Settimi, Michela Talò

Secretaries

Com-it (Salerno, Italy)

International Scientific Advisory Committee

Addessi Daniela (Italy)	Leitao Vitor (Portugal)	Shen Lian (USA)
Andrade Jose (USA)	Lenci Stefano (Italy)	Shioya Ryuji (Japan)
Auricchio Ferdinando (Italy)	Leo Hwa Liang (Singapore)	Shu Chang (Singapore)
Benvenuti Elena (Italy)	Li Chenfeng (United Kingdom)	Sibanda Precious (South Africa)
Birken Philipp (Sweden)	Li Eric Quanbing (Hong Kong)	Siegfried Schmauder (Germany)
Bui Tinh Quoc (Japan)	Li Hua (Singapore)	Skelton Robert E. (USA)
Callari Carlo (Italy)	Li Qing (Australia)	Sladek Vladimir (Slovakia)
Chen Bin (China)	Li Wei (China)	Starzewski Martin Ostoja (USA)
Chen Chuin-Shan David (Taiwan)	Li Xikuili (China)	Stefanou George (Greece)
Chen Jeng-Tzong (Taiwan)	Li Yan (USA)	Sun Waiching (USA)
Chen Lei (USA)	Linder Christian (USA)	Sweilam Nasser Hassan (Egypt)
Chen Shaohua (China)	Liu Moubin (China)	Tadano Yuichi (Japan)
Chen Zhen (USA)	Liu Yan (China)	Tan B.C. Vincent (Singapore)
Cheng Yuan (Singapore)	Liu Yijun (USA)	Tanaka Satoyuki (Japan)
Chisari Corrado (United Kingdom)	Liu Yinghua (China)	Tian Rong (China)
Colaco Marcelo (Brazil)	Liu Zhaomiao (China)	Tian Zhaofeng (Australia)
Cui Fangsen (Singapore)	Lu Weizhen Jane (Hong Kong)	Toklu Cengiz (Turkey)
Das Raj (Australia)	Luciano Raimondo (Italy)	Tornabene Francesco (Italy)
Dimitri Rossana (Italy)	Luo Zhen (Australia)	Trovalusci Patrizia (Italy)
Dulikravich George (USA)	Ma Qingwei (United Kingdom)	Tsubota Ken-Ichi (Japan)
Effenhauser Carlo (Brazil)	Macorini Lorenzo (United Kingdom)	Vairo Giuseppe (Italy)
Fantuzzi Nicholas (Italy)	Manzari Majid (USA)	Wan Decheng (China)
Fraternali Fernando (Italy)	Matsubara Hitoshi (Japan)	Wang Cheng (China)
Fuschi Paolo (Italy)	Miller Karol (Australia)	Wang Hu (China)
Gambarotta Luigi (Italy)	Misnhaewsky Leon (Denmark)	Wang Jie (China)
Gan Yixiang (Australia)	Nagashima Toshio (Japan)	Wang Lifeng (China)
Gao Xiaowei (China)	Natarajan Sundararajan (India)	Wang Xianqiao (USA)
Gattulli Vincenzo (Italy)	Ng Alex (Australia)	Wang Yuesheng (China)
Geers M.G.D. (Netherlands)	Nithiarasu Perumal (United Kingdom)	Wang Yunzhi (USA)
Gerasimov Alexander (Russia)	Niu Yang-Yao (Taiwan)	Wu Hengan (China)
Ghoniem Nasr (USA)	Nogueira Xess (Spain)	Xiao Feng (Japan)
Ghosh Somnath (USA)	Noto Francesco (Italy)	Xiao Jinyou (China)
Greco Fabrizio (Italy)	Ogino Masao (Japan)	Yang Qingsheng (China)
Gu Yuantong (Australia)	Onishi Yuki (Japan)	Yang Richard Chunhui (Austria)
Guan Zhongwei (United Kingdom)	Oudjene Marc (France)	Yang Zhenjun (China)
Guo Wanlin (China)	Packo Pawel (Poland)	Yao Jianyao (China)
Hagihara Seiya (Japan)	Pandolfi Anna (Italy)	Ye Hongling (China)
Hirose Sohichi (Japan)	Paolone Achille (Italy)	Ye Wenjing (Hong Kong)
Hou Shujuan (China)	Papadrakakis Manolis (Greece)	Yeo Jingjie (Singapore)
Huang Yu (China)	Perego Umberto (Italy)	Yosibash Zohar (Israel)
Huang Zheng-Ming (China)	Picu Catalin R. (USA)	Yu Chengxiang Rena (Spain)
Jabareen Mahmood (Israel)	Pirrotta Antonina (Italy)	Yvonne Julien (France)
Jacobs Gustaaf (USA)	Popp Alexander (Germany)	Zeidan Dia (Jordan)
Jiang Chao (China)	Prakash Jagdish (Botswana)	Zelepugin Sergey (Russia)
Jin Feng (China)	Qin Qinghua (Australia)	Zhang Chuanzeng (German)
Kanayama Hiroshi (Japan)	Realì Alessandro (Italy)	Zhang Guiyong (China)
Kang Zhan (China)	Rebielak Janusz (Poland)	Zhang Jian (China)
Khennane Amar (Australia)	Reddy Daya (South Africa)	Zhang Liangchi (Australia)
Khoo Boo-Cheong (Singapore)	Rizzi Nicola Luigi (Italy)	Zhang Lihai (Australia)
Khurram Rooh (Saudi Arabia)	Sacco Elio (Italy)	Zhang Lucy (USA)
Koh Soojin Adrian (Singapore)	Sadowsky Tomasz (Poland)	Zhang Xiong (China)
Koshizuka Seichi (Japan)	Saitoh Takahiro (Japan)	Zhang Zhao (China)
Kougioumtzoglou Ioannis (USA)	Sakai Mikio (Japan)	Zhou Kun (Singapore)
Lacarbonara Walter (Italy)	Sarler Bozidar (Slovenia)	Zhuang Zhuo (China)
Lee Chin-Long (New Zealand)		
Lee Ik-Jin (South Korea)		

Table of Contents

Preface	iii
Committees	iv
Table of Contents	vi
Study of Error based on Six Segments Ring Pair Electrical Resistance Sensor <i>Xiaona Wang, Yi Ding, Shide Song, Yi Huang</i>	1
Optimal design and error analysis based on pulsed eddy current sensor <i>Shide Song, Tianyang Qu, Xiaona Wang, Yi Huang</i>	12
Mode shapes complexity for damage identification of structures experiencing plasticization <i>Fabrizio Iezzi, Claudio Valente</i>	19
Parametrization of radiative properties of mono- and multi-component plasmas for astrophysics and nuclear fusion applications <i>Rafael Rodriguez, Guadalupe Espinosa, Juan Miguel Gil</i>	29
Droplet morphology simulation with SPH: A simple method of implementing contact angles <i>Xiangwei Dong</i>	42
Relationship between Train Speed and Aural Discomfort in Tunnels Based on Tympanic Membrane Model <i>Yong Peng, Pengpeng Xie, Shengen Yi, Junjiao Hu</i>	62
Numerical Study on Debonding Failure between FRP and Concrete <i>Huang Lihua, Yang Zhiquan, Wang Yuefang</i>	76
Simulation of liquid cargo - vehicle interaction under lateral and longitudinal accelerations <i>Jose A. Romero, Frank Otremba, Alejandro A Lozano-Guzman</i>	88
Fractional Order Derivative Computation with a Small Number of Discrete Input Values <i>Dariusz Wojciech Brzezinski</i>	103
Interface effect on failure of ceramic coating/alloy substrate systems <i>Lihong Liang</i>	116
Numerical Simulation and Experimental Verification of Ti/APC-2/Kevlar Hybrid Composite Laminates due to Low-Velocity Impact <i>Ming-Hwa R. Jen, Dong-Yi Cai, Che-Kai Chang, Feng-Chi Hsu</i>	120
A data-driven diagnostic tool for wind turbines under operational variability <i>Simona Bogoevska, Eleni Chatzi, Elena Dumova-Jovanoska, Rudiger Hoffer</i>	128
Numerical modeling of free-surface wave effects on flexural vibration of floating structures <i>Shahrokh Sepehri Ahnema, Eng Teo Ong, Heow Pueh Lee, Kian-Meng Lim</i>	138
Computational Fluid Dynamics Modelling of Slurry Transport by Pipeline <i>Ming-zhi Li</i>	151

Recovering historical urban texture by parametric computing modeling" Historical Tehran as case study <i>Saeed Dolatkah</i>	168
Stochastic homogenization of polymeric composites with randomly ellipsoidal reinforcement <i>Damian Sokolowski, Marcin Kaminski</i>	184
Numerical simulation of crack propagation interacting with microdefects using adaptively refined XFEM <i>Zihao Teng, Dunming Liao</i>	198
A Semi-Lagrangian Method Based on Moving Particles and Its Application in Fluid Simulation of Casting Filling Process <i>Yulong Tang, Dunming Liao, Tao Chen, ShuYuan Fan</i>	211
A positional FEM formulation applied to 2D dynamic nonlinear analysis of structures and mechanisms with improved frictional internal sliding connections <i>Tiago Morkis Siqueira, Humberto Breves Coda</i>	219
Effect of soil properties on seismic response of underground station <i>Sunbin Liang, Zhiyi Chen</i>	238
Nomograph to calculate amount of reinforcing bar against bending moment in Circular Void slabs <i>Shigehiro Morooka</i>	248
Simulation of Smoke Flow in a Longitudinal Ventilated Tunnel in Macau <i>Hou Kuan Tam, Pak Hang Fu, Seng Kin Lao, Lap Mou Tam, Vai Kuong Sin</i>	255
A finite element based procedure for accurate determination of mode I SIF of orthotropic materials based on two parameter strain series <i>Debabrata Chakraborty, Debaleena Chakraborty, K. S. R. K. Murthy</i>	264
KBL: A knowledge based learning method for extracting formulas of aerodynamic heating <i>Changtong Luo, Zonglin Jiang</i>	270
A multi-level Method of Fundamental Solutions using quadtree-generated sources <i>Csaba Gaspar</i>	278
FEM non-linear modelling of cob using ANSYS <i>Alejandro Jimenez Rios, Dermot O'Dwyer</i>	285
Parametric Study on RPV Integrity Assessment under Pressurized thermal Shock <i>Jong-Wook Kim, Hanbum Suhr, Shinbeom Choi</i>	299
Spectral quasi-linearization method for entropy generation using the Cattaneo-Christov heat flux model <i>Precious Sibanda, Hiranmoy Mondal</i>	305
Homogenization of graphene sheet reinforced composites considering material and geometrical uncertainty <i>George Stefanou, Dimitrios Savvas</i>	322
Meshless Method with reduced integration - high performance <i>Wilber Velez, Tiago Oliveira, Elvis Pereira, Artur Portela</i>	332

Analysis of the results obtained from the application of the two-stage method with calculations of some statically indeterminate trusses <i>Janusz Rebielak</i>	347
Two-stage method applied in calculations for statically indeterminate truss of larger span <i>Janusz Rebielak</i>	356
Implementing Axisymmetric Smoothed Finite Method (S-FEM) Element in ABAQUS to Analyze defective Pressure Piping <i>S.H. Huo, Xin Cui, Y.H. Qie, Shuyong Duan</i>	364
An inverse method for identification of continuously varying material properties in post-manufactured structures through neural networks <i>Shuyong Duan, Zhanming Zhang</i>	368
An oscillation-free finite volume method with staggered grids for solving problems of poroelasticity <i>Clovis Raimiundo Maliska, Herminio Tasinafo Honorio</i>	372
SPH modelling of consolidation problem based on two-phase mixture theory <i>Jianhua Wang, Hao Wu, Jinjian Chen, Chencong Liao, Jian Wang</i>	386
CFD simulation of chemical gas dispersion under atmospheric boundary conditions <i>George Xiangguo Xu</i>	394
Numerical simulations of coupling effects in FGM plates by meshfree methods <i>Vladimir Sladek, Ladislav Sator, Miroslav Repka, Jan Sladek</i>	400
2016-2017 Central Italy Earthquake: Seismic Assessment of "Pietro Capuzi" School in Visso (Marche) <i>Chiara Ferrero, Paulo B. Lourenco, Chiara Calderini</i>	418
Rock mounted iconic lighthouses under extreme wave impacts: Limit Analysis and Discrete Element Method <i>Athanasios Pappas, Dina D'Ayala, Alessandro Antonini, Alison Raby</i>	435
Simulations of Dynamical Fracture of Concrete using Implicit Time Integration <i>Josef Kveton</i>	448
Computational Package for the Simulation of Plasma Microscopy Properties and Ion Beam-Plasma Interaction in High Energy Density Plasmas <i>Juan Miguel Gil</i>	454
Numerical simulation of metal-intermetallic laminate composites failure under dynamic loading <i>Sergey A. Zelepugin, Alexey S. Zelepugin, Alexey Popov, Dmitri Yanov</i>	467
Computational Methods in Architecture. From theories to architectural design <i>Anna Irene Del Monaco</i>	472
Numerical analyses of cement-based piezoelectric smart composites <i>Jan Sladek, Peter Bishay, Pavol Novak, Vladimir Sladek</i>	486
Wire rope model with elliptic cross sectional outer wires <i>Cengiz Erdonmez</i>	492

Damage assessment by Non-Smooth Contact Dynamics method of the iconic crumbling of the clock tower in Amatrice after 2016 central Italy seismic sequence <i>Marina Poiani, Francesco Clementi, Gabriele Milani, Valentina Gazzani, Stefano Lenci</i>	496
New computational algorithms for the Limit Analysis of large-scale 3D truss-frame structures <i>Egidio Rizzi, Rosalba Ferrari, Giuseppe Cocchetti</i>	506
Surface discretization based on bionic patterns in search of structural optimization <i>Ewelina Gawell, Anna Nowak, Wieslaw Rokicki, Anna Klara Stefanska</i>	518
A numerical study of thermal impact of forest fires on buildings <i>Valeriy Perminov</i>	527
Drag reduction of KCS based on extended FFD method and EGO <i>Wan Decheng, Aiqin Miao</i>	536
Fracture of alpha titanium alloys at high strain rates and stress triaxiality <i>Vladimir V. Skripnyak, Alexander A. Kozulyn, Evgeniya G. Skripnyak, Vladimir A. Skripnyak</i>	546
Numerical simulation of mechanical behavior of Zr-Nb alloys in a wide temperature range <i>Vladimir A. Skripnyak, Evgeniya G. Skripnyak, Vladimir V. Skripnyak, Vitas Serbenta, Natalia Skripnyak</i>	558
An optimal reconstruction of Chebyshev-Halley type methods with local convergence analysis <i>Ramandeep Behl, Ali Saleh Alshomrani, Ioannis K Argyros</i>	566
A Measure of Engineering Students' Pro-industry Behavior Adjust Industry 4.0 <i>Chun-Mei Chou, Chien-Hua Shen, Hsi-Chi Hsiao, Tsu- Chuan Shen</i>	586
A Novel Method Combining Reverse Modeling and Topological Optimization for Lightweight Design of Automobile Wheel Hubs with Hollow Ribs <i>S. Liu, Pengfei Xu, Shuyong Duan</i>	591
Shape identification of unsteady heat convection fields to control temperature distribution <i>Eiji Katamine, Naoya Okada</i>	595
Simulations of Thermal-Hydraulics Two-Phase Flows using Mixture Formulations <i>Dia Zeidan</i>	603
Numerical investigation of beam-column connections using a new multi-axial-spring model <i>Hu Qi, Haishan Guo, Kan Liu, Lida Tian, Jiao Geng</i>	608
Predicting Plaque Area Increase and Plaque Burden Increase Using Patient-Specific Fluid-Structure-Interaction Models Based on IVUS and OCT Images with Follow-Up <i>Dalin Tang, Xiaoya Guo, David Molony, Chun Yang, Habib Samady, Jie Zheng, Gary Mintz, Akiko Maehara, Jian Zhu, Genshan Ma, Don Giddens</i>	622
Failure modes analysis of a multi-story subway station under stochastic earthquake based on probability density evolution method <i>Zhiqian Liu, Zhiyi Chen</i>	629
Hydrodynamic characteristics of twin rudders <i>Linfeng Chen</i>	638
Winkler model for seismic response of shafts under stochastic earthquakes <i>Bu Zhang, Zhiyi Chen</i>	650

Optimization of Left Ventricle Pace Maker Location Using Echo-Based Fluid-Structure Interaction Models <i>Dalin Tang, Longling Fan, Jing Yao, Chun Yang, Di Xu</i>	661
Two parameters modelling of clay brick masonry confinement <i>Gian Piero Lignola, Giancarlo Ramaglia, Francesco Russo Spena, Andrea Prota</i>	668
Frictional contact analysis of functionally graded materials using smoothed finite element methods <i>Yufei Zhang, Junhong Yue, Ruiping Niu</i>	675
A sequential method in inverse estimation of the absorption coefficient for the spot laser welding process <i>Son Hoai Nguyen, Long Nhut Phi Nguyen, Quan Nguyen, Tin Trung Le</i>	681
Limit analysis of masonry structures based on fictitious associative-type contact interface laws <i>Giuseppe Quaranta, Francesco Trentadue</i>	693
Molecular dynamics simulation of the initiation of plastic deformation in nanocrystalline material <i>Takuya Uehara</i>	700
2-D inverse scattering analysis using pure SH wave for delamination in carbon fiber reinforced plastic <i>Sohichi Hirose</i>	706
Tsunami Run-Up Simulation Using Particle Method and its Visualization with Unity <i>Takahiro Saitoh, Gohki Noguchi, Takumi Inoue</i>	714
A class of novel tetrahedron elements with curved surfaces for three-dimensional solid mechanics problems with curved boundaries <i>Chunqiao Wang, Junhong Yue, Ming Li</i>	721
Multiscale Damage Modelling of Sustainable Composite <i>Raj Das, Shyam M. Panamoottil, Krishnan Jayaraman</i>	727
Cosmotic, Aquatic. Exploring the Potential of Computational Design in the Preservation of Aquatic ecotones. <i>Aya Mohanna</i>	741
The topological principles in the contemporary architectural design process <i>Maja Dragisic, Vladimir Lojanica</i>	752
Application of Uniform Design on Improvement Design of Detector Slides in Switch Machine system <i>Yung-Chang Cheng</i>	767
Numerical simulation of Two-dimensional risers under oscillatory flows with low Reynolds and KC for predicting the involve forces response <i>Maria Catalina Valencia-Cardenas, Carlos Alberto Riveros-Jerez</i>	776
An Optimal Eighth-order Family of Iterative Methods For Multiple Roots <i>Saima Akram, Fiza Zafar, Nusrat Yasmin</i>	787

Selective cell-based smoothed finite element method using 10-node tetrahedral elements for large deformation of nearly incompressible solids <i>Yuki Onishi</i>	804
An Evaluation of Hegemonic Epochs - A Time Series Analysis <i>Mihai Caramihai, Irina Severin</i>	815
An optimal iteration function for multiple zeros with eighth-order convergence <i>Ramandeep Behl, Ali Saleh Alshomrani</i>	822
Identification of Neumann boundary condition assuring the destruction of target region of biological tissue <i>Lukasz Turchan</i>	835
A novel approach for regulation of the diffusive effects of limiters in viscous-compressible-flow computations using a boundary-layer sensor <i>Anoop K Dass, Paragmoni Kalita</i>	847
The effects of solution techniques on the results of the simulation of human motion <i>Behzat Bahadir Kentel</i>	855
Seismic risk of masonry buildings: methods compared <i>Mariateresa Guadagnuolo</i>	859
An innovative Pedestrian-Bicycle Bridge Shape for Environmental Sustainability and Structural Efficient Improvement <i>Alberto Viskovic, Melania Lucci</i>	868
Ball convergence for a multi-step Harmonic mean Newton-like method in Banach space <i>Ali Saleh Alshomrani, Ramandeep Behl, I K Argyros</i>	874
Numerical simulation of natural convection from a heated cylinder <i>Gregor Kosec, Jure Slak</i>	887
Highly Accurate Smoothed Finite Element Methods Based on Simplified Eight-noded Hexahedron Elements <i>Yanhua Li, Ming Li</i>	897
A robust inversion-based Fourier Transformation algorithm used in the interpretation of non-equidistantly measured magnetic data <i>Mihaly Dobroka, Daniel Nuamah, Tamas Ormos, Judit Somogyi Molnar</i>	903
Solving the singular Motz problem using Radial Basis functions <i>T. S. Li, S. M. Wong</i>	908
Application of Parametric Modeling in the Early Design Phase for an Interdisciplinary Design Approach for Adaptive Buildings <i>Frederik Ernst</i>	917
Comparison of Characteristic-wise WENO and Central Difference Schemes with Numerical Viscosity Models for the Unsteady Compressible Flow <i>Se-Myong Chang, Hyun-Jin Kwon</i>	924
Timber shear walls: numerical assessment of the equivalent viscous damping <i>Giorgia Di Gangi, Cristoforo Demartino, Giuseppe Quaranta, Marco Vailati, Giorgio Monti, Marc'Antonio Liotta</i>	929

Computation of Deformable Image Registration by Weighted Kernel-based Meshless Method <i>S. M. Wong, K. S. NG and T. S. Li</i>	939
Approximation of the parallel robot working area using the method of nonuniform covering <i>Mikhail Posypkin, Larisa Rybak, Dmitry Malyshev, and Elena Gaponenko</i>	949
Numerical Investigation of Vortex-induced Vibration (VIV) of a Flexible Cylinder in Combined Flow <i>Di Deng, Zhe Wang, Decheng Wan</i>	961
Numerical Analysis on Two Floating Offshore Wind Turbines with Different Layouts <i>Yang Huang, Ping Cheng and Decheng Wan</i>	973
Numerical Calculations for Smooth Circular Cylinder Flow at 3900 Reynolds Numbers with SST-IDDES Turbulence Model <i>Jiawei HE, Weiwen ZHAO, Decheng WAN</i>	989
Numerical Simulation of Ship Bow Wave Breaking using DES and RANS <i>Zhen Ren, Jianhua Wang and Decheng Wan</i>	1001
Numerical Study of Bubble Coalescence by Multiphase MPS Method <i>Xiao Wen, Decheng Wan</i>	1013
Authors Index	1024

Study of Error based on Six Segments Ring Pair Electrical Resistance Sensor

WANG Xiao-na¹, DING Yi¹, SONG Shi-de², HUANG Yi³

(1. School of Physics and Optoelectronic Engineering, Dalian University of Technology, Dalian 116024, China;

2. School of Civil Engineering, Dalian University of Technology, Dalian 116024, China; 3. School of Naval Architecture and Ocean Engineering, Dalian University of Technology, Dalian 116024)

Abstract

There are some defects in the classical theory of ring pair electrical resistance sensor. Aiming at describing the error introduced by the input and output ports, the paper explains the reason, shows the simulation results and gives some suggestions to reduce the error. Then, comparing with the linear arc resistance values, the exact values are calculated using the nonlinear equation. Finally, Considering the temperature difference between the inner and outer walls, the nonlinear result is shown in the follow, and it is consistent with the results of the linear theory, that is, the pipe wall temperature difference does not affect the measurement of corrosion depth.

Keywords: **Corrosion; Ring pair electrical resistance sensor; Error; Resistance measurement**

1. Introduction

In recent years, with the depletion of petroleum resources on land, offshore oil exploration and development has gradually started to implement. Under the marine special environmental conditions, the pipeline corrosion monitoring is particularly important. At present, in the field of pipeline corrosion monitoring sensor technology, a variety of new excellent sensors began to be put into use.

Ring pair electrical resistance sensor (RPERS) is a new type of resistance sensor. Compared with the traditional resistance sensor, its inner and outer diameters is consistent with the pipeline, so it can be well integrated into the pipeline system. In addition, As the RPERS has no protrusion in the pipeline, it is conducive to the transport of oil and the operation of the pig. What's more, the sensor is in the same condition with the pipe wall, so the corrosion measurement depth is highly consistent with the corrosion depth of the pipe wall.

Huang Yi^[1] has successfully developed a set of corrosion monitoring system by using RPERS. Compared with the traditional system, the experimental results have better temperature compensation and can also monitor the local corrosion. However, the arc resistance and the wall temperature distribution use the approximate calculation method, so there may be some theoretical error in the calculation part. This paper focuses on the theoretical error caused by input/output ports and nonlinear effect, some suggestions are shown after simulation experiments.

2. Sensor theory

2.1 The position and shape of RPERS in pipeline

As shown in Fig.1, the RPERS consists of two metal ring sensors with the same inner diameter and material as the pipeline. The two rings are all coated with chromium oxide except the internal face of corrosion ring. The corrosion ring and the compensation ring are embedded into the pipeline through a connecting ring, which is shown in Fig.1.b. These rings are insulated from neighbor rings by coating of Cr_2O_3 , the insulation resistance is higher than $1\text{M}\Omega$.

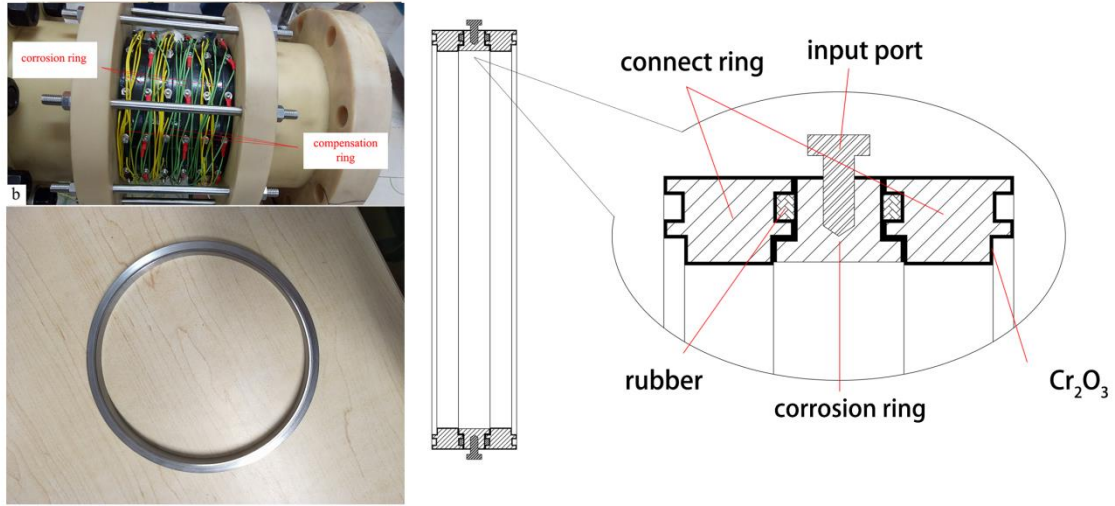


Fig.1 The structure of RPERS

2.2 The principle of Kelvin four-wire method (KFWM)^[2]

Micro-resistance is generally measured by KFWM to reduce the impact of wire resistance and contact resistance^[3]. as shown in Fig.2, the traditional RPERS corrosion monitoring system using a similar KFWM to measure the resistance.

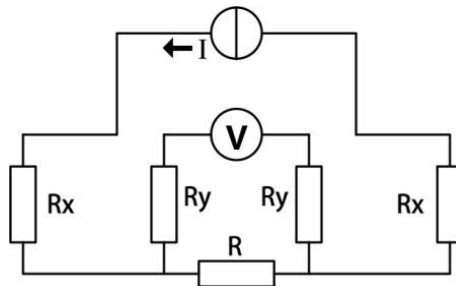


Fig.2 The schematic diagram of KFWM

In Fig. 2, R_x and R_y are the sum of the wire resistance and the contact resistance of each branch. Since R is a micro-resistance, the voltage drop and current through R_y can be ignored, the current $I_R \approx I$, and the voltmeter value $U_v \approx U_R$. By Ohm's law.

$$R = \frac{U_v}{I} \quad (1)$$

2.3 The measurement principle of RPERS^[4]

When I_1 is injected as the red path, the voltage value across the segment 2, 3, 5 and 6 in corrosion ring can be measured by using micro ohm meter RM3545. Segment 2 and 3 have the same current, so the resistance ratio $R_2:R_3$ can be obtained. The segment 5 and 6 have the same current, so the resistance ratio $R_5:R_6$ can be obtained too. In compensation ring, the resistance ratio $r_2:r_3$, $r_5:r_6$ can be obtained in the same way. Similarly, when I_2 is injected as the black path, the resistance ratio $R_3:R_4$, $R_1:R_6$, $r_3:r_4$ and $r_1:r_6$ can be calculated. When I_3 is injected as the blue path, the resistance ratio $R_4:R_5$, $r_4:r_5$, $R_1:R_2$ and $r_1:r_2$ can be calculated. Finally, calculated from the measurement value above, the resistance ratio $R_i:R_k$ and $r_i:r_k$ between the random segments in each ring can be obtained, where $i=1,2,\dots,6$, $k=1,2,\dots,6$. The corrosion ring and the compensation ring are connected in series, so their total current is equal, I_1 can be expressed as:

$$\frac{U_2}{R_2} + \frac{U_6}{R_6} = \frac{u_2}{r_2} + \frac{u_6}{r_6} = I_1 \quad (2)$$

where U_2 , U_6 represent the voltage value of segment 2, segment 6 in corrosion ring, and u_2 , u_6 represent the voltage value of segment 2, segment 6 in compensation ring. In summary, it is easy to get the resistance ratio $R_i:r_i$ between the corresponding segments of the two rings.

The cross-section parameter of RPERS is shown in Fig.3.

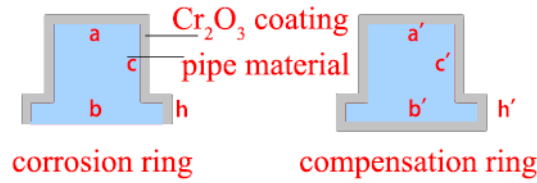


Fig.3 Cross-section parameter

If the inner diameter is much larger than the difference between the inner and outer diameters, Arc resistors can be approximately treated as linear resistors. According to the Fig.3, the resistance R_i and r_i can be expressed as:

$$R_i = \frac{L_i \rho(T)}{a_i c_i + b_i h_i} \quad (3)$$

$$r_i = \frac{L'_i \rho(T)}{a'_i c'_i + b'_i h'_i} \quad (4)$$

Where L_i and L'_i represent length of the materials, $\rho(T)$ represent resistivity. The corrosion ring and the compensation ring have the same $\rho(T)$, if there is no corrosion, the resistance ratio between the corresponding segments of the two rings can be expressed as:

$$\frac{R_i}{r_i} = \frac{L_i}{L'_i} \times \frac{a'_i c'_i + b'_i h'_i}{a_i c_i + b_i h_i} \quad (5)$$

If the corrosion depth is Δh_i , the resistance ratio can be expressed as:

$$\frac{R'_i}{r'_i} = \frac{L_i}{L'_i} \times \frac{a'_i c'_i + b'_i h'_i}{a_i c_i + b_i (h_i - \Delta h_i)} \quad (6)$$

Combining Eqs.(5),(6), the relation between the corrosion depth and the resistance value can be obtained:

$$\Delta h_i = \left(\frac{a_i c_i}{b_i} + h_i \right) \left(1 - \frac{R_i r'_i}{r_i R'_i} \right) \quad (7)$$

3. Theoretical error analysis

3.1 The error caused by input and output ports.

As shown in Fig.4, the resistance of the RPERS is measured by the KFWM to calculate the resistance ratio of each segment. In actual operation, the current flows from the port a to the port d. the potential of the surface C and surface B can be substituted by the potential of the port c and port b, only if the surface B and C are equipotential surfaces.

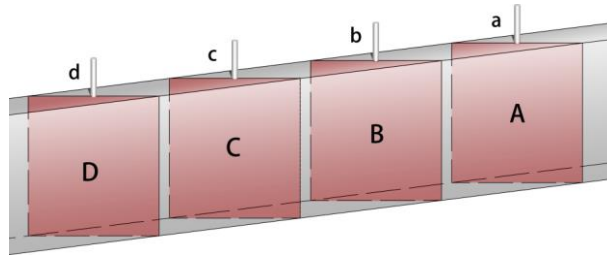


Fig.4 KFWM in actual operation

As shown in Fig.5, the current flows into the port a is approximately the point current flows into the half infinite plane, so the equipotential surfaces close to the port are hemispherical.

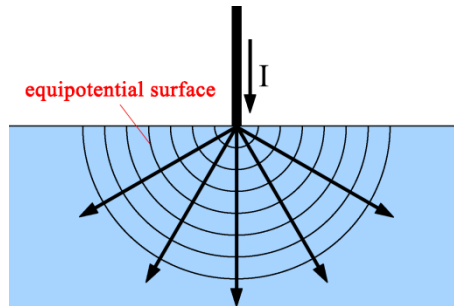


Fig.5 The equipotential surfaces close to the port

Thus, when the cross section B is too close to the cross section A, the cross section B will

be affected by the port A and become a non-equipotential surface, which will introduce the error of the potential measurement.

3.2 The error caused by resistance calculation

For the traditional RPERS, when the inside diameter of the sensor is much larger than the difference between the inside and outside diameter, the theoretical calculation of the resistance value is approximately calculated by using the linear resistance Eq.8. The relationship between L and R is linear.

$$R = \rho \frac{L}{S} \quad (8)$$

At present, the ratio between the ring inner diameter and the sensor thickness is generally about 10:1 or at the same order of magnitude, so the theoretical calculation error of the resistance is not negligible, i.e. the relationship between L and R is no longer linear.

3.2.1 Theoretical calculation of arc resistance^[5-6]

According to the classical physics theory, the arc resistance is calculated by calculus. As shown in Figure.6, the current flows in from the surface A steadily and flows out from the surface B.

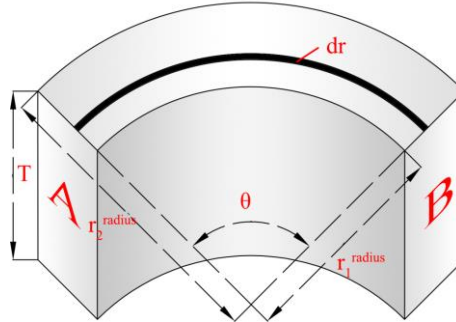


Fig.6 Resistor parameter diagram

In the picture above, the resistance value dR can be calculated as

$$dR = \rho \frac{\theta r}{T dr} \quad (9)$$

dR is connected with each other in parallel, so R can be calculated as

$$R = \frac{1}{\int_{r_1}^{r_2} \frac{T dr}{\rho \theta r}} = \frac{\rho \theta}{T \ln \left(\frac{r_2}{r_1} \right)} \quad (10)$$

As shown in Fig.3, after substituting the geometric parameter, the resistance value of segment i can be calculated as

$$R_i = \frac{\rho\theta_i}{b_i \ln\left(\frac{r_f^{\text{radius}} + h_i}{r_f^{\text{radius}}}\right) + a_i \ln\left(\frac{r_f^{\text{radius}} + h_i + c_i}{r_f^{\text{radius}} + h_i}\right)} \quad (11)$$

where r_f^{radius} represent the inner diameter of RPERS. Let $\frac{R_i r_i'}{r_i R_i'}$ be M_i , according to the Eqs.5,6,7,11. The relationship between M_i and Δh_i can be calculated as

$$M_i = \frac{b_i \ln\left(\frac{r_f^{\text{radius}} + h_i}{r_f^{\text{radius}} + \Delta h_i}\right) + a_i \ln\left(\frac{r_f^{\text{radius}} + h_i + c_i}{r_f^{\text{radius}} + h_i}\right)}{b_i \ln\left(\frac{r_f^{\text{radius}} + h_i}{r_f^{\text{radius}}}\right) + a_i \ln\left(\frac{r_f^{\text{radius}} + h_i + c_i}{r_f^{\text{radius}} + h_i}\right)} \quad (12)$$

Combining Eqs.7 and Eqs.12, the non-linear relationship between M_i and Δh_i can be obtained, which is shown in Fig.7,8.

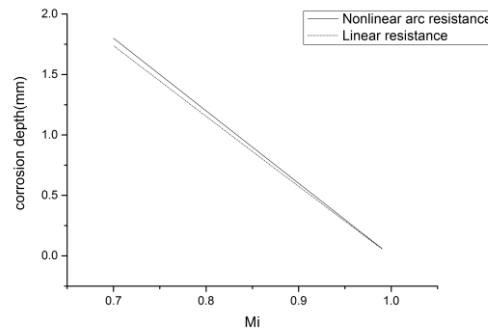


Fig.7 Relationship between linear theory and nonlinear theory

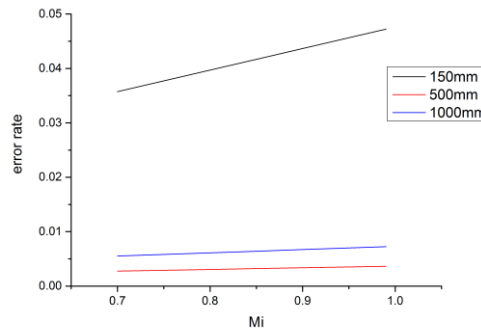


Fig.8 The error rate under different inner diameter conditions

According to Fig.7,8, it can be seen that the theoretical error of linear computation decreases with the increase of M_i , and the error rate between the linear and the nonlinear theoretical value increase with the increase of M_i . When the cross section of the RPERS is

constant, the error rate decreases with the increase of the diameter. When the diameter is 150mm, the average error rate reaches about 4%, this theoretical error cannot be ignored. When the diameter increases to 1m and 2m, the average error rate drops to 0.5% and 0.3%, and the error rate decreases slower as the change of M_i value. In Fig.3, Limited by the shape of the ring cross section, the range of Δh_i is from 0mm to 2mm. In this range, the nonlinear curve approximately shows a straight line. In order to reduce the calculation in the actual operation, the nonlinear curve can be approximately calculated as a linear line.

3.2.2 Influence of temperature difference between inner and outer pipe wall on ring resistance measurement.

When the arc resistance is approximately treated as linear resistance, the temperature distribution between the inner and outer pipe wall is linear, and the influence on the calculation of the corrosion depth will be counteracted by the compensation ring^[7-8]. In fact, the temperature distribution in the pipe wall is not linear. according to the Dirichlet boundary condition, it is calculated by the differential equation under cylindrical coordinates^[9], the relation between the temperature and the radius can be calculated as

$$t = t_2 - (t_2 - t_1) \frac{\ln\left(\frac{r}{r_1^{\text{radius}}}\right)}{\ln\left(\frac{r_2^{\text{radius}}}{r_1^{\text{radius}}}\right)} \quad (13)$$

Where t_2 represents the temperature of inner pipe wall, t_1 represents the temperature of outer pipe wall, r_1^{radius} represents the inner radius, r_2^{radius} represents the outer radius. The relationship between metal resistivity and temperature is linear, combining Eq.10 and Eq.13, the arc resistance can be written as

$$R = \frac{1}{\int_{r_1}^{r_2} \frac{Tdr}{\rho_0 \theta r \left\{ 1 + \alpha \left[t_2 - (t_2 - t_1) \frac{\ln\left(\frac{r}{r_1^{\text{radius}}}\right)}{\ln\left(\frac{r_2^{\text{radius}}}{r_1^{\text{radius}}}\right)} \right] \right\}}} \quad (14)$$

Where ρ_0 represents the resistivity of arc resistance material at 0°C, α represents the resistance coefficient of temperature. Through calculating, R can be written as

$$R = \frac{\rho_0 \theta \alpha (t_2 - t_1)}{T \ln\left(\frac{r_2^{\text{radius}}}{r_1^{\text{radius}}}\right) \ln\left(\frac{1 + \alpha t_2}{1 + \alpha t_1}\right)} \quad (15)$$

The new M_i can be calculated by using Eqs.15, and the result is consistent with Eqs.13. So we can come to the conclusion that, theoretically, the temperature difference between inner and outer pipe wall has no influence on the measurement of corrosion depth.

4. Simulation process and results

4.1 Simulation of error caused by input and output ports.

In order to investigate the error of potential measurement in the input and output ports of RPERS, finite element analysis software is applied to simulate the process of resistance measurement. The paper only analyses the input port because of the symmetry of the input and output ports. The sensor model is built in the software interface, the length of inner diameter is 150mm, the length of outer diameter is 168mm. As shown in Fig.3, the lengths of a, b, c and h are 8mm, 12mm, 7mm and 2mm respectively. A series of cross sections close to the input port are created to analyse the potential and current density distributions, which is shown in Fig.9.

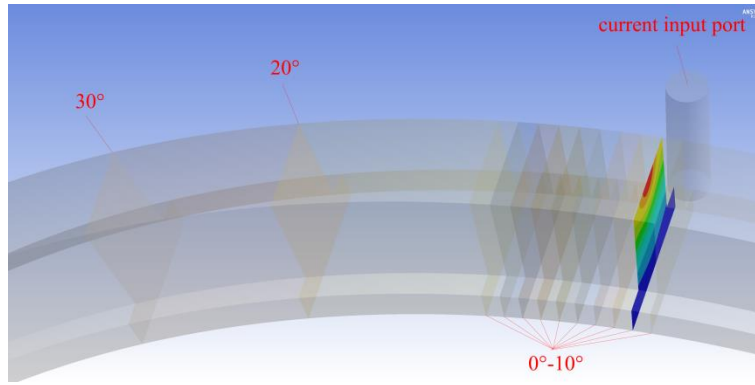


Fig.9 The location of cross sections

The simulation result is shown in Fig.10 and Fig.11.

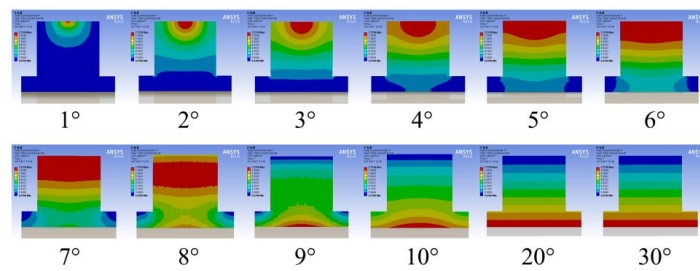


Fig.10 Distribution of current density in each section

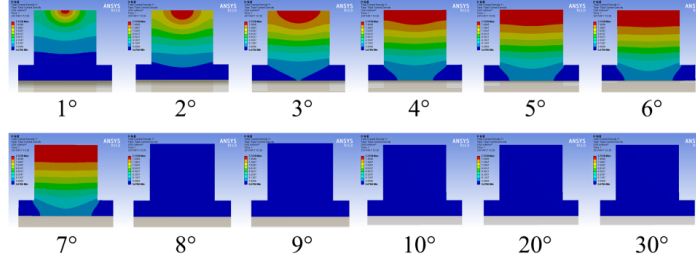


Fig.11 Potential distribution of each section

The simulation results show that, the potential distribution and current density distribution are affected by the input port in varying degrees. The current density distribution is approximately spherical in the cross section 1°, 2°, and 3°. With the increase of the angle between the input port and cross section, the current density gradually becomes more uniform. The location of the maximum current density transforms between the cross section 8° and 9°, the maximum value begins to appear in the bottom section, after 10° the current density distribution tends to stable and shows a gradient distribution map. The potential distribution changes faster than the current density. After cross section 7°, the simulation software is unable to distinguish the potential difference and the faces present equipotential surfaces.

According to the data given by Fig.12, the potential error rate of cross section 1° is about 1.5%, after 8 degrees the error rate between the simulation potential and the theoretical value is less than 0.0349%, which fully meets the requirements of measurement and calculation.

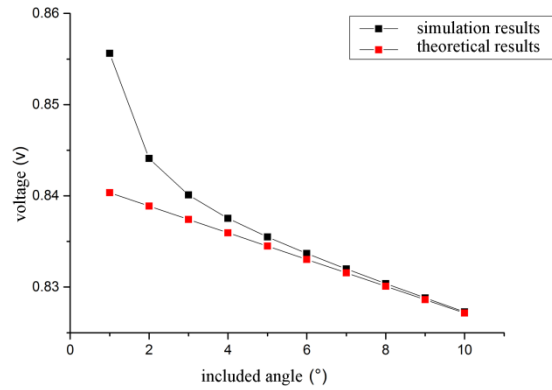


Fig.12 Simulation potential distribution and theoretical potential distribution

4.2 Simulation of error caused by temperature difference between inner and outer pipe wall

The volume parameters of the arc resistor are shown in the Fig.6, the element is made of pipeline steel. The center angle θ of the arc is 30°, the inner diameter r_1^{radius} is 60mm, the outer diameter r_2^{radius} is 110mm, and the thickness T is 30mm. The inner surface temperature is 200°C. The outer surface temperature is 0°C. All the surfaces except the inner and outer surfaces are set to adiabatic surfaces. The resistivity at 25°C is $17 \times 10^{-7} \Omega \cdot \text{m}$, the resistivity temperature coefficient is $5 \times 10^{-3} \text{R}/^\circ\text{C}$. The result is shown in the follow.

Table 1

Corrosion depth	The resistance at 25°C	The resistance at 200°C temperature difference	$M_i(25^\circ\text{C})$	$M_i(200^\circ\text{C}$ temperature difference)
0mm	$8.1369 \times 10^{-5} \Omega$	$1.0425 \times 10^{-4} \Omega$	1	1
1mm	$9.9212 \times 10^{-5} \Omega$	$1.2711 \times 10^{-4} \Omega$	0.81962	0.82015
2mm	$1.25950 \times 10^{-4} \Omega$	$1.6139 \times 10^{-4} \Omega$	0.64604	0.64595
3mm	$1.70498 \times 10^{-4} \Omega$	$2.1848 \times 10^{-4} \Omega$	0.47724	0.47716
4 mm	$2.59547 \times 10^{-4} \Omega$	$3.3303 \times 10^{-4} \Omega$	0.31350	0.31303

According to the result we can see that, the different temperature conditions lead to different resistance value, but the M_i value is identical.

5. Conclusions

Through theoretical calculation and simulation analysis, the following conclusions can be drawn:

(1) There is an error in the potential measurement near the current port. In order to reduce the influence, the potential measuring port needs to be away from the current port. the distance departing from the current port is determined by the conductor shape, size and current input mode. The simulation results show that, when the included angle between the current port and the potential measuring port is larger than 8° , the error is negligible, so in the classic RPERS theory included angle between the two ports is 30° , which fully meet the requirements of the engineering calculation.

(2) When the difference between the inner and outer diameter is far less than the inner diameter, the arc resistor can be treated as linear resistor, otherwise the error between the linear and arc resistance cannot be ignored. Based on the same cross section, the error rate increases with the inner diameter. When the temperature difference between the inner and outer pipeline walls exists, the linear and nonlinear calculation show consistent results, that is, the temperature difference between the inner and outer pipeline walls does not affect the measurement of corrosion depth.

References

- [1] Yunze Xu, Yi Huang, Xiaona Wang, Xuanqin Lin. Experimental study on pipeline internal corrosion based on a new kind of electrical resistance sensor[J]. Sensors and Actuators B: Chemical, 2016,224:37-47.
- [2] Dai Shengyue. The advantages and disadvantages of two wire method and four wire method for measuring resistance. [J]. Science and technology information, 2012 (34): 232-233.
- [3] Zhao Yingwei, Pang Kejian. Test technology and application of Kelvin four wire connection resistance [J].

- semiconductor technology, 2005, 30 (11): 43-45.
- [4] HEMBLADE B. Electrical resistance sensor and apparatus for monitoring corrosion : United States, 6946855 B1[P]. 2005-09-20.
- [5] Zhang Sanhui, University Physics [M]. Beijing: Tsinghua University press, 2009.
- [6] Ni Tao, Li Changjin. Comparison of half ring resistance calculated by infinitesimal method and mean value method [J]. west china education research 2012 (4): 84-85.
- [7] Xu Guoliang, Wang Xiaomo, Wu Tian Hua, et al. Engineering heat transfer [M]. Beijing: China Electric Power Press, 2005.
- [8] Huang Yi, Xu Yunze, Wang Xiaona, Li Xiaolong, Tu Shan. Research on pipeline internal wall corrosion monitoring technology based on ring pair resistance sensor [J]. Chinese Journal of mechanical engineering, 2015, (24): 15-23.
- [9] Baijing. Calculation and analysis of cylinder wall temperature [J]. Shanxi: Shanxi architecture, 2007 ,33(25):87.

Optimal design and error analysis based on pulsed eddy current sensor

Song Shi-de¹, Qu Tian-yang², Wang Xiao-na², Huang Yi³

1. School of Civil Engineering; 2. School of Physics and Optoelectronic Engineering; 3. School of Naval Architecture and Ocean Engineering; Dalian University of Technology, Dalian 116024, China

Abstract

Application and analysis method of pulsed eddy current is introduced. The parameters of sensors are optimized according to the detection requirement and suitable results and conclusions are obtained. The feature named attenuation rate is proposed to measure and calculate the thickness of steel plate. Error caused by lift-off effect and other factors is also analyzed.

Key words: **Pulsed eddy current; NDT; Finite element; Thickness measurement**

Introduction

With the rapid development of oil and gas transportation pipeline industry, the damage prevention and safety maintenance work of pipeline is of great significance. On the one hand, the HTHP fluid flowing through the pipe will cause corrosion on the inner wall. On the other hand, the protective coating with the function of thermal insulation will cause chemical corrosion to the external pipe wall [1] -[3] . Therefore, although the safest and most convenient way of offshore oil and gas transportation the submarine pipelines are, they are always taking the risk of thinning and damaging [4] -[9] .

The long service time and backward detection technology make the existing submarine pipelines leakage accidents easy to occur. In order to ensure the safety in operation, pulsed eddy current (PEC) testing technology is used to detect the pipeline. As a kind of new nondestructive testing technique, pulsed eddy current testing can be used to measure the corrosion status of pipelines without affecting the normal operation and destroying the protective layers of submarine pipelines [10] -[16] . And the research of PEC testing technology is of great value to ensure the safe and stable operation of submarine pipelines.

The sensor for PEC testing is composed of two parts: the driver coil and the magnetic field detector (pickup coil, Hall sensor, GMR, etc.) [17] -[21] . In inspection process, rectangular wave current is transmitted to the exciting coil by the drive circuit. The square wave field generated by the excitation signal in driver coil is the source magnetic field. The change of magnetic field intensity will excite the induced electric field in specimen and the eddy current is generated. Magnet field produced by eddy current and the source magnet field together breed the induced electromotive force. By analyzing the detection signal and extracting feature information, detection of thickness or defect is realized.

Theory

The analysis of the PEC theory requires understanding of the electromagnetic field and the analysis of the Maxwell equations. The wave equation which can be derived from the Maxwell equation is written as

$$\nabla^2 \dot{H} = j\omega\mu(\sigma + j\omega\varepsilon)\dot{H} \quad (1)$$

For general metallic material, $\varepsilon_0 = 8.85 \times 10^{-12} \text{ F/m}$, so $\omega\varepsilon$ is negligible compared with σ . Eq. (1) can be reduced to

$$\nabla^2 \dot{H} = j\omega\mu\sigma \dot{H} \quad (2)$$

Similarly, the following equation is derived as

$$\nabla^2 \dot{E} = j\omega\mu\sigma \dot{E} \quad (3)$$

$$\nabla^2 \dot{j} = j\omega\mu\sigma \dot{j} \quad (4)$$

where \dot{H} , \dot{E} and \dot{j} are the complex vector of magnetic field intensity, electric field intensity and current density respectively. These are the basic equations for the analysis of PEC phenomenon.

However, analytical solution is hardly obtained except for a few simple problems by solving the above equations. In order to research the electromagnetic phenomenon of PEC testing, the eddy current loop model is used to study the equivalent circuit. The specimen is modeled as series connections of inductors and resistors. The driver coil, pickup coil and specimen together are simplified as a multi-coil coupling system. Finally, the complex physical field problem is converted into the circuit theory calculation.

The circuit model is shown in Fig. 1. The driver coil in the PEC sensor can be equivalent to the primary coil of the multi-coil coupling system. And the pickup coil and eddy current ring can be equivalent to the secondary coil. The coils are coupled together by magnetic field, and the magnetic properties of the coils are described by self-inductance and mutual inductance.

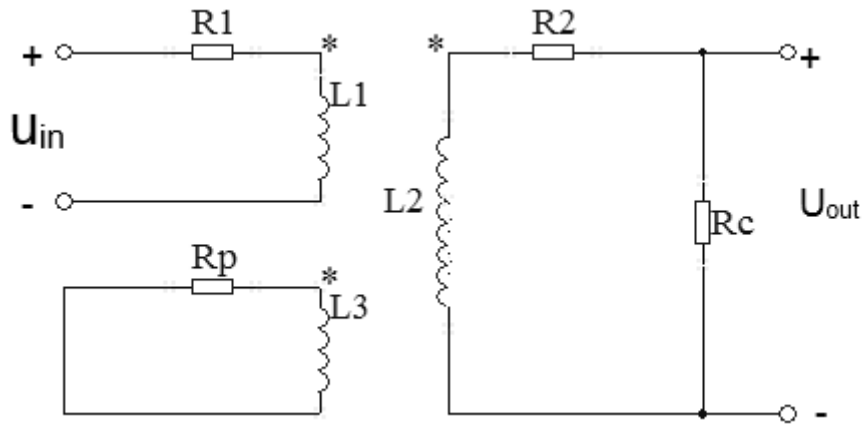


Fig. 1 Circuit model of PEC testing

The following equation is given according to the circuit theory

$$\begin{cases} u_1(t) = R_1 i_1(t) + L_1 \frac{di_1(t)}{dt} + M_{12} \frac{di_2(t)}{dt} + M_{13} \frac{di_3(t)}{dt} \\ u_{out}(t) = -R_c i_2(t) = R_2 i_2(t) + L_2 \frac{di_2(t)}{dt} + M_{12} \frac{di_1(t)}{dt} + M_{23} \frac{di_3(t)}{dt} \\ 0 = R_p i_3(t) + L_3 \frac{di_3(t)}{dt} + M_{13} \frac{di_1(t)}{dt} + M_{23} \frac{di_2(t)}{dt} \end{cases} \quad (5)$$

Where R_1 , R_2 and R_p represent resistance value of driver coil, pickup coil and eddy current loop, L_1 , L_2 and L_3 represent self-inductance respectively. M is the mutual inductance among the coils.

Although the circuit model is intuitive and convenient, too much approximation reduces the scope of application in the process of design and analysis.

Simulation optimization

In the process of analysis and calculation in PEC testing, a plenty of approximations are needed in modeling. The difficulties of obtaining analytical solution are great and the inaccuracy can't be neglected. Hence the numerical simulation model by finite element method is established in the time domain. It can simplify the calculation process and calculate the detection signal quantitatively. COMSOL is used as simulation software, which is practical with abundant CAD modeling tools, powerful meshing ability, a large number of physical modules and extended functions.

During pulsed eddy current testing, most of the time the diameter of the pipeline is obviously larger than that of the excitation coil. Therefore, the pipeline can be equivalent to a flat plate structure and the wall thickness of the pipeline is assumed as the thickness of a plate approximately.

The two-dimensional axisymmetric model is used to calculate the solution, which reduces the calculation time and lower the difficulty of analysis. Simulation has showed that the three-dimensional model has no obvious advantages compared with the two-dimensional axisymmetric model under the circumstance of a cylindrical coil.

The Maxwell equations and the following equations are applied in the calculation of simulation model.

$$(j\omega\sigma - \omega^2\varepsilon)A_\phi + \nabla \times (\mu^{-1} \nabla \times A_\phi) = J_\phi^e \quad (6)$$

$$J_\phi^e = \sigma\varepsilon = -(\nabla J V + \frac{\partial A_\phi}{\partial t}) \quad (7)$$

$$H = \frac{B}{\mu} = \frac{\nabla \times A_\phi}{\mu_0 \mu_r} \quad (8)$$

Where ω is angular frequency, σ is dielectric constant, μ is permeability, J_ϕ^e is current density.

In order to analyze the relationship among the structural parameters of the coil, the magnetic field around the coil and the eddy current in the test piece, the most important parameters of the coil are simulated and analyzed to obtain more accurate detection results. The height, inner diameter and outer diameter are considered as the main structural parameters of the coil. Only one parameter is changed at a time when variation of eddy current distribution is researched.

Fig. 2 shows the simulation results of the eddy current distribution in the test parts when inner and outer diameter of the coil are the same and the height of coil is 30mm, 20mm and 10mm respectively. It can be seen from the graph that the magnetic induction line will come close to

the coil with the decrease of coil height when the inner and outer diameter stay the same, which is conducive to the improvement of the resolution of the detection system.

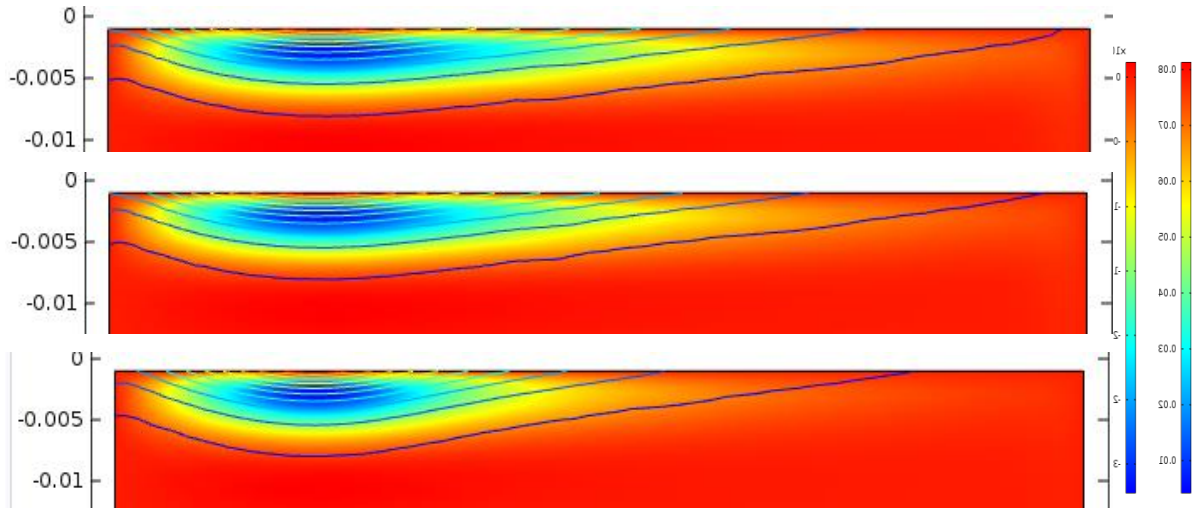


Fig. 2 Current distribution with different height of coil

Fig. 3 shows the simulation results of the eddy current distribution in the test parts when inner diameter and height of the coil are the same, outer diameter is 5mm, 15mm and 25mm respectively. It can be seen from the graph that the eddy current intensity will increase with the increase of outer diameter when the inner diameter and height of the coil stay the same, which is conducive to the improvement of the sensitivity of the detection system.

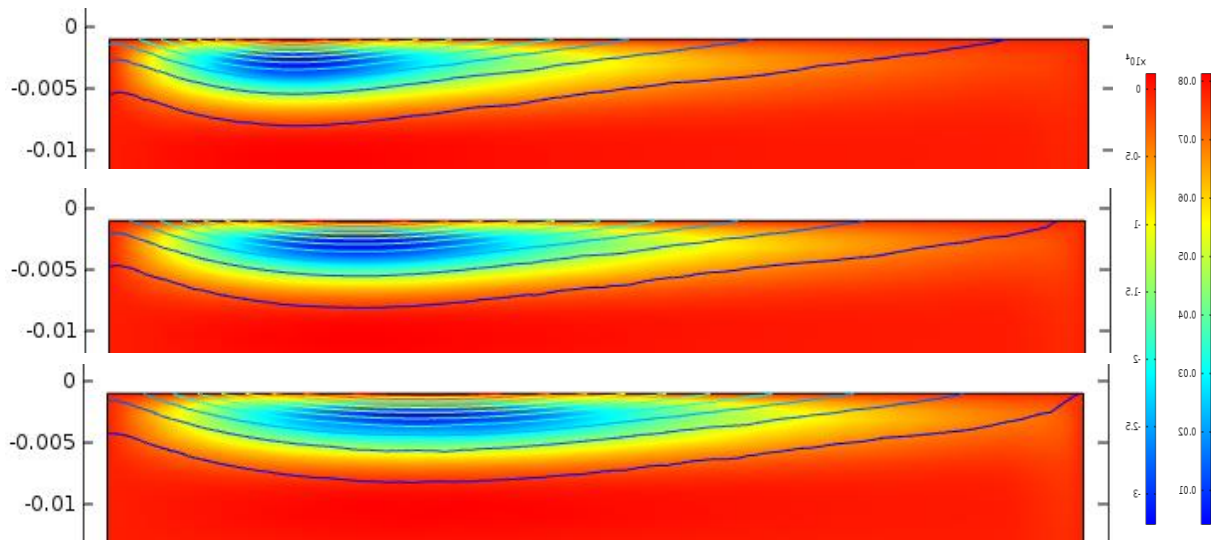


Fig. 3 Current distribution with different outer diameter of coil

Fig. 4 shows the simulation results of the eddy current distribution in the test parts when outer diameter and height of the coil are the same, inner diameter is 25mm, 15mm and 5mm respectively. It can be seen from the graph that the magnetic field distribution concentrate with the decrease of outer diameter but there is no obvious change in eddy current intensity. It indicates that the decrease of inner diameter is conducive to the improvement of the resolution of the detection system.

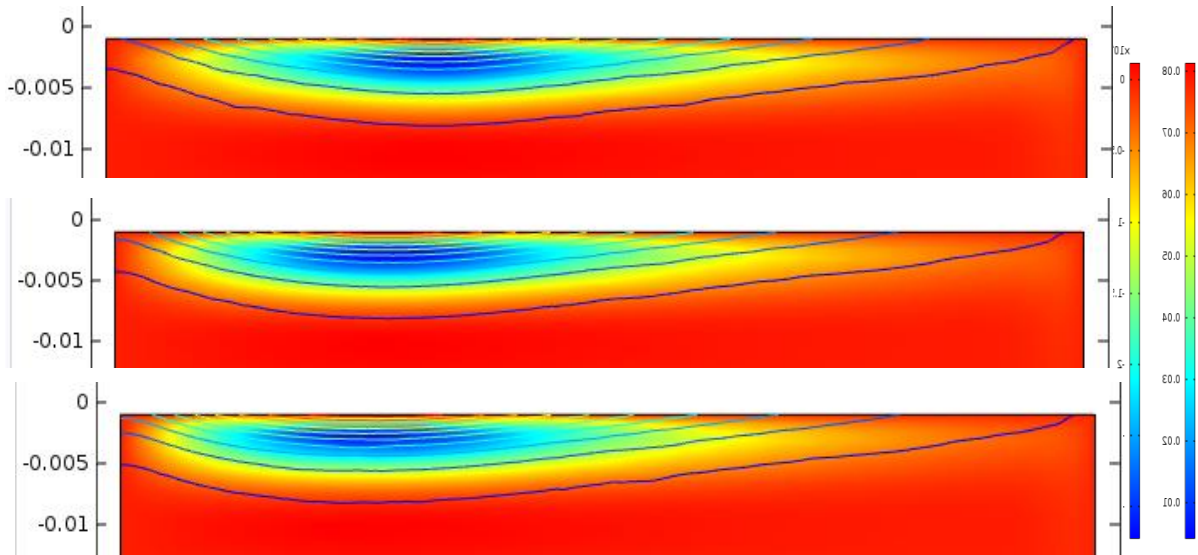


Fig. 4 Current distribution with different inner diameter of coil

It can be seen from the simulations that a relatively smaller inner diameter, larger outer diameter and lower height for the coil are beneficial to the improvement of the resolution and sensitivity of a PEC testing system. If necessary, an iron core can be added to reinforced the effect.

Error analysis

After optimizing the structure of the PEC sensor, steel plate with different thickness is simulated to find proper features from the detection signal. The thickness of the steel plate is in the range from 10mm to 25mm. Due to the large range of voltage variation in detection coil, the logarithm of voltage is taken. As shown in Fig. 5 plotted by time(s) on the horizontal axis and logarithm of voltage(V) on the vertical, the rear part of the induction voltage curve on semi log coordinate is approximate to right line. And the attenuation rate of the signal increase as the thickness of the steel plate decreases. Therefore, the thickness can be calculated by taking the attenuation rate of detection signal in the rear part as the feature.

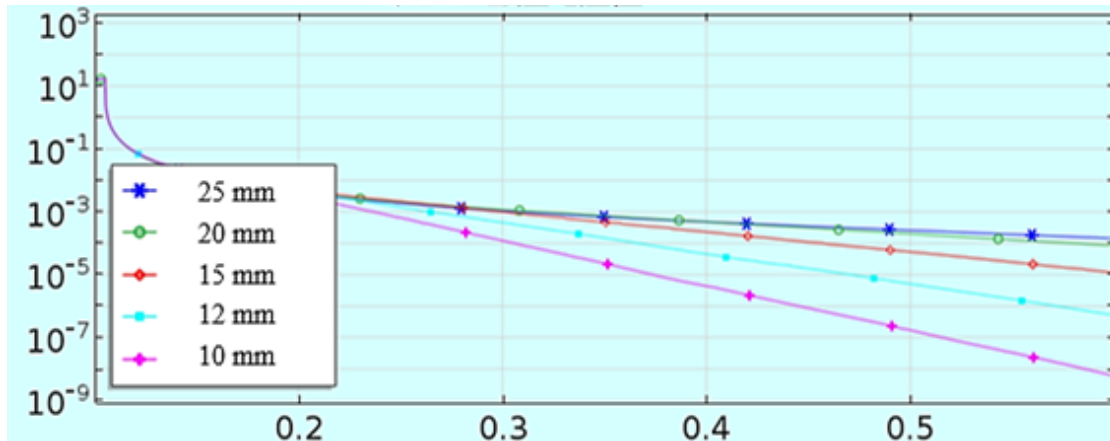


Fig. 5 Voltage curves in pickup coil for steel plates

According to the simulation results, a set of PEC testing system is designed and optimized to finish the thickness measurement. Thickness of each segment is executed for four times. Signal segments from 0.15V to 0.05V are captured to do linear fitting and the feature related to thickness, attenuation rate of the signal is calculated and averaged. Measurement results are as shown in Table 1, in which M1 to M4 represents the first to fourth measurement.

Table 1 Decay rate for measurement of steel plates with different thickness

Thickness (mm)	M1 (dB/s)	M2 (dB/s)	M3 (dB/s)	M4 (dB/s)	Average (dB/s)
9.5	-385.159	-384.046	-385.061	-390.297	-386.141
12.3	-257.593	-258.678	-258.521	-262.188	-259.245
14.9	-194.135	-191.178	-191.188	-194.907	-192.852
17.5	-151.232	-151.171	-152.706	-149.093	-151.0505
20.1	-109.259	-110.543	-111.276	-110.614	-110.423
22.5	-92.5183	-92.4155	-91.494	-92.4588	-92.22165
25.0	-84.9723	-84.057	-84.800	-84.551	-84.59507
27.6	-85.1229	-84.4194	-84.108	-82.743	-84.09833

Fitting with exponential function, the following equation can be obtained

$$y = -64.363 - 1670.85 \times e^{-0.17352x} \quad (9)$$

Where x is the thickness of steel plate, y is the attenuation rate.

According to the equation above, the thickness of measurement can be calculated. Compared with the actual thickness, relative error is computed as shown in Table 2

Table 2 Thickness measurements and relative errors

measurement	Actual thickness (mm)	decay rate (dB/s)	Thickness for measurement (mm)	relative error
1st	10.9	-306.004	11.14	2.19%
2nd	12.2	-267.289	12.15	0.41%
3rd	14.9	-198.577	14.53	2.53%
4th	16.2	-174.51	15.67	3.37%
5th	13.55	-227.587	13.4	1.08%
6th	21.3	-102.445	21.79	2.26%
7th	22.55	-96.44	22.78	1.01%
8th	25.05	-87.035	24.78	1.09%
9th	23.8	-89.4	24.2	1.69%
10th	26.3	-87.756	25.7	6.91%

It can be seen from the table that when the measured value is less than 25mm, relative error is less than 5%, The relative error increases when the thickness of the steel plate is a little bit larger.

Simulation proves that lift-off effect has little effect on measurement when attenuation rate is taken as a feature. However, actual measurement results shows that the influence of lift-off effect cannot be ignored. All the amplifying and filtering or noise in hardware circuit may make the lift-off effect more influential. Exponential fitting is not always the best method. Polynomial fit in different order or neural network algorithm may bring distinct precision results. Magnetic shielding acting on coils can make sensitivity improve.

Conclusion

In PEC testing system, a relatively smaller inner diameter, larger outer diameter and lower height for the coil bring about better resolution and sensitivity. Attenuation rate from rear part of the induction voltage curve can be used as a feature to measure thickness of steel plate. Thicker steel plate will lead to lower measurement accuracy. the influence of lift-off effect needs to be take into consideration to reduce errors.

Reference

- [1] Zhang Q, Wu X and Li J. An analytical method for pulsed eddy current testing of the steel plate with a flat-bottom hole.
- [2] Huang C and Wu X. An improved ferromagnetic material pulsed eddy current testing signal processing method based on numerical cumulative integration. *NDT&E INT* 2015, 69: 35-39.
- [3] Xu Z, Wu X, Huang C and Ke H. Analysis of excitation current parameters in pulsed eddy current testing for ferromagnetic metallic materials.
- [4] Xu Z, Wu X, Li J and Kang Y. Assessment of wall thinning in insulated ferromagnetic pipes using the time-to-peak of differential pulsed eddy-current testing signals. *NDT&E INT* 2012, 51: 24-29.
- [5] Huang C, Wu X, Xu Z and Kang Y. Ferromagnetic material pulsed eddy current testing signal modeling by equivalent multiple-coil-coupling approach.
- [6] Li J, Wu X, Zhang Q and Sun P. Measurement of lift-off using the relative variation of magnetic flux in pulsed eddy current testing.
- [7] Xu ZY, Wu XJ, Huang C and Kang YH. Measurement of Wall Thinning through Insulation with Ferromagnetic Cladding Using Pulsed Eddy Current Testing.
- [8] Chen H, Wu X, Xu Z and Kang Y. Pulsed eddy current signal processing method for signal denoising in ferromagnetic plate testing.
- [9] Li J, Wu X, Zhang Q and Sun P. Pulsed eddy current testing of ferromagnetic specimen based on variable pulse width excitation. *NDT&E INT* 2015, 69: 28-34.
- [10] Tian GY and Sophian A. Defect classification using a new feature for pulsed eddy current sensors.
- [11] Li Y, Tian GY and Simm A. Fast analytical modelling for pulsed eddy current evaluation.
- [12] He Y, Tian G, Pan M and Chen D. Impact evaluation in carbon fiber reinforced plastic (CFRP) laminates using eddy current pulsed thermography. *COMPOS STRUCT* 2014, 109: 1-7.
- [13] Ludwig R and Dai XW. Numerical and analytical modeling of pulsed eddy currents in a conducting half-space.
- [14] He Y, Luo F, Pan M, Weng F and Hu X. Pulsed eddy current technique for defect detection in aircraft riveted structures.
- [15] Yang HC and Tai CC. Pulsed eddy-current measurement of a conducting coating on a magnetic metal plate.
- [16] Tian GY and Sophian A. Reduction of lift-off effects for pulsed eddy current NDT. *NDT&E INT* 2005, 38: 319-324.
- [17] He Y, Pan M, Luo F and Tian G. Reduction of Lift-Off Effects in Pulsed Eddy Current for Defect Classification.
- [18] Nen KK, Kinen JM, Korhonen P, Juntunen E and Heikkinen V. *Sensors and Actuators A: Physical*.
- [19] He Y, Tian G, Zhang H, Alamin M and Simm A. Steel Corrosion Characterization Using Pulsed Eddy Current Systems.
- [20] Tian GY, Li Y and Mandache C. Study of Lift-Off Invariance for Pulsed Eddy-Current Signals.
- [21] Tian GY, Sophian A, Taylor D and Rudlin J. Wavelet-based PCA defect classification and quantification for pulsed eddy current NDT.

Mode shapes complexity for damage identification of structures experiencing plasticization

†*Fabrizio Iezzi¹ and Claudio Valente¹

¹Department of Engineering and Geology, University “G. d’Annunzio” of Chieti-Pescara, Italy

*Presenting author: fabrizio.iezzi@unich.it

†Corresponding author: fabrizio.iezzi@unich.it

Abstract

Classically, for structures prone to earthquakes, the damage identification techniques rely on changes of modal parameters between different structural states. Recently, the effects of energy dissipation have been considered as possible alternatives to modal parameters for damage identification. In particular, it is assumed that energy dissipation causes an increase of damping non-proportionality and, consequently, an increase of complexity in the mode shapes. A number of indices have been proposed to measure the damping non-proportionality or complexity in the mode shapes. To be successful in damage detection such indices should possess at least two characteristics: monotony and sensitivity. The work aims to investigate the effectiveness of the indices for structural damage identification. To this end, numerical simulations concerning a plane frame structural model are carried out. The damage is such to produce plastic hinges at selected joints of the model where the energy dissipation is concentrated. Seismic type base motion of progressive increasing intensity is considered to show the relation between damage severity, energy dissipation, damping non-proportionality and mode shapes complexity. Signal processing time domain techniques, EMD and CPR, are applied to the structural dynamic response in order to identify the indices based on the complex mode shapes. The indices are then applied to detect the damage.

Keywords: Damage identification; Non-proportional damping; Mode shapes complexity; Damage indices.

Introduction

The recent and numerous seismic events that worldwide hit the existing buildings have made very timely and important the development of new techniques for the structural damage identification. In the seismic field, the analysis of the dynamic behavior of a structure can be used to identify the damage occurred. When the damaged structure preserves a quasi-linear behavior, it can be retained that its modal parameters (frequencies, mode shapes and damping) are function of the physical properties (mass, stiffness and energy dissipation). Changes in these latter properties, caused by damage, are reflected, therefore, in modal parameters changes. The comparison of the values attained by the frequencies, mode shapes and damping, or their appropriate function, between different structural states is an effective means to estimate the damage. There exist a number of techniques based on changes in modal parameters and targeted to damage assessment [1]. In general, the frequencies are sensitive to the damage but hardly allow its localization [2]. On the contrary, the modal shapes are identifiable with greater difficulty but are extremely effective to localize the damage [3]. The damping presents the same characteristics of the frequencies but it is more difficult to identify exactly [4][5][6].

Recently, the damping has been used in conjunction with the effects induced on the mode shapes in energy dissipating structures [7][8]. The basic hypothesis is to associate the damage to the energy dissipated during the structural vibrations and to measure the effects generated on the mode shapes in terms of modal complexity. In particular, it is assumed that: the greater the damage is, the greater the energy dissipation is and consequently, the more damped the dynamic response is, the higher the loss of damping proportionality is and so the more complex the mode shapes are. The measure of this non-proportionality (or complexity) is used to estimate the structural damage. A number of indices have been proposed and analyzed to provide an effective measure of non-proportionality [9]. A group of indices requires the knowledge of the damping matrix of the structure; whereas another group of indices requires the knowledge of the mode shapes. In consideration of the difficulty of the experimental identification of the damping matrix, the first group has only a theoretical value. The practical applicability is therefore restricted to the second group of indices. However, to get successful indices, two properties should be fulfilled: monotony and sensitivity both related to uniqueness aspects. In this contingency, it is possible to perform a reliable identification based on the comparison of the values attained by the indices between two different states of a structure.

The work aims to examine the effectiveness of the indices based on the mode shapes for damage identification purposes. To this end, numerical simulations are carried out. The pseudo-experimental data (i.e. the dynamic responses) are generated using a plane frame model endowed with plastic hinges localized in selected joints of the structural model where the energy dissipation, i.e. the damage, is concentrated. The damage is caused by seismic type base motion. The intensity of the base motion is progressively raised in order to increase the damage severity. Each intensity level produce a different state of the structural model. The structural response at each state is processed to identify the mode shapes according to the assumption of quasi-linear behavior of the structure. Then the indices (of the second group) are computed and compared between the different states. In this way it is possible to show the relation between damage severity, hysteretic energy dissipation, damping non-proportionality and mode shapes complexity.

Plasticity and modal complexity

The discrete form of the motion of a structure with linear behavior endowed with viscous damping has equation:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{P}(t) \quad (1)$$

in which t is the time variable; \mathbf{x} , $\dot{\mathbf{x}}$ and $\ddot{\mathbf{x}}$ are, respectively, the displacement, velocity and acceleration vectors; in turn \mathbf{M} , \mathbf{C} and \mathbf{K} are the mass, damping and stiffness matrices; \mathbf{P} is the load vector.

Even though the matrices \mathbf{M} and \mathbf{K} are diagonalizable, \mathbf{C} can be diagonalizable or not depending whether the damping is proportional or non-proportional (i.e. \mathbf{C} is a combination or not of \mathbf{M} and \mathbf{K}). In the first case, the mode shapes are real and have components with equal phase; in the second case, the mode shapes are complex and have the components with different phase [10]. Further, the more \mathbf{C} is non-proportional, the more the imaginary part of the mode shapes is high, that is to say the complexity of the mode shapes increases along with the increase of the \mathbf{C} non-proportionality [9][11].

In framed structures, the damage is often confined in the beam-column joints where local plasticity occurs accompanied by hysteretic type energy dissipation. In these conditions, the discrete form of the equation of motion Eq. (1) takes the form:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{H}[\mathbf{x}(t)] = \mathbf{P}(t) \quad (2)$$

where $\mathbf{H}(\mathbf{x}) = d\mathbf{F}(\mathbf{x})/d(\mathbf{x})$ is the matrix of the instantaneous non-linear stiffness that depends on the reaction force $\mathbf{F}(\mathbf{x})$. The increase of plasticity (i.e. of damage) implies the stiffness reduction and the energy dissipation increase. These two effects are taken simultaneously into account in Eq. (2) through $\mathbf{H}(\mathbf{x})$ that represents their combined effect. Both effects contribute individually to make non-proportional the damping and, hence, to make complex the mode shapes. Actually, the dominant contribution is that due to the energy dissipation [7], therefore the contribution due to the stiffness is not considered in this work.

Modal complexity indices for structural damage identification

Among the indices based on the knowledge of the mode shapes, collected in [9], and for coherence with those studied in [7][8][11], in the present work five indices are considered: modal imaginary ratio (I_1), modal collinearity (I_2), modal dispersity (I_3), modal phase difference (I_4) and modal polygon area (I_5). Their formulation is given below:

$$I_1 = \frac{\|\text{Im}(\boldsymbol{\Psi}_k)\|}{\|\boldsymbol{\Psi}_k\|}; \quad I_2 = 1 - \frac{|\text{Re}(\boldsymbol{\Psi}_k)^T \text{Im}(\boldsymbol{\Psi}_k)|}{\sqrt{(\text{Re}(\boldsymbol{\Psi}_k)^T \text{Re}(\boldsymbol{\Psi}_k))(\text{Im}(\boldsymbol{\Psi}_k)^T \text{Im}(\boldsymbol{\Psi}_k))}}; \quad I_3 = \frac{\sum_{j=1}^N |\text{Im}(\psi_{kj})|}{N};$$

$$I_4 = \frac{|\phi_{k,max}| - |\phi_{k,min}|}{\pi}; \quad I_5 = \frac{A_k}{A_{k,max}}; \quad (3)$$

in which N is the number of the degrees of freedom (dofs) of the structures, $\boldsymbol{\Psi}_k$ is the k -th mode shape, ψ_{kj} is the j -th component of $\boldsymbol{\Psi}_k$; $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ stand for the real and imaginary part of the quantity in (\cdot) ; $\phi_{k,max}$ and $\phi_{k,min}$ are, respectively, the maximum and minimum phase angle of $\boldsymbol{\Psi}_k$; A_k and $A_{k,max}$ are respectively the modal polygon area of $\boldsymbol{\Psi}_k$ and its maximum value; $\|\cdot\|$ is the Euclidean 2-norm operator and $|\cdot|$ the componentwise absolute value. The meaning of the indices is as follows: I_1 weighs the importance of the imaginary part with respect to the overall length of the complex mode shape. I_2 measures the degree of interdependence of the real and imaginary parts of a complex mode shape and is directly affected by the damping proportionality: the higher the damping proportionality is, the more correlated the real and imaginary parts of the mode shapes are; if the imaginary part of a complex mode shape is completely dependent on the real parts, the damping is proportional. I_3 measures the degree of the scatter of the complex mode shape that is directly related to the amplitude of the imaginary part. The idea behind the I_4 and I_5 relies on a geometric interpretation. If the components of the mode shape are plotted in the complex plane, the effects of the non-proportional damping become apparent. I_4 considers the phase differences between the dofs of a mode shape as a consequence of damping non-proportionality. In fact, each component of a mode shape of a system endowed with proportional damping lies on a straight line, whereas those of a system endowed with non-proportional damping do not; in effect, these latter exhibit an angular dispersion equivalent to the phase differences. If the individual components of a mode shape are connected by straight lines, an N -side polygon is formed. If the damping is proportional, the components of a mode shape lie on a straight line and the polygon area is zero. As the non-proportionality of the damping increases, the area of

this polygon also increases. I_5 measures the area of this polygon. In practical terms, once the mode shapes are identified they are first normalized using the procedure proposed in [12] that ensures the minimization of the errors in the identification of the imaginary part of the mode shape [8]. Subsequently the normalized mode shapes are used to compute the indices of Eq. (3). Finally, the indices, scaled in the interval $[0; 1]$ and expressed in percentage, are analyzed to infer the damage presence.

Case study and methodology

The effectiveness of the five modal complexity indices of Eq. (3) to identify the damage in structures experiencing plasticization is analyzed using as pseudo-experimental data generator a basic model of framed structure with the scheme of “strong beam – weak column”.

The reference structure is a plane frame with a single span and three levels. The inter-storey height is 3 m and a mass of 10 t is condensed at each level. The columns have a constant cross-section 0,30 by 0,30 m and a Young modulus equal to $3 \cdot 10^7$ kN/m², so the inter-storey stiffness is $1,8 \cdot 10^4$ kN/m. The damage is simulated by means of plastic hinges that dissipate energy for hysteresis and are localized at the columns base (Fig. 1).

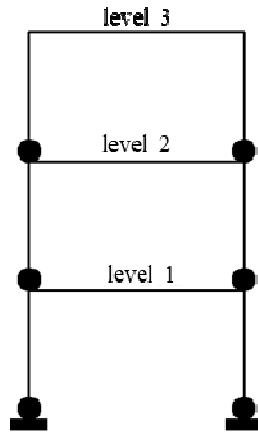


Figure 1. Reference framed structure with simulated damage using plastic hinges localized at the columns base

The excitation of the structure is a seismic type base motion. Seismic input in resonance conditions (ideal) and not (actual) are considered. The “resonance conditions” case corresponds to mono-harmonic input with the excitation frequency equal to any of the structural frequencies. This case allows to obtain output signals typical of the experimental modal analysis in real structures. The “actual conditions” case corresponds to a real earthquake input. The energy dissipation entity of the plastic hinges (i.e. of the damage) is controlled by a damage parameter function of the excitation amplitude: the higher the excitation amplitude is, the higher the value of the damage parameter is.

In order to obtain mono-harmonic output signals of the structural dynamic response, the excitation is endowed with the same natural frequency of the structure to excite, according to the mode shape to identify. The damage parameter ranges from 0 (no damage) to 1 (maximum damage). The pseudo-experimental mode shapes can be derived in principle by any identification technique. In the present context, the Complex Plane Representation (CPR) method [13] is used. Briefly speaking, the CPR method is an output-only time domain technique in which the original signal is mapped in the complex plane by computing its imaginary counterpart via the Hilbert transform [14]. This new representation makes it very

simple to identify the phase shift of the motion between the different measurement points and, therefore, it is particularly effective for the identification of complex mode shapes of general viscously damped systems.

Multi-harmonic output signals of the structural dynamic response are obtained by exciting the framed structure with the Northridge earthquake, Arleta and Nordhoff Fire Station, 1994 (Fig. 2). The seismic excitation is gradually increased to increase the damage. In particular, the input intensity is progressively increased by scaling the earthquake profile to the following PGA levels: 0,1g; 0,3g, 0,5g; 0,7g; 0,9g. The elastic response spectrum of pseudo-acceleration (5% damping) is shown in Fig. 3, where the portion of interest of the spectrum is highlighted. In particular, it is important to note that this portion allows to verify the effectiveness of the modal complexity indices. In fact, as damage progresses, the first mode shape of the structure increases its natural period from 0,3 s to 0,5 s and simultaneously reduces, about 25%, the maximum pseudo-acceleration acting on the structure. As an example, the time-histories of the dynamic response in acceleration at the three levels of the structure in Fig. 1 are given below in Fig. 4 and in Fig. 5, respectively, for the minimum and maximum PGA values (0,1g and 0,9g).

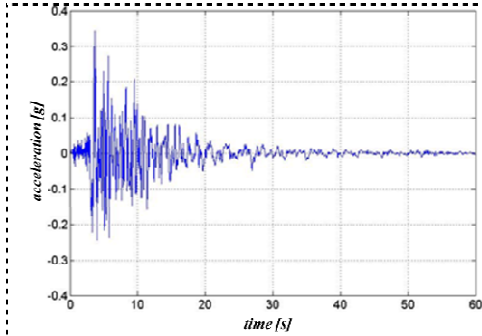


Figure 2. Northridge earthquake, comp. 90°, 1994.

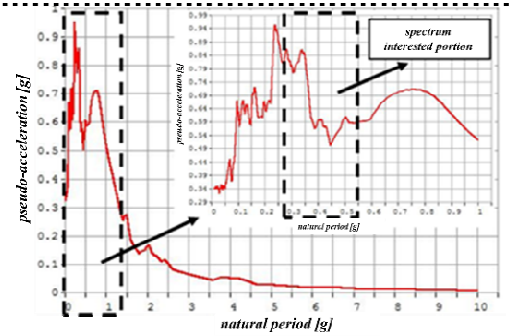


Figure 3. Elastic response spectrum - Northridge earthquake, comp. 90°, 1994.

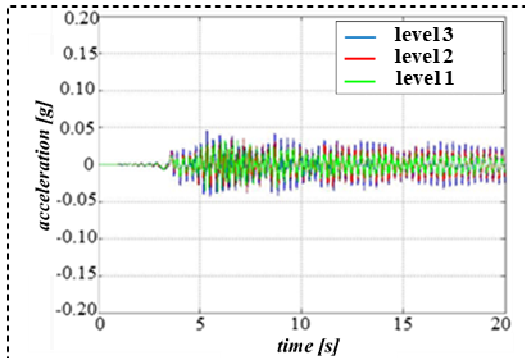


Figure 4. Dynamic response. PGA = 0,1g

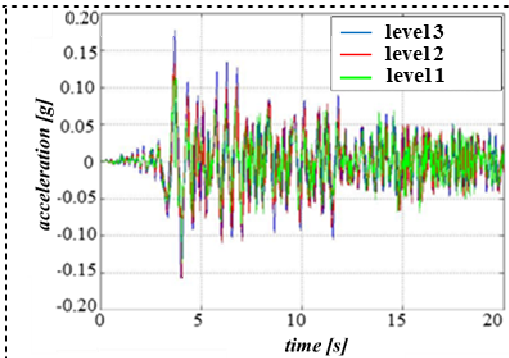


Figure 5. Dynamic response. PGA = 0,9g

As shown in Fig. 4 and Fig. 5, the structural response is composed by multi-harmonic signals. Therefore, in order to use the CPR method, to identify the pseudo-experimental mode shapes, it is necessary to decompose the signals in their individual harmonic components (each referred to a particular mode shape). Subsequently, the individual harmonics related to the same mode shape are considered to form a structural response composed by quasi-mono-harmonic signals in each dof of the structure, in such a way to recover the “resonance conditions” and hence the applicability of the CPR method. In this work, the decomposition

procedure for obtaining the individual harmonic components of the response signals is carried out using the time domain Empirical Mode Decomposition (EMD) [15]. The EMD method is applied in the time interval [5,5 s; 7,5 s] where the response signals are closer to the stationary (i.e. resonance) conditions necessary to apply the CPR method, see Fig. 4 and Fig. 5.

The results obtained after to apply the EMD method are given from Fig. 6 to Fig. 11 and referred to the minimum (PGA = 0,1g) and maximum (PGA = 0,9g) earthquake intensity. More in detail, the Fig. 6-7, Fig. 8-9 and Fig. 10-11 show, in the order, the results relative to extraction of the first, second and third quasi-harmonic component (i.e. with lower, middle and higher frequency associated to the first, second and third mode shape, respectively) of the multi-harmonic response signals of each frame level.

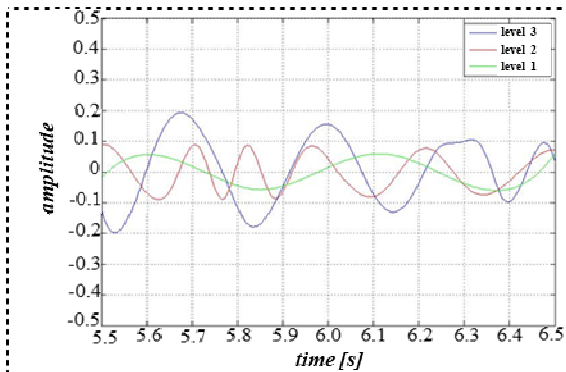


Figure 6. Quasi-harmonic component with lower frequency (first component) of the dynamic response. PGA = 0,1g

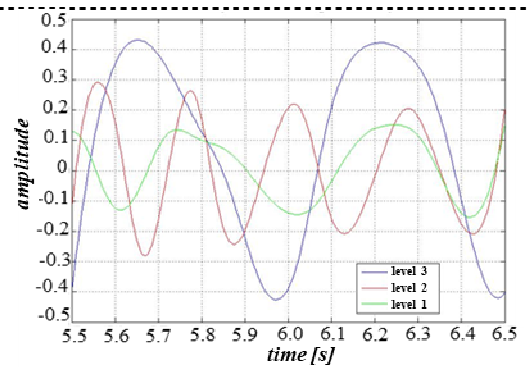


Figure 7. Quasi-harmonic component with lower frequency (first component) of the dynamic response. PGA = 0,9g

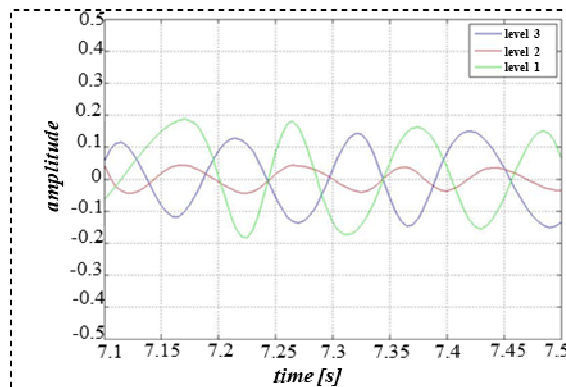


Figure 8. Quasi-harmonic component with middle frequency (second component) of the dynamic response. PGA = 0,1g

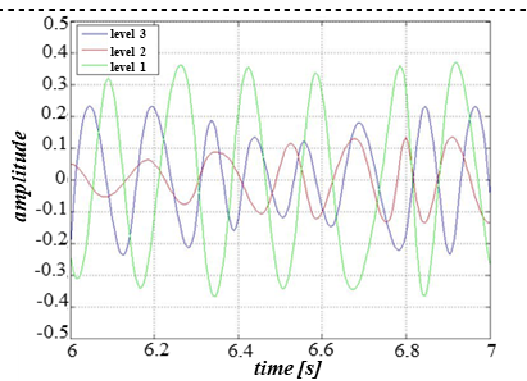


Figure 9. Quasi-harmonic component with middle frequency (second component) of the dynamic response. PGA = 0,9g

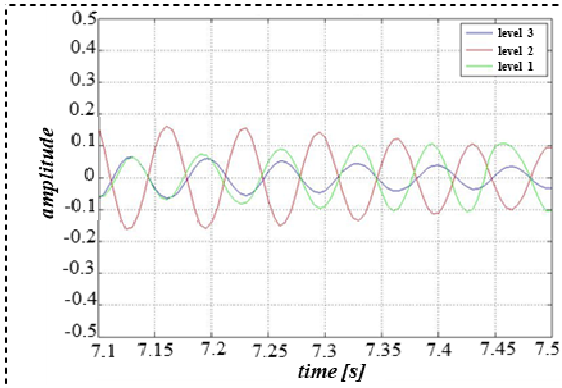


Figure 10. Quasi-harmonic component with higher frequency (third component) of the dynamic response. PGA = 0,1g

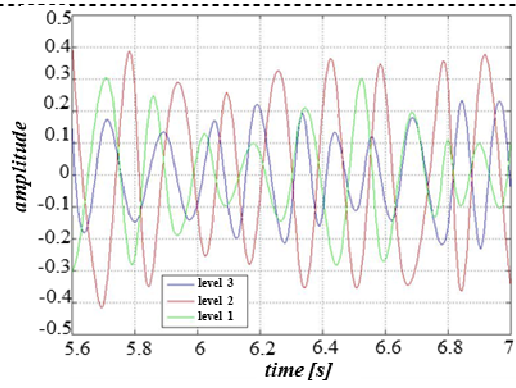


Figure 11. Quasi-harmonic component with higher frequency (third component) of the dynamic response. PGA = 0,9g

Results

As a matter of example, the results obtained in “resonance conditions” (ideal seismic excitation) are shown in the figures below for the two cases of low and high damage severity. In each case, the moment-rotation hysteretic cycle of the plastic hinge localized at the column base, Fig. 12 and Fig. 14, and the displacement time-histories at each floor level, Fig. 13 and Fig. 15, of the structure in Fig. 1, are reported.

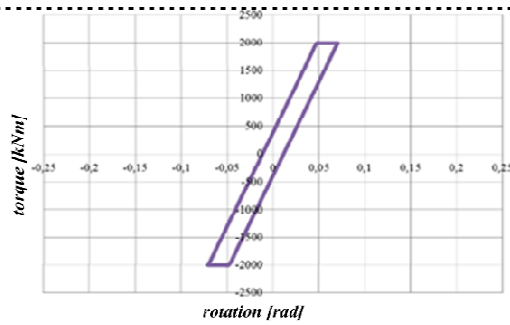


Figure 12. Moment-rotation diagram of the plastic hinge at the column base (low damage)

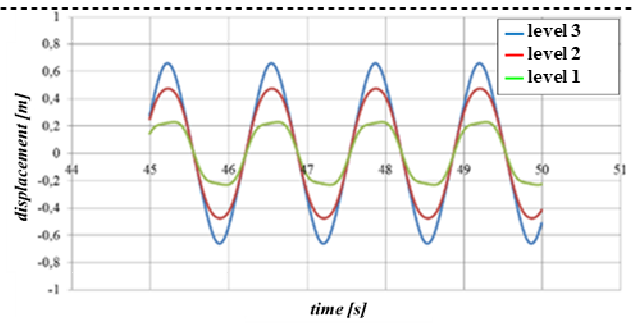


Figure 13. Dynamic response: displacement time history (low damage)

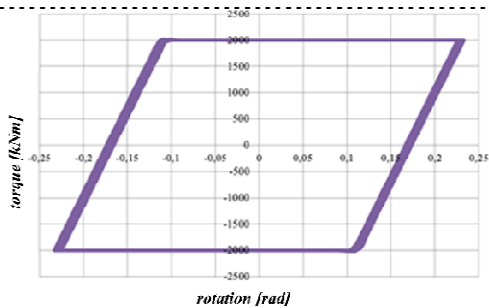


Figure 14. Moment-rotation diagram of the plastic hinge at the column base (high damage)

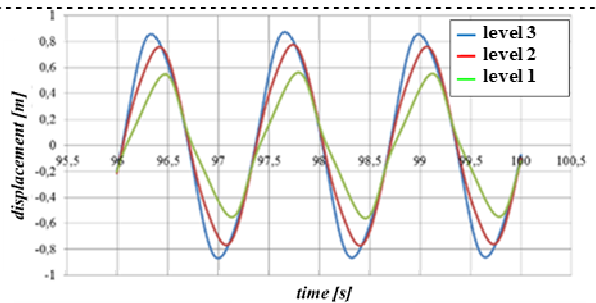


Figure 15. Dynamic response: displacement time history (high damage)

The analysis of the results highlights that as damage increases, both the area of the hysteresis cycle and the phase difference between the response signals of the three levels of the structure increase as well. Now, since the energy dissipation is a function of the area of the hysteresis cycle and the modal complexity is a function of the phase difference, both the energy dissipation and the modal complexity increase in turn along with the damage severity.

In the following figures, the variations of the five modal complexity indices versus the damage severity (damage parameter or PGA) are reported for the first mode shape. The cases of ideal, Fig. 16, and actual, Fig. 17, seismic input are compared.

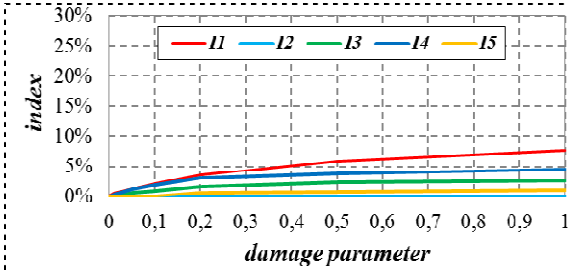


Figure 16. Indices vs. damage parameter. First mode shape

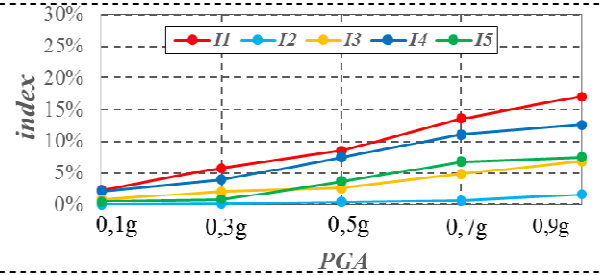


Figure 17. Indices vs. PGA. First mode shape

The analysis of Fig. 16 and Fig. 17 leads to conclude that the indices show always an increasing monotonous behavior with the damage severity. This property is fundamental when the damage identification is performed through the comparison between different states. As concerns the sensitivity, it is apparent that the indices show a significant difference among them; the more and less sensitive ones are respectively I_1 and I_2 .

Similar results are obtained for the higher order modes as it is shown in Fig. 18 and Fig. 19 for the case of the Northridge seismic input. In particular, it is observed that the indices tend to increase more rapidly along with the order of the mode shape. This is a consequence of the major articulation presented by the higher order mode shapes. In fact, the higher gradient of inter-story displacement demands higher rotations of the plastic hinges and ultimately an increase in energy dissipation.

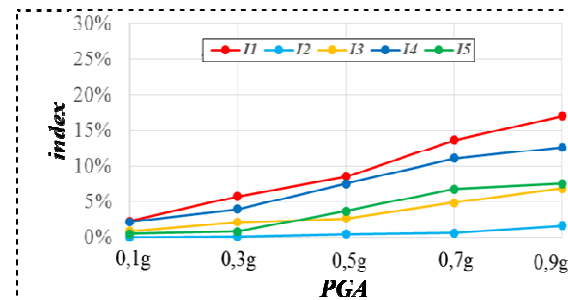


Figure 18. Indices vs. PGA. Second mode shape

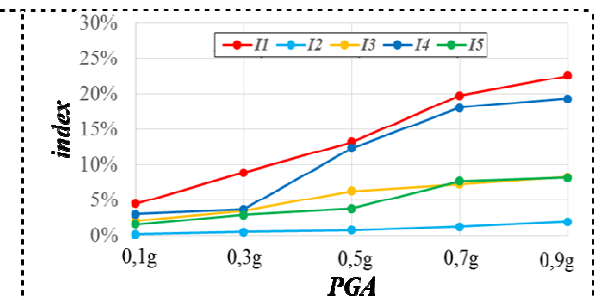


Figure 19. Indices vs. PGA. Third mode shape

However, it is observed from Fig. 17 to Fig. 19 that the indices estimated using the first mode shape are characterized by a sensitivity comparable with those of higher order mode shapes. Therefore, at least in the present case, the indices can be applied to identify the structural damage indifferently from the order of the mode shape selected. Finally, it is worth considering that the monotonic increase of the indices is preserved regardless the spectral

form of the input. In fact, in the present case, a drop in the spectral ordinates is observed passing from 0,3 s to 0,5 s (period elongation due to damage), nonetheless the indices appear not to be affected by this reduction of the seismic intensity.

Conclusions

The work investigates the effectiveness of some indices based on the mode shape complexity to identify the damage in structures experiencing plasticization. The basic idea is the relation between the damage and the complexity of the mode shapes generated by the damping non-proportionality that occurs when a structure subjected to vibration plasticizes dissipating energy. The indices derive from the technical literature and transform the identified complex mode shapes into scalar quantities readily usable in one to one relationships with any measure of the input severity. Presently, the identification of mode shapes is carried out by the joint and sequential use of the CPR and EMD methods. Nonetheless, any available signal processing technique is in principle applicable.

A simple plane frame structure endowed with plastic hinges is used as a case study. The behavior of the indices with respect to the damage is studied using seismic type base motion of progressive increasing intensity. The severity of the input is measured by the PGA of the selected earthquake.

The results confirm the monotony and sensitivity of the indices respect to the structural damage. As consequence, it is possible to identify the damage presence in a very simple form by the comparison of the indices between the current state and a previous (reference) state. More in detail, the results show a higher sensitivity of the indices measured using the higher order modes. This different sensitivity is however modest and it can be concluded that any mode shape can be effectively used to identify the damage in structures. In conclusion, it can be stated that the complexity of the mode shapes is an effective tool for a reliable detection of the structural damage. However, the use of modal complexity indices for the assessment of the damage allows only its identification but not its localization. The effectiveness of these indices rely therefore on their capability of providing synthetic and reliable indications on the structural damage presence.

References

- [1] Gunes, B. and Gunes, O. (2013) Structural health monitoring and damage assessment Part I: A critical review of approaches and methods, *International Journal of Physical Sciences Mechanics* **8**(34), 1694-1702.
- [2] Salawu, O. S. (1997) Detection of structural damage through change in frequency: a review, *Engineering Structures* **19**(9), 718-723.
- [3] Shi, Z. Y., Law, S. S., Zhang, L. M. (2000) Damage localization by directly using incomplete mode shapes, *Journal of Engineering Mechanics* **126**(6), 656-660.
- [4] Curadelli, R. O., Riera, J. D., Ambrosini, D., Amani, M. G. (2008) Damage detection by means of structural damage identification, *Engineering Structures* **30**, 3497-3504.
- [5] Egba, E. I. (2012) Detection of structural damage in building using changes in modal damping mechanism, *I. J. E. M. S.* **3**(3), 250-255.
- [6] Lage, Y., Cachão, H., Reis, L., Fonte, M., de Freitas, M., Ribeiro, A. (2014) A damage parameter for HCF and VHCF based on hysteretic damping, *International Journal of Fatigue* **62**, 2-9.
- [7] Iezzi, F., Spina, D. and Valente, C. (2015) Damage assessment through changes in mode shapes due to non-proportional damping, *Journal of Physics: Conference Series* **628**(1) 12019-12026(8).
- [8] Iezzi, F., Valente, C. and Zuccarino, L., The measure of the modal complexity as structural damage indicator (in Italian), *Proceedings of the XVI ANIDIS Conference*, L'Aquila, Italy, 13-17 September 2015.
- [9] Iezzi, F., *Structural damage identification using complex modes* (in Italian), PhD Thesis, University "G. d'Annunzio" of Chieti-Pescara, Italy, 2016.
- [10] Craig, R. R. and Kurdila, A. J. (2006) *Fundamentals of Structural Dynamics*, 2nd ed., Wiley, USA.
- [11] Iezzi, F. and Valente, C. (2017) Modal density influence on modal complexity quantification in dynamic systems, *Procedia Engineering* **199**, 942-947.

- [12] Liu, K., Kujath, M. R. and Zheng, W. (2000) Quantification of non-proportionality of damping in discrete vibratory systems, *Computers and Structures* **77**, 557-569.
- [13] Gabriele, S., Iezzi, F., Spina, D. and Valente, C., Effects of modal density in system identification using the Hilbert transform, *Proceedings of 2014 IEEE Workshop on Environmental, Energy and Structural Monitoring Systems*, Naples, Italy, 17-18 September 2014.
- [14] Feldman, M. (2011) *Hilbert Transform Applications in Mechanical Vibrations*, Wiley, UK.
- [15] Rato, R. T., Ortigueira, M. D. and Batista, A. G. (2008) On the HHT, its problems, and some solutions, *Mechanical Systems and Signal Processing* **22**, 1374-1394.

Parametrization of radiative properties of mono- and multi-component plasmas for astrophysics and nuclear fusion applications

Guadalupe Espinosa¹, †*Rafael Rodríguez^{1,2}, and Juan Miguel Gil^{1,2}

¹IUNAT, Departamento de Física, Universidad de Las Palmas de Gran Canaria, Spain.

²Instituto de Fusión Nuclear, Universidad Politécnica de Madrid, Spain.

*Presenting and †corresponding author: rafael.rodriiguezperez@ulpgc.es

Abstract

Plasma radiative properties are fundamental in many topics in plasma physics, such as nuclear fusion energy, astrophysics and extreme-ultraviolet lithography. Therefore, they are needed in radiation hydrodynamic simulations of those plasmas. However, the calculation of those properties involve the generation of huge databases of atomic data such as atomic cross sections of several processes that occur in the plasma and the resolution of very large non-linear coupled rate equations to determine the atomic level populations in the plasma. Since these ones depend on the plasma conditions which are determined by the radiation-hydrodynamic simulations at each instant and position, the whole set of equations must be solved self-consistently which makes in-line radiation-hydrodynamics simulations unfeasible. One of the solutions is to perform parametrizations of the plasma radiative properties as a function of the plasma conditions which leads to considerable reductions in computational costs. However, most of the parametrizations available are carried out for particular thermodynamic regimes (Coronal or local thermodynamic equilibria) of the plasmas and are not accurate out of those regimes. In this work, we present parametrizations of average radiative properties as a function of plasma density and temperature useful for astrophysics and nuclear fusion applications. The databases of the properties were generated using a recent code we have developed where a collisional-radiative model is implemented which ensures that the radiative properties obtained are accurate for any thermodynamic regime of the plasma.

Keywords: Non-local thermodynamic equilibrium plasmas, Plasma radiative properties, Generation of databases and parametrization of radiative properties, Mono- and multicomponent plasmas.

Introduction

Plasma radiative properties, i.e. the opacity and the emissivity, play a pivotal role in nuclear fusion and astrophysics. In astrophysics, the opacities of the stellar mixtures control the energy transfer in the stars, affecting their structure and evolution [1] and also govern the levitation of metals in stellar interiors [2]. Furthermore, the plasma emissivity is a key quantity in the structure, behavior and stability of radiative shock waves which are present in many astrophysical scenarios. Hence, for example, the onset of thermal instabilities, that can be the responsible of the origin of some astrophysical objects, is related to temperature dependence of the radiative power loss (i.e. the frequency integrated emissivity) in the post-shock medium. In the field of inertial fusion confinement, opacities are relevant in the design of hohlraum walls in the indirect drive scheme and also for the dopants embedded in the ablator of the target since they control the absorption of the thermal radiation coming from the hohlraum [3]. On the other hand, in magnetic confinement fusion, the radiative power loss plays an important role in the current decays after disruptions caused by strongly radiating

impurities [4] and also in the radiation losses from impurities that can help in the development of thermal instabilities at the plasma edge of the fusion devices [5].

The numerical simulations of the plasma phenomena above commented require of the resolution of the radiation-hydrodynamic equations (RHE). For a single fluid, that does not involve interior mass sources, the continuity equation is given by [6]

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0 \quad (1)$$

where ρ and \mathbf{u} are the fluid density and velocity, respectively. The general transport equations for momentum and energy, in the non-relativistic limit, are given by

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) = -\nabla(p + p_R) + \nabla \cdot \underline{\sigma}_v + \mathbf{F}_{EM} + \mathbf{F}' \quad (2)$$

and

$$\frac{\partial}{\partial t} \left(\rho \epsilon + \frac{\rho u^2}{2} + E_R \right) + \nabla \cdot \left[\rho \mathbf{u} \left(\epsilon + \frac{u^2}{2} \right) + p \mathbf{u} \right] = -\nabla H - \mathbf{J} \cdot \mathbf{E} + \mathbf{F}' \cdot \mathbf{u} \quad (3)$$

where p and p_R are the fluid and radiation pressures, respectively, $\underline{\sigma}_v$ is the viscous stress tensor, \mathbf{F}_{EM} denotes the force density due to the interaction of the electromagnetic fields and charges and \mathbf{F}' the density of other possible forces. In Eq. (3), ϵ and E_R denote the energy density of the fluid and the radiation field, respectively, and $\mathbf{J} \cdot \mathbf{E}$ represents the volumetric heating by the current induced in the plasma by the electromagnetic waves. The divergence of the energy flux is given by

$$\nabla H = \nabla \cdot [\mathcal{F}_R + (p_R + E_R) \mathbf{u} + \mathbf{Q} - \underline{\sigma}_v \cdot \mathbf{u}] \quad (4)$$

where \mathcal{F}_R is the spectral radiation energy flux which is given by

$$\mathcal{F}_R(\mathbf{r}, t) = \frac{1}{4\pi} \int_0^\infty \int I_\nu(\mathbf{r}, t, \nu) d\nu \mathbf{n} d\Omega \quad (5)$$

where ν is the photon frequency, \mathbf{n} is a unit vector in the direction of propagation for any value of Ω and I_ν is the spectral radiation intensity. The radiation energy density and the radiation pressure also depend on the spectral radiation intensity

$$E_R(\mathbf{r}, t) = \frac{1}{c} \int_0^\infty \int I_\nu(\mathbf{r}, t, \nu) d\nu d\Omega \quad (6)$$

$$p_R(\mathbf{r}, t) = \frac{1}{c} \int_0^\infty \int I_\nu(\mathbf{r}, t, \nu) \cos^2 \theta d\nu d\Omega \quad (7)$$

with c the speed of light. Therefore, the spectral radiation intensity is the basic macroscopic quantity to describe the radiative transfer and is obtained solving the radiative transfer equation (RTE) given by

$$\frac{1}{c} \frac{\partial I_\nu(\mathbf{r}, t, \nu, \mathbf{n})}{\partial t} + \mathbf{n} \cdot \nabla I_\nu(\mathbf{r}, t, \nu, \mathbf{n}) = -\kappa(\mathbf{r}, t, \nu) I_\nu(\mathbf{r}, t, \nu, \mathbf{n}) + j(\mathbf{r}, t, \nu) \quad (8)$$

where $j(\mathbf{r}, t, \nu)$ and $\kappa(\mathbf{r}, t, \nu)$ are the monochromatic emissivity and absorption coefficients, respectively. Both coefficients include electron transitions in the plasma between atom bound levels (line transitions or bound-bound contributions), between bound and free levels (photoionization and radiative recombination which are bound-free contributions) and between electron free levels (direct and inverse bremsstrahlung or free-free contributions). The expressions used to calculate them can be found elsewhere [7]. Two ingredients are needed to compute them. First, the cross sections of the radiative processes, which are obtained through atomic simulations. Secondly, the populations of the atomic levels in the plasma. In general, plasmas are in non-local thermodynamic equilibrium (NLTE) regime and the atomic-level populations (for a given plasma condition, i.e. for a plasma density and temperature) are obtained from the solution of a system of rate equations of a so-called collisional-radiative model (CRM) [8][9]. This set of kinetic equations is given by

$$\frac{dN_{\zeta i}(\mathbf{r}, t)}{dt} = \sum_{\zeta' j} N_{\zeta' j}(\mathbf{r}, t) \mathbb{R}_{\zeta' j \rightarrow \zeta i}^+ - \sum_{\zeta' j} N_{\zeta i}(\mathbf{r}, t) \mathbb{R}_{\zeta i \rightarrow \zeta' j}^- \quad (9)$$

where $N_{\zeta i}$ is the population density of the atomic level i of the ion with charge state ζ . The terms $\mathbb{R}_{\zeta' j \rightarrow \zeta i}^+$ and $\mathbb{R}_{\zeta i \rightarrow \zeta' j}^-$ take into account all the collisional and radiative processes which contribute to populate and depopulate the level ζi , respectively. Since the rate equations included the radiative processes in the plasma, i.e. the absorption and emissivity coefficients, these equations are coupled to the RTE. Furthermore, since the CRM depends on the plasma density and temperature, which are obtained from the single fluid RHE of the plasma (and an equation of state), the rate equations are also coupled to them.

Radiation-hydrodynamic simulations (RHS) of nuclear fusion and astrophysical plasmas require the calculation of plasma radiative properties for density and temperature profiles that could include between $10^2 - 10^3$ different plasma conditions. For each plasma condition, the set of rate equations of the CRM must be solved, which could involve around 10^4 non-linear equations including 10^6 radiative and collisional processes. Furthermore, the RTE for each photon frequency, the RHE and the rate equations must be solved self-consistently. Therefore, in-line RHS are unfeasible and approximations are usually made. In the RTE is common to assume that the radiation does not depend explicitly on time. Moreover, the RTE is solved under the grey approach, in which appropriate mean opacities are used (Planck or Rosseland mean opacities) instead of monochromatic radiative properties, thus preventing the solution of the RTE for each photon frequency. Even so, huge radiative properties databases should be generated solving the CRM coupled to the RHE, which is still a complex problem that involves large computational times. An appropriate solution to this problem is to perform parametrizations of the plasma mean radiative properties in terms of plasma density and temperature. Thus, these properties can be obtained for given plasma conditions from

polynomial fittings obtained from reduced databases, greatly decreasing the computational cost.

There are several available analytical expressions for the mean radiative properties, but more of them are accurate only for Coronal equilibrium [10]-[15], i.e. for plasmas at very low density in which they are density independent, or for the range of high photon energies assuming the plasma in LTE [16][17], i.e. for high density plasmas, but not for NLTE situations. In this work, we present a parametrization of the average radiative properties which are essential for RHS of plasmas in nuclear fusion and astrophysics scenarios, such as the average ionization, Rosseland and Planck mean opacities (which are weighted averages in frequency of the monochromatic opacities) and the radiative power loss, for mono- and multi-component plasmas. The parametrization was performed using the PARPRA code [18] and is based on polynomial functions where the coefficients of the fitting were determined by means of a least square regression. A quad-tree algorithm was used to optimize the division of the space of plasma conditions to parametrize. The databases of the radiative properties to parametrize are generated for a representative set of plasma conditions using the RAPCAL code [19][20] using the the plasma atomic level populations obtained from the resolution of the rate equations of the CRM implemented in the POLAR code (a recent extension of the MIXKIP code [21]) and for this reason the parametrizations obtained are valid for Coronal equilibrium, LTE and NLTE regimes. The paper is structures as follows. Next section is devoted to a presentation of the theoretical and computational models used to perform the parametrizations. Thereafter, some examples of parametrizations for both mono- and multicomponent plasmas are shown. Finally, conclusions and future extensions are commented.

Theoretical and computational models

The POLAR code was developed in order to obtain the atomic level populations at typical plasma conditions obtained in nuclear fusion and astrophysics. For that purpose, a CRM was implemented including, in the rate equations (Eq. (9)), the following atomic processes that populate and de-populate the atomic levels: collisional ionization and three-body recombination, spontaneous decay, collisional excitation and deexcitation, radiative recombination, autoionization and electron capture, photoexcitation, photodeexcitation and photoionization, which are the most common processes in this kind of plasmas. In this work, we have considered that the atomic level populations do not depend on time. Under this assumption, the rate equations of the collisional-radiative steady-state (CRSS) model are given by

$$\sum_{\zeta'j} N_{\zeta'j}(\mathbf{r}, t) \mathbb{R}_{\zeta'j \rightarrow \zeta i}^+ - \sum_{\zeta'j} N_{\zeta i}(\mathbf{r}, t) \mathbb{R}_{\zeta i \rightarrow \zeta'j}^- = 0 \quad (10)$$

Steady-state approach implies that the characteristic times of the atomic processes are significantly lower than the characteristic time of the plasma evolution. This approach becomes invalid, for example, if the plasma is under the interaction with ultra-short pulsed laser (\sim fs). However, for the plasmas of interest in this work that assumption is accurate. Two complementary equations have to be satisfied together with Eq. (10). First, the requirement that the sum of fractional level populations equals to the total ion particle density, n_{ion}

$$\sum_{\zeta=0}^Z \sum_{i=1}^{M_{\zeta}} N_{\zeta i} = n_{ion} \quad (11)$$

where Z is the atomic number and M_ζ is the total number of levels for the charge state. Second, the charge neutrality condition in the plasma

$$\sum_{\zeta=0}^Z \sum_{i=1}^{M_\zeta} \zeta N_{\zeta i} = n_e \quad (12)$$

with n_e the electron particle density. The average ionization of the plasma, \bar{Z} , is then defined as the ratio between the electron and the ion particle densities.

The effect of radiative processes in the low density plasmas obtained in magnetic fusion devices and stellar atmospheres can be neglected since they can be assumed as optically thin (plasma radiation self-absorption is not significant). Furthermore, we have not considered external radiation fields in this work. On the other hand, for the high density plasmas obtained in stellar interiors and inertial confinement targets the relevance of radiative processes is significantly lower than collisional processes and can also be neglected. Therefore, the atomic processes considered in the CRSS model were collisional ionization and three-body recombination, spontaneous decay, collisional excitation and deexcitation, radiative recombination, autoionization and electron capture and the rate equations and the RTE will be uncoupled.

The calculation of the rate coefficients of those processes requires of atomic data such as atomic energy levels, oscillator strengths and cross sections. Before then, the first problem to address is the selection of a suitable set of atomic configurations for the CRM since, in principle, the number of atomic levels for a given ion is infinite. There is not a priori criterion to determine which configurations should be considered in the model and, in general, the kind of configurations to include depends on the plasma conditions, the presence of external radiation fields (such as thermal radiation or ultra-intense lasers) or the interaction with particle beams. However, since we are interested in the generation of databases of radiative plasma properties and their parametrization in wide range of plasma conditions, the criterion employed was based on a rule of thumb in which the configurations included for each ion in the model are those with energies up to twice the ionization energy of the ground configuration of the ion. With this criterion we cannot guarantee that all the configurations that have some influence in the radiative properties for a given plasma condition are included but the ones that have a large contribution are considered. The second question to address is related with the degree of detail of the atomic description. The most detailed description is the so-called detailed level accounting (DLA) approach. However, this description entails very large computational times and, therefore, it is only useful for chemical elements of low atomic number. In this work, the atomic data were generated in the detailed configuration accounting (DCA) approach. This is a statistical average of the atomic properties obtained in the DLA approach and is more accurate for elements of high atomic number. Nevertheless, since we are interested in average radiative properties, DCA approach is accurate enough for that purpose. Once the set of atomic configurations and the atomic description were selected, the atomic data used in the POLAR code were generated with the FAC code [22]. For a given ion with N electrons the energy levels are obtained in FAC by means of the diagonalization of the relativistic Hamiltonian. The wave functions are then obtained as antisymmetric sums of products of N mono-electronic Dirac spinors. Configuration interactions effects are also considered. The cross sections of the forward-going processes (collisional ionization, collisional excitation and autoionization) are calculated quantum-mechanically in the

relativistic distorted wave approach [22]. The cross sections for the corresponding inverse processes were determined from the appropriate micro-reversibility relations [23].

The atomic data provided by the FAC code are obtained for isolated atoms. However, the plasma surrounding (ions, free electrons and photons) modify the atomic data since they change the potential experienced by the bound electrons. This effect is commonly modelled in plasma physics through the so-called continuum lowering, that represents the depression of the potential with respect to the isolated situation due to the electric fields generated by the plasma charged particles. In this work, the model used for the continuum lowering is based on the widely used proposal developed by Stewart and Pyatt [24] where the correction to the ionization potential I_ζ is given by

$$\Delta I_\zeta = \frac{3a_0 I_H}{2R_\zeta} (\zeta + 1) \left\{ \left[1 + \left(\frac{D}{R_\zeta} \right)^3 \right]^{2/3} - \left(\frac{D}{R_\zeta} \right)^2 \right\} \quad (13)$$

where a_0 is the Bohr radius, I_H is the Rydberg constant, $R_\zeta = [3(\zeta + 1)/4\pi n_e]^{1/3}$ is the sphere-ion radius assuming the plasma composed of ions with charge ζ only and the Debye radius is $D = [4\pi(\bar{Z} + \bar{Z}^2)n_{ion}/T_e]^{-1/2}$, with T_e the electron temperature and \bar{Z}^2 the second order moment of the population distribution. The inclusion of the continuum lowering the kinetics rate equations of the CRSS model imply to solve them iteratively since the atomic data are now modified by the shifts caused by plasma effect.

The set of rate equations constitutes a system of M equations for the level populations, where M denotes the total number of levels included in the CRM. Therefore, the size of the collisional-radiative matrix scales like M^2 . Even in the DCA approach, the number of levels involved in collisional-radiative simulations can easily reach 10^4 which entails matrix with 10^8 elements. However, in CRM the atomic processes usually connect only levels belonging to ions with either the same charge state or to adjacent ones which means that the matrix is sparse and for this reason POLAR use sparse techniques to store and operate on only the non-zero elements, decreasing the memory requirements.

Once the plasma level populations are calculated with the POLAR code, these ones along with the oscillator strengths and the photoionization cross sections obtained using the FAC code are the input to the RAPCAL code to calculate the plasma radiative properties. Since in this work we show some examples of the parametrization of the radiative power loss, in the following we present a brief explanation about its calculation. The method to calculate the other plasma radiative properties in RAPCAL can be found in [19][20]. The radiative power loss is the frequency integrated plasma emissivity, if the plasma may be considered as optically thin. The emissivity, $j(\mathbf{r}, t, \nu)$, has three contributions. The bound-bound contribution, $j_{bb}(\mathbf{r}, t, \nu)$, is given by

$$j_{bb}(\nu) = \sum_\zeta \sum_{i,j} j_{\zeta j \rightarrow \zeta i}(\nu), \quad j_{\zeta j \rightarrow \zeta i}(\nu) = \frac{h\nu}{4\pi} N_{\zeta j} A_{\zeta j \rightarrow \zeta i} N \phi_{ij}(\nu) \quad (14)$$

where we have omitted the dependence on the position and time in the formula for simplicity. $A_{\zeta j \rightarrow \zeta i}$ is the Einstein coefficient for spontaneous deexcitation between the bound states j, i of the ion ζ and h is the Planck's constant. $\phi_{ij}(\nu)$ is the line profile and in its evaluation of the

line profile, natural, Doppler, and electron-impact [25] broadenings were included and also the Unresolved Transition Array width [26], which is a statistical method to take into account the atomic fine structure of the spectra in the DCA atomic approach used in this work. The line-shape function is applied with the Voigt profile that incorporates all these broadenings. The bound-free contribution to the emissivity, $j_{bf}(\mathbf{r}, t, \nu)$, is determined by means of

$$j_{bf}(\nu) = \sum_{\zeta+1,j} \sum_{\zeta,i} j_{\zeta+1,j \rightarrow \zeta,i}(\nu)$$

$$j_{\zeta+1,j \rightarrow \zeta,i}(\nu) = \frac{h^4 \nu^3 n_e}{2\pi c^2 \varepsilon^{\frac{1}{2}}} \left(\frac{1}{2m_e} \right)^{\frac{3}{2}} N_{\zeta+1,j} f(\varepsilon) \frac{g_{\zeta,i}}{g_{\zeta+1,j}} \sigma^{\text{pho}}_{\zeta+1,j \rightarrow \zeta,i}(\nu) \quad (15)$$

where ε is the free electron energy and m_e the electron mass. In this work, a Maxwell-Boltzmann distribution $f(\varepsilon)$ at the electron temperature is assumed. Photoionization cross section, $\sigma^{\text{pho}}_{\zeta+1,j \rightarrow \zeta,i}(\nu)$, were calculated quantum-mechanically using the FAC code in the relativistic distorted wave approach. $g_{\zeta,i}$ denotes the statistical weight of level i . Finally, for the free-free contribution to the emissivity a semi-classical expression, based on the Kramer's inverse bremsstrahlung cross section [27], was used

$$j_{ff}(\nu) = \frac{32\pi^2 e^4 a_0^2 \alpha^3}{\sqrt{3}(2\pi m_e)^{3/2} h} \left(\frac{m_e}{2\pi k_B T_e} \right)^{1/2} \bar{Z}^2 n_{\text{ion}} n_e e^{-h\nu/k_B T_e} \quad (16)$$

where k_B is the Boltzmann's constant and α is the fine structure constant. Therefore, the total emissivity is the sum of these three contributions and the radiative power loss is the emissivity integrated in frequency, as said above.

The ranges of plasma conditions of interest in RHS are, in general, very wide, covering several orders of magnitude in temperature and density. For this reason, we have obtained that an optimum grid in which generating the databases of the radiative properties using POLAR is based on logarithmic meshes with steps of 0.1 and 0.5 for the temperature and density grids, respectively, since the radiative properties are more sensitive to the plasma temperature than to the density. Once the databases have been generated we can proceed to their parametrization. Due to the nature of the radiative properties, we have obtained that the parametrization of the decimal logarithm of the radiative property than the property itself [28] shows numerical advantages. The polynomial employed for the fitting is given by

$$P(d, T_e) = \sum_{i=0}^n \sum_{j=0}^m C_{ij} (\log d)^i (\log T_e)^j \quad (17)$$

where d denotes either the electron particle density, the ion particle density or the density of matter of the plasma. The coefficients of the parametrization, C_{ij} , are obtained by the minimization of the following function through least square regression

$$F(C_{ij}) = \sum_{k=1}^{n_d} \sum_{l=1}^{n_T} [P(d_k, T_{e,l}) - \log A(d_k, T_{e,l})]^2 \quad (18)$$

where $A(d_k, T_{e,l})$ is the radiative property evaluated at a given plasma condition. Then, from the equation of the minimization

$$\frac{\partial F(c_{ij})}{\partial c_{ij}} = 0 \quad (19)$$

the following set of $(n + 1) \cdot (m + 1)$ coupled algebraic equations is obtained

$$\sum_{k=1}^{n_d} \sum_{l=1}^{n_T} \sum_{i=0}^n \sum_{j=0}^m c_{ij} (\log d_k)^{i+q} (\log T_{e,l})^{j+r} = \sum_{k=1}^{n_d} \sum_{l=1}^{n_T} \log A(d_k, T_{e,l}) (\log d_k)^q (\log T_{e,l})^r \quad (20)$$

with $q = 0, \dots, n$ and $r = 0, \dots, m$. Polynomial of high degree for the fittings can be highly oscillatory and may provide very inaccurate values for the radiative properties for plasma conditions not belonging to the databases. This fact is avoided fixing the maximum degree of the polynomial both for the temperature, m , and density, n , to 7, which are enough for the properties in which we are interested. To obtain the coefficients of the fitting we start from the lowest degree and we increase it until the relative difference between the fitting and the database value is lower than an imposed criterion. If the maximum degree is reached and the criterion has not been fulfilled, then the range of density and temperature considered must be sub-divided in short ranges and then the procedure starts again. In order to optimize this division, we have employed a quad-tree algorithm. Obviously, the number of polynomials obtained (and of sub-divisions of the space of plasma conditions) depends on the criterion imposed. As the criterion becomes more restrictive, the number of polynomial needed increases. This procedure is integrated in a computational code named PARPRA [18] which, in addition, was developed with a graphic interface to ease its use by the user. A detailed explanation of the code can be found in [29].

The databases of the radiative properties are generated for monocomponent plasmas. However, both in nuclear fusion and astrophysics, multicomponent plasmas (i.e. plasma mixtures) are commonly found. In this case, the fittings of the individual elements of the mixture can be used for determining its radiative properties. The procedure followed to obtain the mixture radiative properties depend on the type of density given as an input. If the input is the electron density, then the procedure is very simple since we only have to add the fitted radiative property of each single element, weighted by the abundance of each element in the mixture, in order to obtain the total one. On the other hand, if the input is the ion number density or the mass density we have to make an iterative procedure. Hence, we start assuming an electron density for the mixture equal to the ion density or to the mass density divided by the Avogadro's number and multiplied by the mixture atomic mass. With this electron density we obtain the average ionizations of the different elements in the mixture from the fitting of the average ionization of the single elements for that electron density. With this set of average ionizations and weighted by their fractional abundances in the mixture we obtain a new electron density and we repeat this procedure until the relative difference between the average

ionizations of the mixture in two consecutive steps of the iterative procedure will be lower than an imposed criterion. We would like to point out the relevance of this procedure since it avoids the resolution of the CRM for a mixture which involves to solve as many set of rates equations as elements of the mixture, that, for example, in astrophysics can easily reach 26 elements.

Results

In this section we present some examples of the fittings of some of relevant microscopic radiative properties in order to show the utility and the accuracy of the fitting proposed. In Figure 1 we show an example of characteristic databases, generated using POLAR, to parametrize. In this case the property is the average ionization for Xe and Kr in a range of plasma conditions typically obtained in experiments of laboratory astrophysics in which scaled astrophysical phenomena are reproduced in laboratory with ultra-intense laser and noble gases.

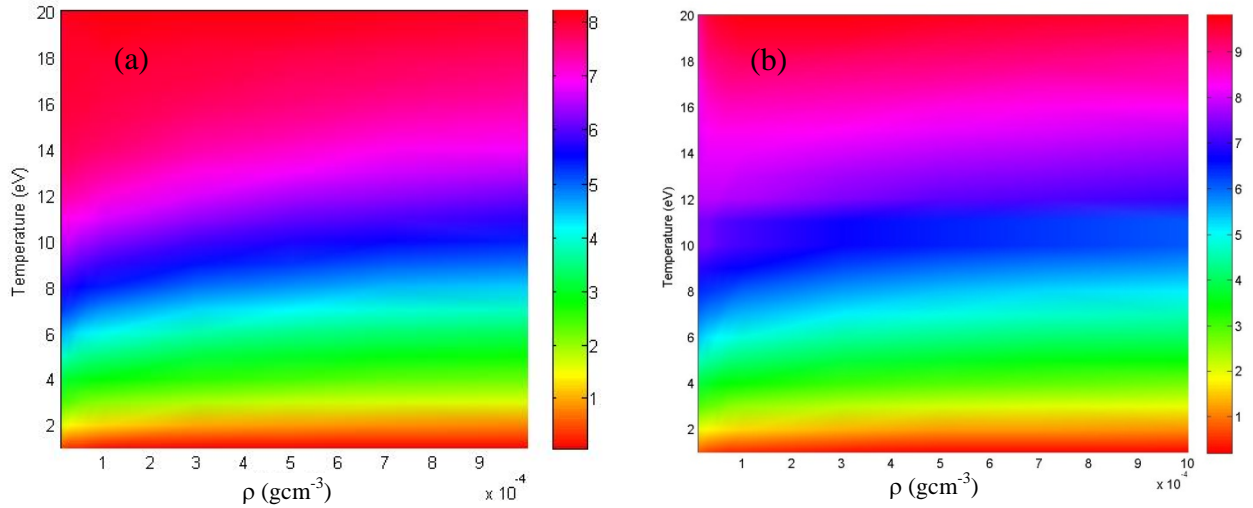


Figure 1. Representation of the databases of the average ionization generated by POLAR as a function of the electron temperature and density for plasmas of (a) Kr and (b) Xe.

In Figure 2, we present the parametrizations carried out with PARPRA of the average ionization and the radiative power loss of a krypton plasma as a function of the temperature and for several mass densities. For the fitting, the criterion imposed was a relative error lower than 1%. We also show in the figure the relative errors obtained between values obtained with POLAR and the parametrization either for plasma conditions used in the parametrization or not. We can observe that for the latter the relative error is sometimes larger than the criterion used in the fitting although they are still small and near to the criterion (lower than 1.5%). The errors obtained are slightly greater for the radiative power loss than for the average ionization since the former is a less average property than the latter and, therefore, is more sensitive to the plasma conditions. The number of polynomial functions required for the parametrization of the whole set of plasma conditions was 4 both for the average ionization and the radiative power loss.

The examples presented in the images correspond to parametrizations of radiative properties of monocomponent plasmas (krypton in particular). As commented above, a great advantage

of this kind of parametrization is that it can be used to obtain the radiative properties of multicomponent plasmas without making any new parametrization or the calculation of the databases for the mixture. In Table 1 we present the comparison of the average ionization and the radiative power loss for a plasma mixture of four elements (neon, argon, aluminium, which are of interest in astrophysics, and xenon, which is commonly used as temperature moderator in nuclear fusion chambers) from the database and the ones obtained from the parametrization of the individual elements. We have assumed the same relative abundance of the four elements in the mixture. The databases of the single elements were parametrized with a relative error of 1% for the plasma conditions of the mesh of the database. From the table, we can observe that although the radiative properties of the multicomponent plasma in PARPRA are generated from the ones fitted for each single element of the mixture, the agreement with the values obtained from the collisional-radiative simulations using POLAR is really good. This is a remarkable result since this means a considerable reduction both in the complexity of the problem of dealing with plasma mixtures and also in the computing time.

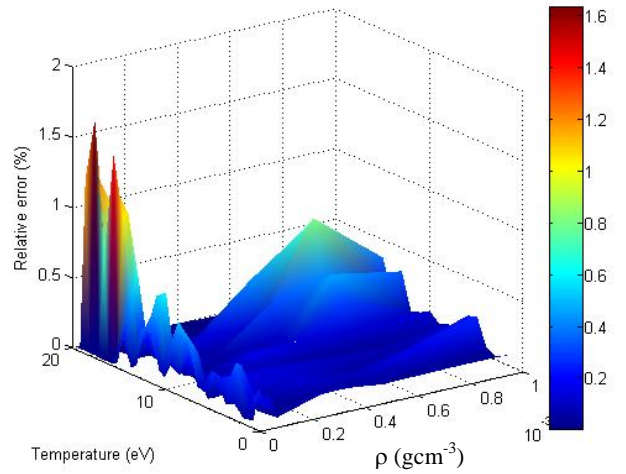
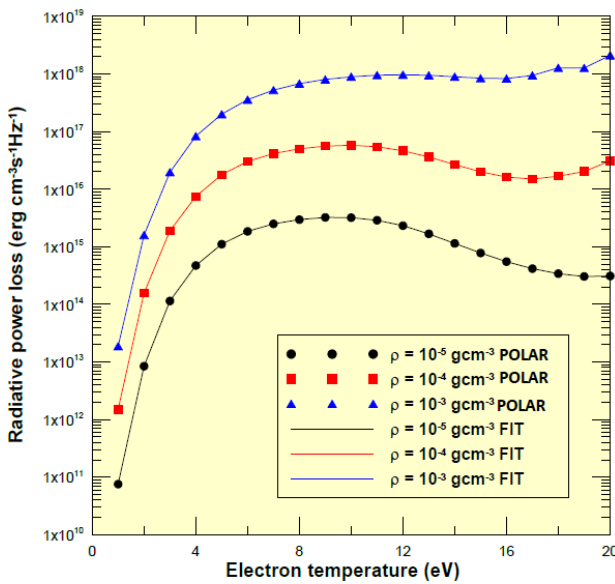
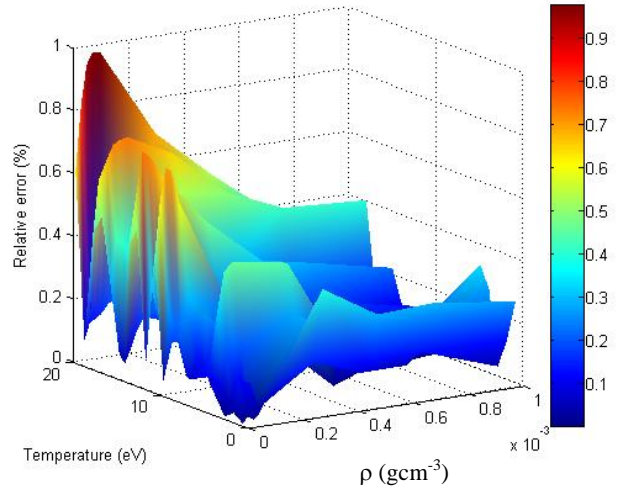
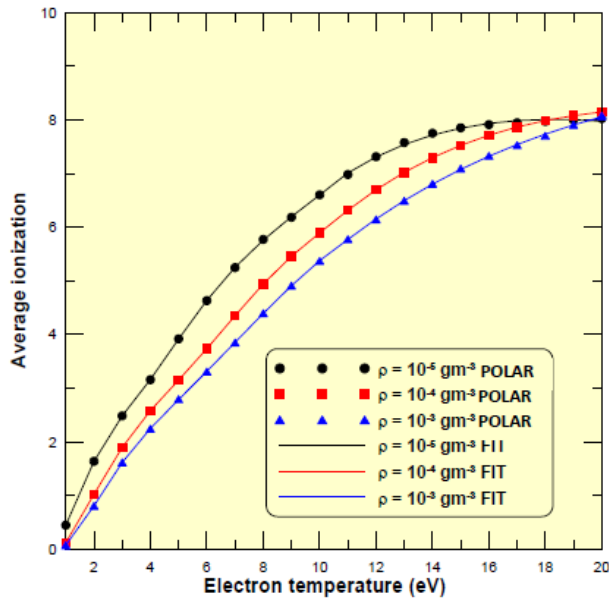


Figure 2. Parametrization of the average ionization and radiative power loss for several mass densities and for the range of temperatures 1-20 eV for a Kr plasma. The figure also shows the relative errors in the fitting for values of the properties belonging or not to the database fitted.

Table 1. Comparison of the average ionization (\bar{Z}) and the radiative power loss (RPL), in erg/s/cm³, for a plasma mixture of four elements, provided by the POLAR simulations and their parametrization with PARPRA.

T (eV)	d (gcm ⁻³)	\bar{Z} (POLAR)	\bar{Z} (PARPRA)	RPL (POLAR)	RPL (PARPRA)
2	10 ⁻⁵	1.293	1.239	3.564×10 ¹³	3.579×10 ¹³
10	10 ⁻⁵	4.046	4.046	5.352×10 ¹⁵	5.351×10 ¹⁵
20	10 ⁻⁵	6.060	6.058	1.286×10 ¹⁶	1.287×10 ¹⁶
20	10 ⁻⁴	6.164	6.158	3.963×10 ¹⁷	3.973×10 ¹⁷

Conclusions

In this work we have presented a method to parametrize average plasma radiative properties, implemented in the PARPRA code in terms of the plasma temperature and density by means of polynomial functions, which is very useful in RHS since these parametrizations considerably reduce the computational costs. The databases of the radiative properties to parametrize are generated using a recently developed code named POLAR which has implemented a CRM. Therefore, the radiative properties and, therefore, their parametrizations, are accurate for any plasma thermodynamic regime. The criterion imposed in the parametrization is fixed by the user. Obviously, as the criterion becomes more restrictive the number of polynomial functions required to parametrize the whole set of plasma conditions increases. We have presented, as example, a parametrization of the average ionization and the radiative power loss of a Kr plasma imposing a relative error in the fitting of 1%. The number of polynomials required was 4 for both properties and the errors obtained in the calculation of the properties from the parametrization at plasma conditions not included in the fitting were also very near to the criterion. We have also showed the utility and the accuracy of the parametrization of the radiative properties of the single elements for obtaining those for plasma mixtures. It is worth pointing out the advantage of providing this kind of parametrizations since the radiative properties are needed in radiation-hydrodynamics simulations of plasmas in nuclear fusion and astrophysics in wide range of plasma conditions being their calculation very complex and computing time consuming.

In this work we have limited the parametrization to mean radiative properties since we have assumed the gray approach for the RTE in the RHS. However, this may be a rude approach in several scenarios. More realistic radiative transfer simulations require more detailed descriptions of the radiative properties such as those based on multigroup descriptions (the gray approach can be considered as one group approach). Parametrizations of multigroup radiative properties as a function of the plasma conditions and photon frequency groups would be highly useful and this will be a future goal of this work.

Acknowledgements

This work has been supported by the EUROfusion Consortium TASK AGREEMENT WPENR: Enabling Research IFE, Project No. AWP15-ENR-01/CEA-02, by the Project of the Spanish Government with reference FIS2016-81019-P.

References

- [1] Fengtao, J., Jiaolong, Z., Jianmin, Y., Guoxing, H., Zeqing, W., Yun, Y., Wan, M. and Peng, Y. (2005) L to M shell transitions and model comparisons for radiative opacities of sodium fluoride plasmas, *Journal of Quantitative Spectroscopy and Radiative Transfer* **95**, 241-253.
- [2] Bailey, J.E., Rochau, G.A., Mancini, R.C., Iglesias, C.A., MacFarlane J.J., Golovkin, I.E., Blancard, C., Cosse, Ph. and Faussurier, G. (2009) Experimental investigations of opacity models for stellar interior, inertial fusion and high energy density plasmas, *Physics of Plasmas* **16**, 058101.
- [3] Benredjem, D., Jarrah, W., Gilleron, F., Pain, J.C., Ferri, S. and Calisti, A. (2015) Opacity calculations. Ge and Si dopants in ICF, *High Energy Density Physics* **16**, 23-27.
- [4] Smith, H.M. and Verwiltche E. (2008) Hot tail runaway electron generation in tokamak disruptions, *Physics of Plasmas* **15**, 072502.
- [5] Tokar, M.Z. and Kelly, F.A. (2003) The role of plasma-wall interactions in thermal instabilities in tokamak edges, *Physics of Plasmas* **10**, 4378-4386.
- [6] Drake, R. P. (2005) *High energy density physics: fundamentals, inertial fusion and experimental astrophysics*, Springer, The Netherlands.
- [7] Rodriguez, R., Espinosa, G., Gil, J.M., Stehle, C., Suzuki-Vidal, F., Rubiano, J.G., Martel, P. and Minguez, E. (2015) Microscopic properties of xenon plasmas for density and temperature regimes of laboratory astrophysics experiments on radiative shocks, *Physical Review E* **91**, 053106.
- [8] Bates, D.R., Kingston, A.E. and McWhirther, R.W. P. (1962) Recombination between electrons and atomic ions. I. Optically thin plasmas, *Proceedings of the Royal Society of London Series A: Mathematical and physical sciences* **267**, 297-312.
- [9] McWhirther, R.W. P. (1978) Data needs, priorities and accuracies for plasma spectroscopy, *Physics Reports* **37**, 165-209.
- [10] Post, D.E., Jensen, R.V. and Tarter, C.B. and Grasberger, W.H. and Lokke, W.A. (1977) Steady-state radiative cooling rates for low density, high-temperature plasmas, *Atomic Data and Nuclear Data Tables* **20**, 397-439.
- [11] Summers, H.P. and McWhirther, R.W.P. (1979) Radiative power loss from laboratory and astrophysical plasmas. I. Power loss from plasmas in steady-state ionisation balance, *Journal of Physics B: Atomic and Molecular Physics* **12**, 387-411.
- [12] Fournier, K.B., Pacella, D., May, M.J., Finkenthal, M. and Goldstein, W.H. (1997) Calculation of the radiative cooling coefficient formolybdenum in a low density plasma, *Nuclear Fusion* **37**, 825-834.
- [13] Fournier, K.B., Cohen, M., May, M.J. and Goldstein, W.H. (1988) Ionization state distribution and radiative cooling rate for argon in a low density plasma, *Atomic Data and Nuclear Data Tables* **70**, 231-254.
- [14] Fournier, K.B., May, M.J., Pacella, D., Finkenthal, M., Gregory, B.C. and Goldstein, W.H. (2000) Calculation of the radiative cooling coefficient for krypton in a low density plasma, *Nuclear Fusion* **40**, 847-863.
- [15] Schure, K.M., Kosenko, D., Kaastra, J.S., Keppens, R. and Vink, J. (2009) A new radiative cooling curve based on an up-to-date plasma emission code, *Astronomy & Astrophysics* **508**, 751-U240.
- [16] Minguez, E., Ruiz, R., Martel, P., Gil, J.M., Rubiano, J.G. and Rodriguez, R. (2001) Scaling law of radiative opacities for ICF elements, *Nuclear Instruments and Methods in Physics Research A* **464**, 218-224.
- [17] Minguez, E., Martel, P., Gil, J.M., Rubiano, J.G. and Rodriguez, R. (2002) Analytical opacity formulas for ICF elements, *Fusion Engineering. Design* **60**, 17-25.
- [18] Rodriguez, R., Espinosa, G., Gil, J.M., Rubiano, J.G., Mendoza, M.A., Martel, P. and Minguez, E. (2014) Parametrization of Mean Radiative Properties of Optically Thin Steady-State Plasmas and Applications, *Communications in Computational Physics* **16**, 612-631.
- [19] Rodríguez, R., Florido, R., Gil, J.M., Rubiano, J.G., Martel, P. and Mínguez, E. (2008) RAPCAL: a computational package to compute radiative properties of optically thin and thick low and high-Z plasmas in a wide range of density and temperature, *Laser and Particle Beams* **26**, 433-448.
- [20] Rodriguez, R., Florido, R., Gil, J.M., Rubiano, J.G., Suarez, D., Martel, P., Minguez, E. and Mancini, R.C. (2010) Collisional-Radiative calculations of optically thin and thick plasmas using the computational package ABAKO/RAPCAL, *Communications in Computational Physics* **8**, 185-210.
- [21] Espinosa, G., Rodríguez, R., Gil, J.M., Suzuki-Vidal, F., Lebedev, S.V., Ciardi, A., Rubiano, J.G. and Martel, P. (2017) Influence of atomic kinetics in the simulation of plasma microscopic properties and thermal instabilities for radiative bow shock experiments, *Physical Review E* **95**, 033201.
- [22] Gu, M.F. (2008) The flexible atomic code, *Canadian Journal of Physics* **86**, 675-689.
- [23] Blatt, J.M. and Weisskopf, V.F. (1991) *Theoretical Nuclear Physics*, Dover Publications, New York.
- [24] Stewart, J.C. and Pyatt, K.D. (1966) Lowering of ionization potentials in plasmas (1966), *The Astrophysical Journal* **144**, 1203-1211.
- [25] Dimitrijevic, M.S. and Konjevic, N. (1987) Simple estimates for Stark-broadening of ion lines in stellar plasmas, *Astronomy & Astrophysics* **172**, 345-349.
- [26] Bauche, J., Bauche-Arnoult, C. and Klapisch, M. (1987) Transition arrays in the spectra of ionized atoms, *Advance Atomic Molecular Physics* **23**, 131-195.

- [27] Rose, S.J. (1992) Calculation of the radiative opacity of laser-produced plasmas, *Journal of Physics B: Atomic and Molecular Physics* **25**, 1667-1681.
- [28] Espinosa, G., Rodríguez, R., Gil, J.M., Suzuki-Vidal, F., Lebedev, S.V., Ciardi, A., Rubiano, J.G. Martel, P. (2017) *Physical Review E* **95**, 033201.
- [29] Espinosa, G. (2015) *Determinación y parametrización de propiedades radiativas de plasmas para la simulación y análisis de experimentos de astrofísica de laboratorio*. Ph. D Thesis, Las Palmas de Gran Canaria.

Droplet morphology simulation with SPH: A simple method of implementing contact angles

Xiangwei Dong¹

¹Department of Engineering Mechanics, China University of Petroleum(East China), Qingdao, China

*Presenting author: dongxw139@163.com

†Corresponding author: dongxw139@163.com

Abstract

The use of the smoothed particle hydrodynamics (SPH) method to simulate the surface tension effects requires an efficient description of the three phase contact line among liquid, gas, and solid. In this study, based on the continuum surface force (CSF) model, the virtual interface method is proposed to implement the contact angle in the SPH multiphase simulation. The first step is to create the virtual gas-liquid interface at the triple contact point according to the desired contact line. Then, the adjusting force acting on the fluid of the triple line can be expressed as the function of the interface curvature, which can gradually adjust the dynamic contact line towards the desired contact angle. The proposed method is successfully applied to simulate the droplet morphologies on smooth and roughness surfaces. The numerical results are in agreement with the previous analytical solutions. These analyses hold great potentials in bio-inspired superhydrophobic surfaces, oil displacement, microfluidics, and ore floatation, etc.

Keywords: Smoothed particle hydrodynamics (SPH); virtual interface method; equilibrium contact angle; droplet morphology; roughness surface substrate.

1. Introduction

Surface tension and contact line dynamics play a major role in many industrial fields, including inkjet printing[1], powder agglomeration[2] and microfluidics[3], etc. These phenomena also occur in the nature world, such as the superhydrophobic effect of the surface of the lotus leaf[4], nutrition transportation in plants via Phloem and Xylem tissues[5], and some aquatic creatures walking on water[6]. Therefore, much effort has been devoted to explore this area, both from experimental and numerical viewpoints. It should be mentioned that it is not easy to perform experiments in some extreme conditions, and thus numerical simulation must be the necessary approach, which can shorten time and lower expenses. A main task is to develop the numerical model in consideration of the surface tension.

Currently, two common computational fluid dynamics (CFD) approaches are used in multiphase simulations with surface tension effects being considered. The first one is the volume-of-fluid method (VOF)[7] which belongs to grid-based methods, and the second one can be grouped into the particle-based Lagrangian methods, such as the smoothed particle

hydrodynamics (SPH) method[8]. The SPH method was originally proposed to simulate the astrophysical phenomena[9] and then widely used in the simulation of hydrodynamics[10] and solid mechanics[11]. Up to now, the SPH method has also been applied in the simulation of multiphase flows[12–18]. And the Lagrangian nature of SPH makes it a good option for simulating complex multiphase problems.

In SPH multiphase simulations, there exist two ways to implement the surface tension. The first is based on the van der Waals force implemented through particle-particle interactions[20]. Surface tension effects can be simulated by adding the van der Waals force in the pressure term. Tartakovsky and Meakin[21] combined the standard SPH equations with pairwise fluid-solid particle-particle interactions to simulate capillarity. In their work, the simulation of contact lines is realized through the interplay between the pinning forces provided by particle-particle interactions, and driving forces from the gravity. The second method is the continuum surface force (CSF) method, where the surface tension is expressed as a function of the interface normal vector, curvature and surface tension coefficient[22]. The CSF method was first implemented in SPH by Morris[23] in use of the color function for the interface tracking. It is widely accepted that the CSF method is more perceptible than the first method because its formulation contains the surface tension coefficient, which is convenient to control the surface tension according to actual physics.

The implementation of the CSF method requires proper boundary conditions to account for the three-phase contact. Liu and Liu[24] simulated the micro-channel flow adopting an asymmetric SPH model, and the simulation of the contact line with solid surface is realized based on the normal correction method proposed by Brackbill et al.[22]. Hu and Adams[25] developed a multi-phase SPH method from a particle smoothing function in which the neighboring particles only contribute to the specific volume but not density, and this method was used to simulate channel flow by using virtual particles to account for the wall-fluid interaction. But only the straight channel wall was considered. Das and Das proposed another method to handle the wall-fluid interaction, in which the equilibrium contact angle was simulated by incrementally translating particles in the triple contact region. In this method, a constant reconstruction of the triple line is required, and the method was only tested for smooth surfaces also. Breinlinger[28] proposed a smoothed normal correction method to modify the interface normal vectors near the triple line, and introduced an adjusting force to obtain the desired contact line position. Yeganehdoust[19] adopted a similar way to introduce the adjusting force, and proposed a new scheme to correct the normal vectors by re-assigning the color values of dummy particles for solid wall. However, the method needs additional fitting parameters to tune the result so that it is only applied to smooth surfaces. As we all know, most wetting and dewetting phenomena of solid materials are closely related to the micro-structures of their surfaces[29]. For example, the leaves of lotus can keep off rain drops due to its hierarchical surface structures. Inspired by the skill of superhydrophobicity of lotus, rough surfaces with appropriate wetting properties have various applications in industry, e.g., porous media, micro-fluidic devices, self-cleaning paints, and glass windows. Therefore, it is necessary to develop the numerical model for simulation of surface wetting for roughness surface, i.e., the surface with micro-structures.

Following the previous studies, and on the basis of the SPH multiphase algorithm and CSF model, a simple and deterministic method of implementing contact angle is proposed, which can effectively simulate the droplet morphology on the solid surface. In the method, dummy particles of solid wall are used to create the virtual gas-liquid interface near the triple line. Then, the adjusting force which is the function of the interface curvature is introduced to

adjust the dynamic contact line towards the desired contact angle. The method is suitable for both smooth surface and roughness surface substrates. Only two-dimensional problems will be considered in this paper, though the presented model and method can be extended to three-dimensional cases without too much efforts.

2. Model formulation

2.1 Field equations

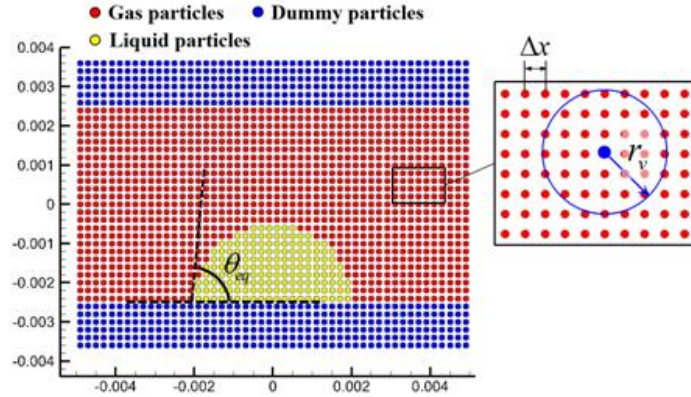


Figure 1 Modeling on a droplet

We first introduce the schematic of the computational model of a droplet deposited on a smooth surface. As shown in Figure 1, at the triple contact line, there is an angle between the liquid-gas interface and solid-liquid interface, which is termed as the Young's contact angle. In the two-dimensional case, the morphology of the droplet can be formulated by the elliptic integrals once the volume and Young's contact angle are given. Herein, we want to simulate the droplet configuration in use of the SPH method. As demonstrated in Figure 1, a computational domain is selected to include the droplet-substrate system. Then the domain is discretized by a finite number of particles, and each particle is assigned physical parameter values and certain volume.

In this study, the Navier–Stokes equation, in combination with the mass conservation equations, are adopted to describe the motion of the fluids, i.e. the gas and liquid. These field equations in Lagrangian view are expressed as follows[12]:

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{u}, \quad (1)$$

$$\rho \frac{D\mathbf{u}}{Dt} = \nabla P + \mathbf{f}^V + \mathbf{f}^B + \mathbf{f}^S + \mathbf{f}^I, \quad (2)$$

where ρ is the mass density, t is the time variable, \mathbf{u} is the velocity, ∇P is the pressure gradient, \mathbf{f}^V , \mathbf{f}^B , and \mathbf{f}^S are the forces per unit mass, which correspond to the viscous force, body force (such as gravity) and surface tension force, respectively.

The fluid is considered as incompressible, so the viscous force per unit mass can be expressed as:

$$\mathbf{f}^V = \mu \nabla^2 \mathbf{u}, \quad (3)$$

where μ is the dynamic viscous coefficient.

The surface tension force is computed using the continuum surface force (CSF) method[22]. According to the CSF method, f^S can be written as

$$\mathbf{f}^S = -\sigma \kappa \mathbf{n} \lambda, \quad (4)$$

where σ is the surface tension of the liquid, which is assumed as a constant, κ is the curvature of an arbitrary point in the interface, \mathbf{n} is the normal unit vector of the interface, and λ is a smeared delta function.

An equation of state is compensated to calculate the fluid pressure[10]:

$$P = \frac{c^2 \rho_0}{\gamma} \left(\left(\frac{\rho}{\rho_0} \right)^\gamma - 1 \right) + P_b, \quad (5)$$

where c is the speed of sound, ρ_0 is the reference density of fluid, γ is a constant parameter[14], and P_b is the background pressure which helps to prevent the tensile instability, and also keep the particles distributed uniformly[13].

2.2 SPH discretization

Evidently, the analytical solutions of Eqs. (1), (2) and (5) are intractable, and thus the numerical approach is adopted. Firstly, these equations should be discretized based on the SPH method, where the computational domain is initialized by a series of uniformly distributed particles. Considering there is a high density ratio between the two fluids, Eqs. (7) and (8) are discretized in the form of[12]:

$$\rho_i = m_i \sum_j W_{ij}, \quad (6)$$

$$\rho_i \frac{D\mathbf{u}_i}{Dt} = \left(-\frac{1}{V_i} \sum_j (P_i V_i^2 + P_j V_j^2) \nabla_i W_{ij} \right) + \mathbf{f}_i^V + \mathbf{f}_i^B + \mathbf{f}_i^S, \quad (7)$$

where the subscripts i and j are the particle indices. The symbol $V_i = \frac{m_i}{\rho_i}$ is the volume of the i th particle, W_{ij} represents the renormalized Gaussian kernel function $W(\mathbf{r}_i - \mathbf{r}_j, h)$ [14], where \mathbf{r}_i is the position vector of the i th particle, and h is the smoothing length. In the simulation, the value of h is set as a constant, which reads: $h = 1.0\Delta x$, and Δx is the initial particle spacing.

The viscous force per unit mass is discretized by the following equation[30]:

$$\mathbf{f}_i^V = \sum_j \mu_{ij} \frac{V_i^2 + V_j^2}{V_i} \frac{(\mathbf{r}_i - \mathbf{r}_j) \cdot \nabla_i W_{ij}}{|\mathbf{r}_i - \mathbf{r}_j|^2 + (0.01h)^2} (\mathbf{u}_i - \mathbf{u}_j), \quad (8)$$

where $\mu_{ij} \left(= \frac{2\mu_i \mu_j}{\mu_i + \mu_j} \right)$ is the mean inter particle viscosity, μ_i and μ_j are the dynamic viscosity of the i th particle and the j th particle.

2.3 Surface tension modeling

According to the Eq. (4), the surface tension force is correlated with the interface normal vector and curvature, so these two parameters should be calculated first. In the two phase flow,

for the i th particle of fluid 1, if some particles of fluid 2 exist in the supporting domain of the i th particle, the surface tension force f_i^S can be expressed as[17]:

$$f_i^S = -\sigma^{1-2}\kappa_i|\nabla C_i|\mathbf{n}_i, \quad (9)$$

where σ^{1-2} represents the surface tension between fluid 1 and 2. The parameter κ_i is the curvature of the i th particle which is located in the interface, and C is the so-called color function and is defined as a pairwise form[17]:

$$C_i^j = \begin{cases} \frac{2\rho_i}{\rho_i + \rho_j} & \text{if the phase of the } i\text{th particle is different from the } j\text{th particle} \\ 0 & \text{if the phase of the } i\text{th particle is the same with the } j\text{th particle} \end{cases} \quad (10)$$

The normal unit vector at the i th particle is calculated using the following equation:

$$\mathbf{n}_i = \frac{\nabla C_i}{|\nabla C_i|}, \quad (11)$$

where ∇C_i is the gradient of the color function and its expression is[17]

$$\nabla C_i = \frac{1}{V_i} \sum_j (V_i^2 + V_j^2) \frac{C_i^j + C_j^i}{2} \nabla_i W_{ij}. \quad (12)$$

Then, the curvature of the i th particle is calculated based on the results of normal unit vectors:

$$\kappa_i = -d \frac{\sum_j (\mathbf{n}_i - \varphi_i^j \mathbf{n}_j) \cdot \nabla_i W_{ij} V_j}{\sum_j |\mathbf{r}_i - \mathbf{r}_j| \cdot |\nabla_i W_{ij}| V_j}, \quad (13)$$

where φ_i^j is defined similar to the color function[17]:

$$\varphi_i^j = \begin{cases} -1 & \text{if the phase of the } i\text{th particle is different from particle } j \\ 1 & \text{if the phase of particle } i \text{ is the same with particle } j \end{cases} \quad (14)$$

The function of φ_i^j is to reverse the direction of the normal vector in the neighbouring phase.

2.4 Time-stepping scheme

A modified prediction-correction time-stepping scheme[14], which is proposed for multiphase droplet dynamics with high density ratio, is applied. The explicit time-stepping scheme is subject to the Courant-Friedrichs-Levy (CFL) condition for stability. The time step is determined according to the following criterion:

$$\Delta t_c = CFL_c \frac{h}{c + u_{max}}, \quad (15)$$

where CFL_c is the coefficient set as $CFL_c = 1.0$, u_{max} is the maximum velocity in the fluid field. Besides, the magnitude of the time step Δt is also determined based on the following criterions:

$$\Delta t_s = CFL_s \left(\frac{\rho h^3}{2\pi\beta} \right)^{1/2}, \quad (16)$$

$$\Delta t_\mu = CFL_\mu \frac{\rho h^2}{\mu}, \quad (17)$$

where the coefficients are set as $CFL_s = 0.5$, $CFL_\mu = 0.125$. The final time step is determined as $\Delta t = \min(\Delta t_c, \Delta t_s, \Delta t_\mu)$.

3. Contact angle implementation

In this section, the method of implementing contact angle is proposed to simulate the droplet morphology with the desired contact angle value. The equilibrium contact angle θ_{eq} are given as the input parameters.

3.1 Interfacial forces at the triple line

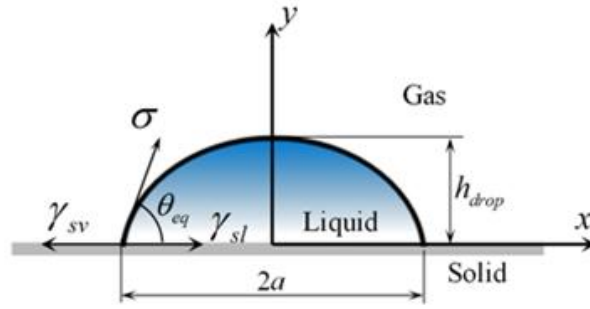


Figure 2 Illustration of interface tensions at the liquid-gas (σ), the solid-liquid (γ_{sl}), and the solid-gas (γ_{sv}) interface.

As shown in Figure 2, according to the Young's equation, the interfacial forces acting on the fluids of the droplet near the triple line can be expressed as:

$$\sigma \cos \theta_{eq} + \gamma_{sl} - \gamma_{sv} = 0, \quad (18)$$

where σ , γ_{sl} and γ_{sv} are the interface tensions at the liquid-gas, the solid-liquid, and the solid-gas interface, respectively.

In the dynamic simulation using SPH, if the instantaneous contact angle θ of a droplet is not consistent with the equilibrium contact angle θ_{eq} , the adjusting force should be introduced based on the following equation[28]:

$$F^a = \sigma(\cos \theta_{eq} - \cos \theta), \quad (19)$$

where F^a is the adjusting force, which is distributed into fluid particles near the triple line as the following equation:

$$\mathbf{F}^a = \sum_{j \in \Omega_{tri}} \mathbf{f}_j^a \quad (20)$$

where \mathbf{f}_j^a is the force acting on the j th particle of fluids; Ω_{tri} represents the domain near the triple line, and it is determined by the circle ($r_v = 3h$) centered on the triple contact point, as shown in Figure 3.

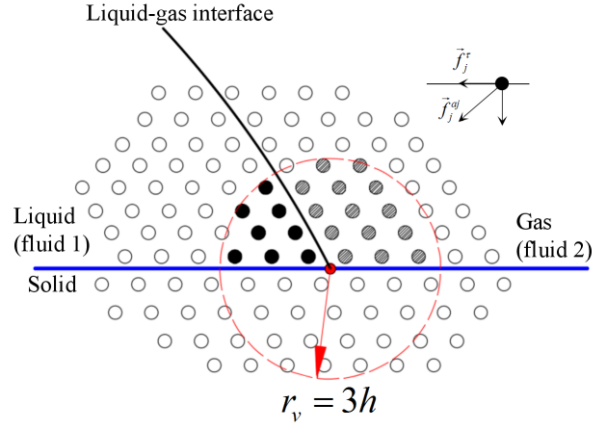


Figure 3 Shadowy circles mark the liquid particles located in the triple contact region. The adjusting force is distributed on these particles.

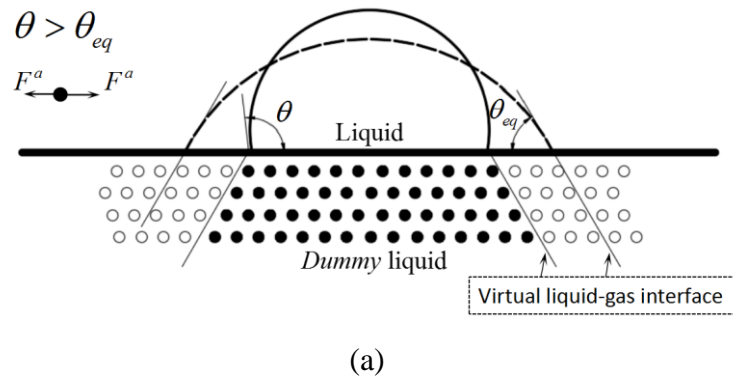
3.2 The virtual interface method

Here, we introduce the following equation to calculate the adjusting force:

$$\mathbf{f}_i^a = -\delta \kappa_i |\nabla C_i| \boldsymbol{\tau}_w, \quad (21)$$

where δ is a coefficient and is set equal to σ^{1-2} in this study, $\boldsymbol{\tau}_w$ is the tangential unit vector of the solid surface, κ_i and $|\nabla C_i|$ are calculated by the Eqs. (11), (12), and (13).

Before implementing Eq. (21), the dummy particles of the solid boundary should be treated first. As shown in Figure 4, the dummy particles are divided into two classes by the desired contact line. One part is attributed to type 1 and is given the same mass density as the liquid (as the solid circle marks in Figure 4), the other is attributed to type 2 which is given the same mass density as the gas. This is equivalent to create a dummy liquid-gas interface at the end of the real liquid-gas interface. Based on above treatment, the adjusting force (Eq. (21)) which has the same formulation as the surface tension force (Eq. (9)), is able to eliminate the concave or convex at the triple contact point. This is the basic principle of the virtual interface method.



(a)

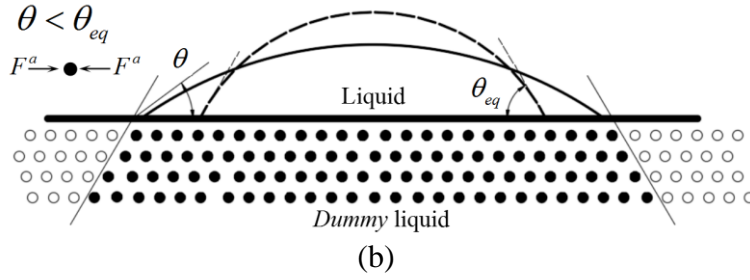


Figure 4 A droplet is deposited on the surface. Illustration of the distribution of dummy particles. (a) Non-equilibrium droplet with $\theta > \theta_{eq}$, (b) Non-equilibrium droplet with $\theta < \theta_{eq}$.

Figure 4 presents two situations that droplets are in non-equilibrium state with the instantaneous contact angle $\theta > \theta_{eq}$ and $\theta < \theta_{eq}$, respectively. If $\theta > \theta_{eq}$, as shown in Figure 4 (a), the adjusting force tends to stretch the droplet outwards due to the concave curvature at the triple contact point. While for the case of $\theta < \theta_{eq}$ (see Figure 4(b)), the adjusting force will stretch the droplet inwards. The equilibrium contact angle can be gradually approached under the effect of the adjusting force.

Figure 5 shows the distribution of interface normal vectors near the triple line. As shown in the Figure, for the case without using the virtual interface method, the fluid particles near the triple line have insufficient supporting particles because of the cutoff of the supporting domain by the wall boundary, which leads to incorrect results of normal vectors. Figure 5 (b) shows the results of normal vectors with the virtual interface method. The dash line represents the virtual liquid-gas interface which can be seen as the extension of the real liquid-gas interface. It shows that better distribution of normal vectors are obtained when the virtual interface method is used.

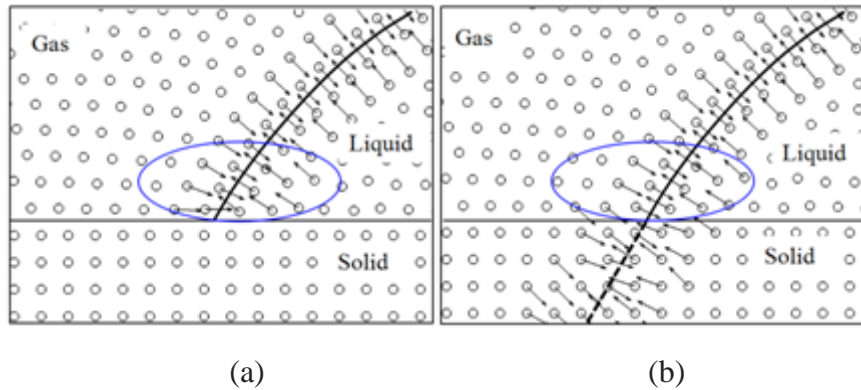


Figure 5 Normal vectors at the triple line (a) without using the virtual interface method and (b) using the virtual interface method

It can be seen that the first step of implementing the virtual interface method is to detect the triple contact points.

3.3 Implementation

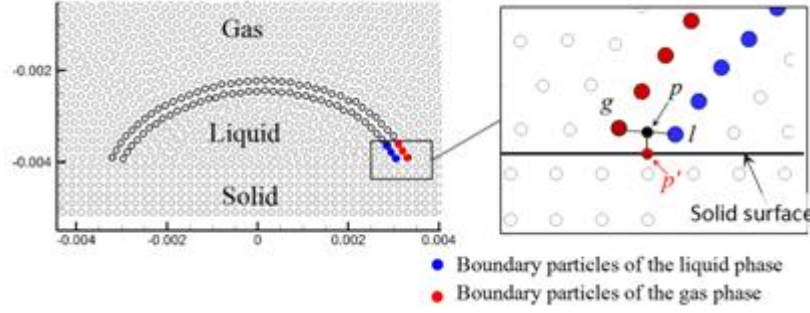


Figure 6 Detection of the triple contact point

For the two-dimensional problem, the triple contact point can be defined as the intersection between liquid-gas interface and solid surface. We propose the following procedure to identify the three-phase contact point in 2D as:

- 1) Search the boundary particles of the liquid and gas phases, respectively.
- 2) Identify two most downward interface particles close to solid surface, labeled as points g and l .
- 3) Take the mean coordinates of g and l as the intermediate point p , and then the projection of point p on the solid surface (p') is regarded as the triple contact point.

Following the method proposed by Dilts [32], the boundary particles can be detected by scanning the $3h$ -radius ($3h$ is the radius of the support domain) circle around an SPH particle. If the circle of the SPH particle is not completely covered by the circles of its neighbors, this particle is labeled as a boundary particle. Otherwise, it is an inner particle.

Figure 7 illustrates the complete procedure of implementing the virtual interface method.

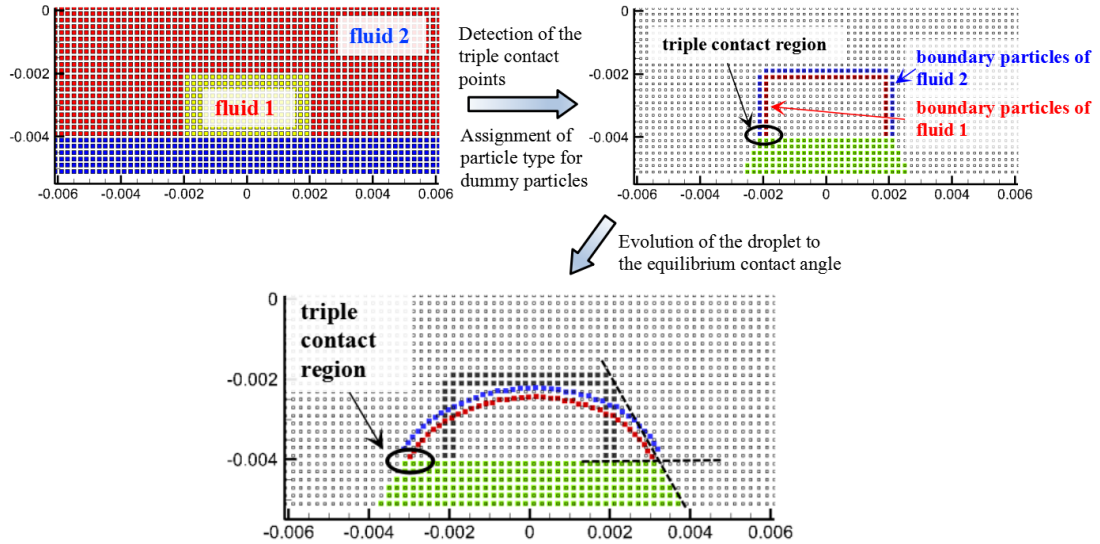


Figure 7 Implementation procedure of the virtual interface method

Figure 8 shows the complete flow chart of the SPH simulation, which is implemented by every time step. .

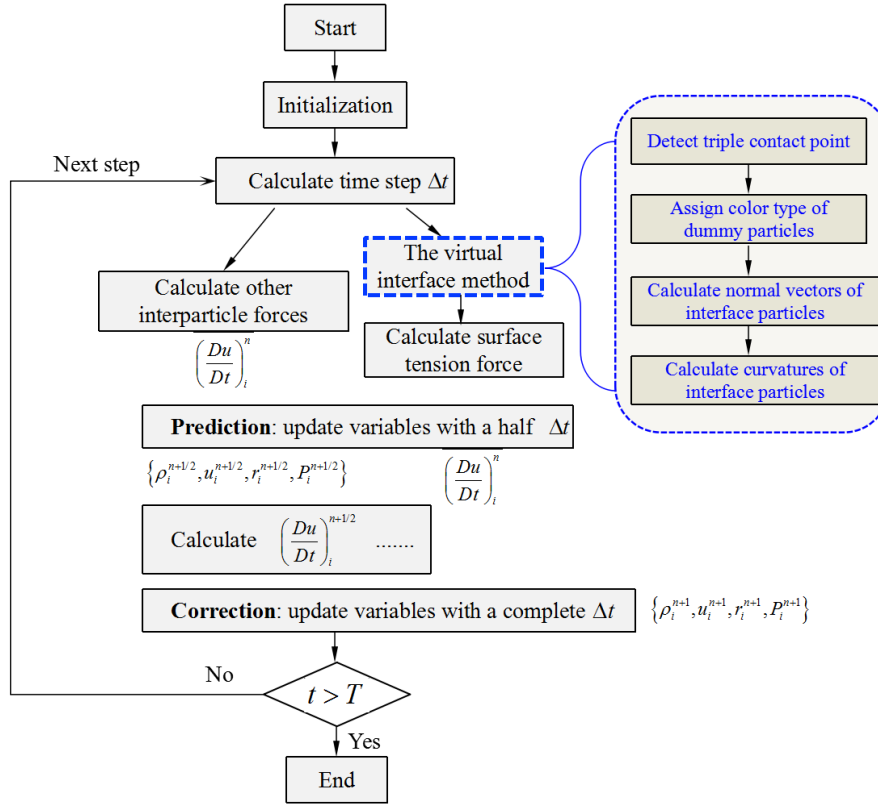


Figure 8 Flow chart of the numerical simulation implemented with the virtual interface method

4. Results and discussion

We consider a droplet deposited on the solid surface and initialized as square shape. Various contact angles can be simulated by changing the input conditions, i.e., the equilibrium contact angles. The simulation domain is a rectangular box of dimensions $L_x \times L_y$ discretized by different resolutions. At the bottom center, a rectangular region of $l_x \times l_y$ is assigned to fluid 1 (liquid), while the surrounding particles are assigned to fluid 2 (gas). The fluid phases include water (liquid) and air (gas). The material properties of liquid and gas are set according to water and air, where the density ratio is 1000.0, the viscosity ratio is 10.0, and the surface tension coefficient is set as 0.072 N/m.

4.1 Evolution of liquid lumps towards the equilibrium contact angle

We first investigate the evolution of the initially square liquid lump. The equilibrium contact angle is set as $\theta_{eq} = 60^\circ$. As shown in Figure 9, the sharp corners become smoothed under the effect of the surface tension. At the same time, the droplet is stretched outwards due to the adjusting force of contact angle. During the evolution process, the position of the triple contact point continually changes, so the color distribution of the dummy particles also varies with time. The droplet experiences a period of oscillation until it evolves to the equilibrium contact angle. From the last picture of Figure 9 one can see that the tangential line of the liquid-gas interface at the triple contact point is visually coincided with the desired contact line, which illustrates that the contact angle method works.

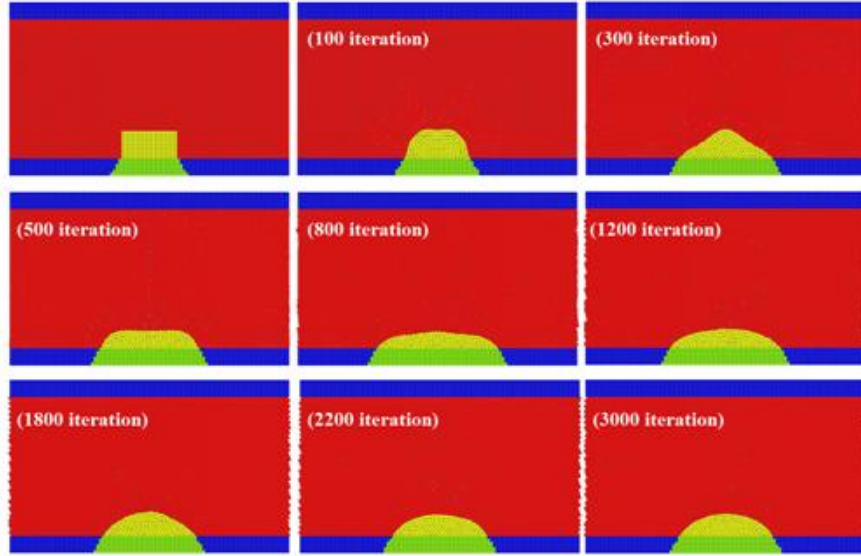
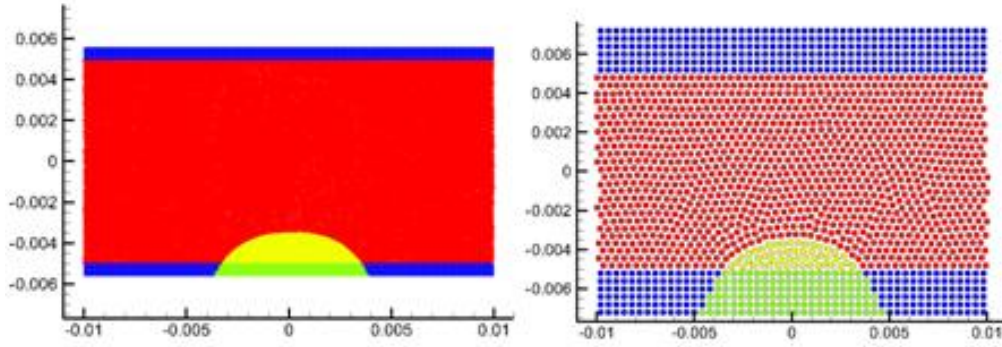


Figure 9 Evolution of a liquid lump to the equilibrium contact angle. (Particle spacing: 100×50 , $\theta_{eq} = 60^\circ$)



(a) Particle spacing: 200×100

(b) Particle spacing: 50×25

Figure 10 Simulation using different particle spacing. ($\theta_{eq} = 60^\circ$)

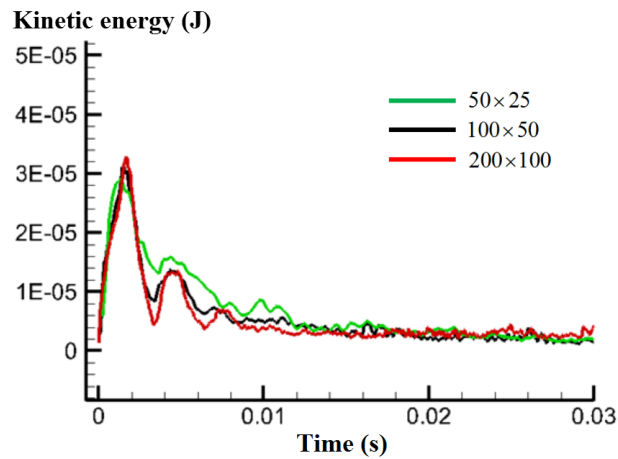


Figure 11 Time history of the kinetic energy of the droplet

Then, we compare the results with different particle spacings, as shown in [Figure 10](#). It illustrates that the final droplet morphology is insignificantly influenced by the particle

spacing. Figure 11 shows the time history of the kinetic energy of the droplet during the simulation. At early stages, the kinetic energy reaches to the peak value because both the surface tension and adjusting force have the largest values. At approximately $t=0.01$ s circular droplet is formed and the particles are nearly at rest, *i.e.* the kinetic energy is very low. For different particle spacings the curve of kinetic energy generally have the same pattern, but the oscillation magnitude increases with the decrease of particle spacing. That is probably because the adjusting force of contact angle is dependent on the particle spacing.

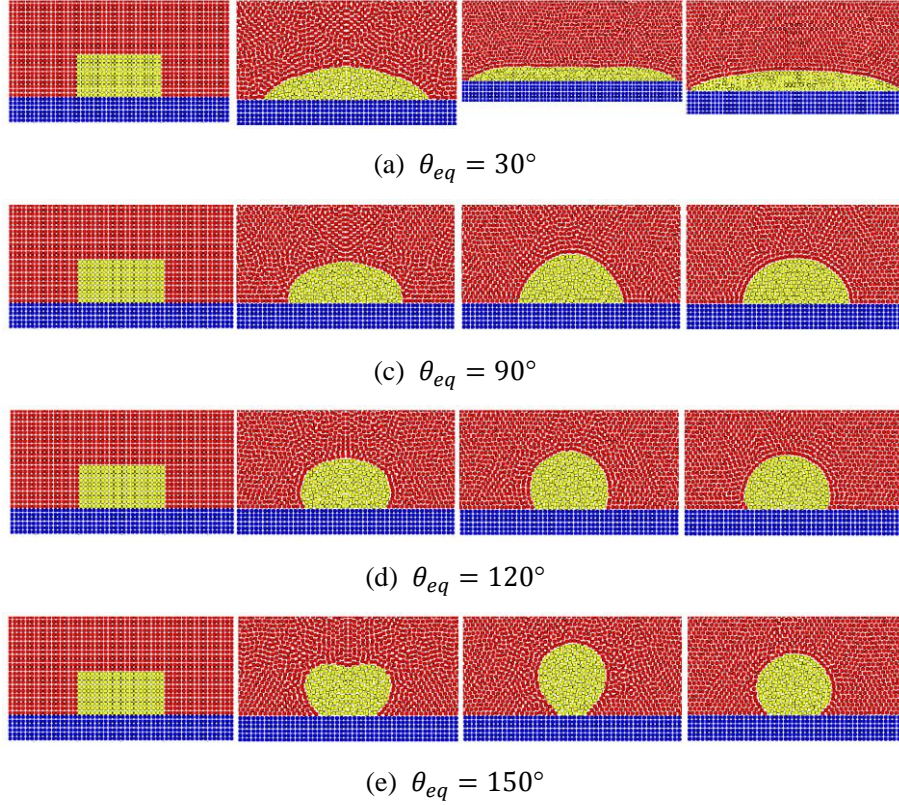


Figure12. Evolution of liquid droplet on a smoothed surface with various contact angles

Then, various contact angles from 30° to 150° are tested. As shown in Figure 12, all presented cases are given the same initial configurations except the input contact angles.

4.2 2D Droplet morphology on the smooth surface

In this section, we first validate the proposed simulation method by considering the morphology of a 2D droplet deposited on the smooth surface.

The analytical solution of 2D droplet morphology had been derived by Liu and Xia^[36] based on the classical Laplace equation across the liquid-gas interface. As shown in Figure 1, the radius of the liquid-solid area is defined as a , the maximum height of the droplet is h_{drop} . The boundary conditions of the semi-droplet are expressed as:

$$y(0) = h_{drop}, y'(0) = h_{drop}, \quad (22)$$

$$y(a) = 0, y'(a) = -\tan \theta_{eq}, \quad (23)$$

The volume of the droplet is given as

$$S = \int_{-a}^a y dx, \quad (24)$$

The analytical solution of the dimensionless morphology for the droplet can be obtained from Eqs. (24) and (25) ^[36], in which the effects of gravity, surface tension, hydrostatic pressure and transversality condition are considered.

$$\kappa x = 2[E(k_0, \varphi) - E(k_0, \varphi_0)] - [F(k_0, \varphi) - F(k_0, \varphi_0)], \quad (25)$$

$$\kappa y = c_0 \kappa^{-1} - 2k_0 \cos \varphi, \quad (26)$$

where $\kappa^{-1} = \sqrt{\sigma/(\rho g)}$ is the capillary length, x, y are the coordinates of droplet surface, $E(k_0, \varphi)$ and $F(k_0, \varphi)$ are the elliptic integrals of the first and second kinds, c_0 is the curvature at the triple contact point, k_0 is the variable which is determined according to the boundary condition. In the Eqs. (24) and (25), the droplet morphology is expressed as the function of the variable φ , while the variable k_0 can be computed from the boundary condition. Parameters including θ_{eq} , c_0 , and κ^{-1} are taken as input parameters.

In this section, a range of simulations are carried out by changing the input conditions of droplet volume and contact angle θ_{eq} . The volume of the droplet can be expressed as $l_x \times l_y$. When the simulation reaches quasi-static state, *i.e.*, the droplet shape changes slightly with time, the droplet morphology is extracted from the particle distribution. As an example, Figure 13 shows the particle distributions of the fluid system when the droplet is in equilibrium state, and the analytical solution of liquid-gas interface is also plotted in the Figure.

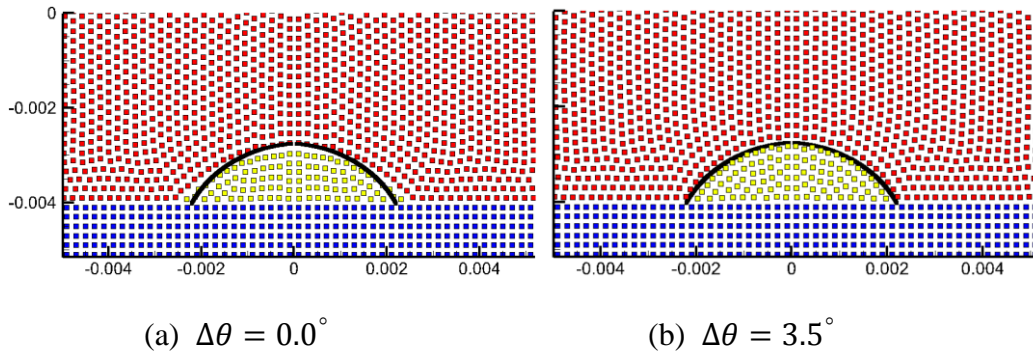


Figure 13 Particle distribution at the equilibrium contact angle. Solid line represents the analytical solution of droplet morphology ($\kappa^{-1} = 0.0027$, $l_x \times l_y = 2.6 \times 1.3 \text{ mm}$, $\theta_{eq} = 60^\circ$)

If the input contact angle (θ'_{eq}) is set equal to the desired contact angle (θ_{eq}), as shown in Figure 13(a), the simulated contact angle is always smaller than the desired value, so that the predicted droplet morphology is not accurately consistent with the analytical result. It indicates that the adjusting force is inadequate to accurately calibrate the contact line. Therefore, we introduce a correction angle to correct the input value of contact angle as:

$$\theta'_{eq} = \theta_{eq} + \Delta\theta, \quad \Delta\theta > 0^\circ \quad (27)$$

where $\Delta\theta$ is the correction angle. θ'_{eq} is the actual input contact angle substituting θ_{eq} . We investigate several cases for various contact angles from 30deg to 150deg, and find that accurate results of the simulated contact angles can be obtained by setting the $\Delta\theta$ between

3° and 5° .

As shown in Figure 13(b), the predicted result of droplet shape is match every well with the analytical result on the condition that the equilibrium contact angle is accurately implemented. Figure 14 shows the comparison of droplet outlines between analytical results and SPH results. It can be seen from Figure 14 that there are good agreements of droplet outlines between SPH results and analytical solutions.

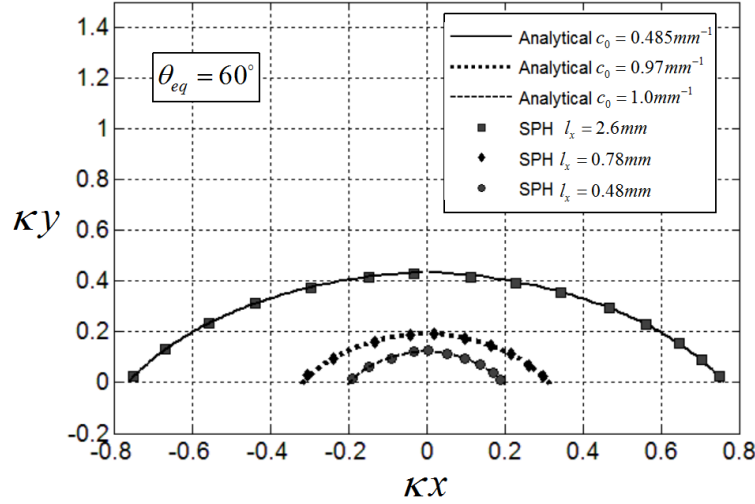


Figure 14 Comparison of droplet morphology between SPH results and analytical results ($\kappa^{-1} = 0.0027$)

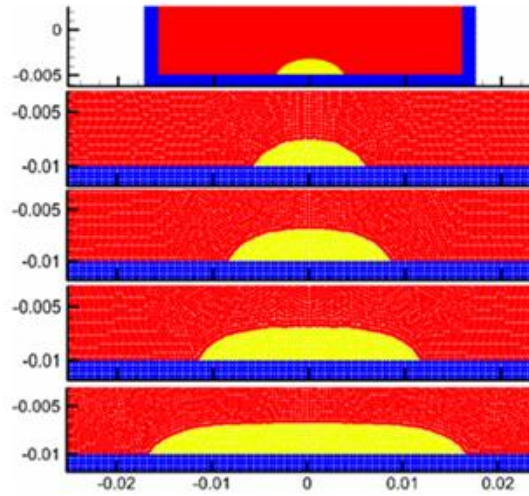


Figure 15 The shape of the droplets changes as the size gets larger ($\theta_{eq} = 60^\circ, \kappa^{-1} = 0.0027\text{m}$)

Figure 15 shows the predicted droplet morphologies for various droplet volumes and constant contact angle of 60° . As shown in the Figure, when the volume of the droplet is small, the curvature of the top of the droplet is relatively large, and the radius of curvature is small, so that the droplet shape is close to a sphere. As the volume of the droplet gradually increases, its shape gets far away from the sphere.

4.3 Droplet morphology on a solid surface with sinusoidal microstructures

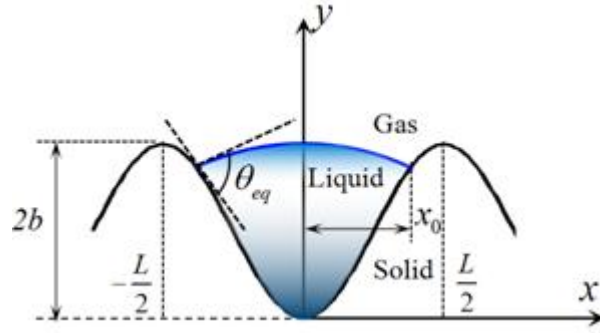


Figure 16 A droplet on a sinusoidal substrate

This section aims to test the applicability of the contact angle method on curved surfaces. As an example, we analyze a single droplet on a sinusoidal surface, as shown in Figure 16. The substrate surface shape is expressed as:

$$y(x) = b[1 - \cos(qx)] \quad (28)$$

where $2b$ is the roughness amplitude, $q = 2\pi/L$ the wave number, and L the wavelength.

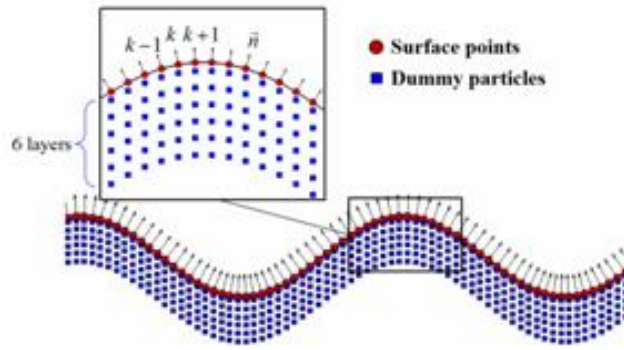


Figure 17 Dummy particles for the rough surface substrate

For the simulation of solid boundary with curved surface substrate, as shown in Figure 17, a set of uniformly distributed surface points are created along the surface first. The surface vectors at an arbitrary point k can be calculated by following equations:

$$\mathbf{n}_k = (n_x, n_y) = \left(-\frac{y_{k+1} - y_{k-1}}{|\mathbf{x}_{k+1} - \mathbf{x}_{k-1}|}, \frac{x_{k+1} - x_{k-1}}{|\mathbf{x}_{k+1} - \mathbf{x}_{k-1}|} \right) \quad (29)$$

$$\boldsymbol{\tau}_k = (\tau_x, \tau_y) = \left(\frac{x_{k+1} - x_{k-1}}{|\mathbf{x}_{k+1} - \mathbf{x}_{k-1}|}, \frac{y_{k+1} - y_{k-1}}{|\mathbf{x}_{k+1} - \mathbf{x}_{k-1}|} \right) \quad (30)$$

where $k, k-1, k+1$ represent surface points of the solid surface adjacent to each other, $\mathbf{n}_k, \boldsymbol{\tau}_k$ are the normal unit vector and tangential unit vector at point k , \mathbf{x}_k is position vector of the point k .

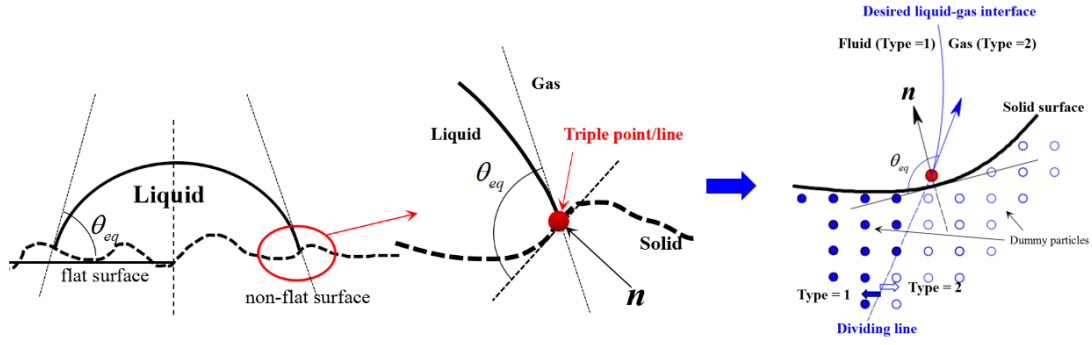


Figure18. The treatment of solid boundary for a curved surface

Let us first consider the case where the droplet is deposited in the single groove. The droplet is still initialized as a square lump, similar to the way we did for smooth surface. The domain is $0.02 \times 0.005\text{m}$ in the rectangular region using a fixed particle initial spacing $d_{\text{ini}} = 0.0001$.

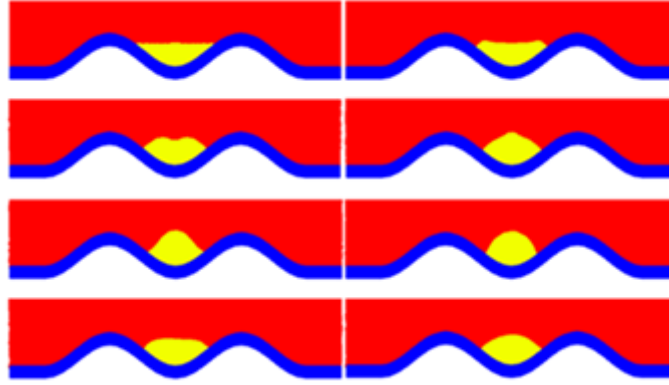


Figure 19 Evolution of liquid droplet in the single groove ($b = 1.0\text{mm}$, $L = 8\text{mm}$, $l_x = 5.2\text{mm}$, $\theta_{eq} = 80^\circ$)

As shown in Figure 18, the droplet is initially in non-equilibrium state. It begins to deform under the action of gravity, surface tension, and contact angle adjusting force. The surface tension causes the sharp corners of the droplet shape to get smoothed. The contact angle adjusting force acts on the fluid near the triple line, causing the droplet to evolve toward the desired contact angle.

As shown in Figure 19, the steady-state morphology of droplets are obtained using different input contact angles in case b, and the color distributions of the dummy particles at steady state are also shown. Consider a droplet located on a rough or curved solid whose characteristic size of surface roughness is of the same order as the droplet size. Then the model can be applied to simulate the droplet morphology. For example, some results of droplet shapes on the sinusoidal surface of case b are plotted in Figure 25. The wavelength L of the roughness is taken as $b = 1\text{mm}$ and the contact angle $\theta_{eq} = 60^\circ$. The color distribution of the dummy particle is also given at each moment, which shows that and the color distribution of dummy particle color changes with the movement of the droplet.

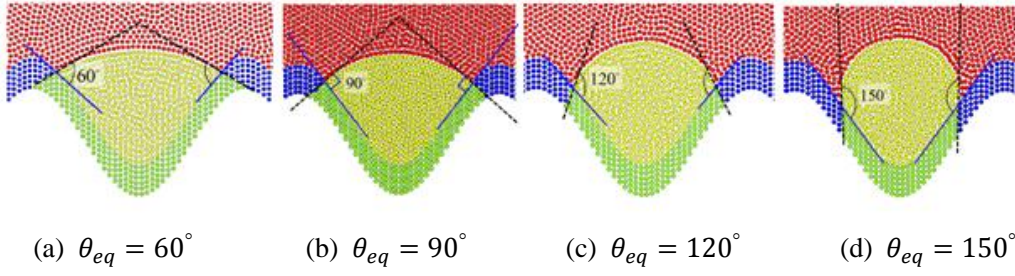


Figure 20 Color distributions of dummy particles for various contact angles ($b = 1\text{mm}$, $L = 3.63\text{mm}$)

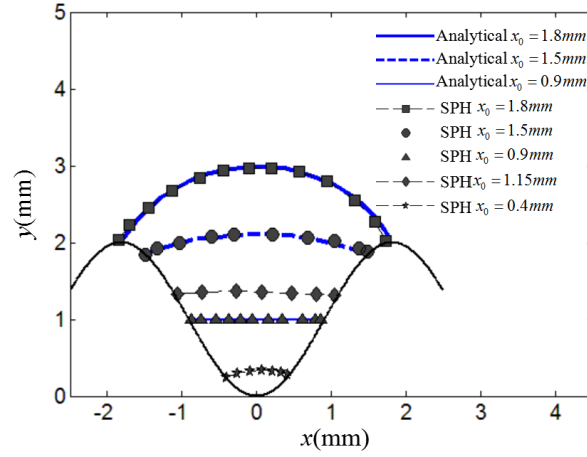


Figure 21 Comparison of droplet morphology between SPH results and analytical results ($\kappa^{-1} = 0.0027$)

For validation, some solutions of droplet shapes on a sinusoidal hydrophilic surface are plotted in Figure 21 for the wavelength $L = 3.63\text{mm}$, the height of the roughness $a = 1\text{mm}$, and the contact angle $\theta_{eq} = 60^\circ$. The analytical solutions from the reference [35] are also plotted in Figure 21. It shows the SPH results are in good agreement with the analytical solutions for three comparative cases of $x_0 = 1.8\text{mm}$, 1.5mm , and 0.9mm . It also can be seen from the Figure that the droplet shape is sensitive to the characteristic size of roughness and the contact angle.

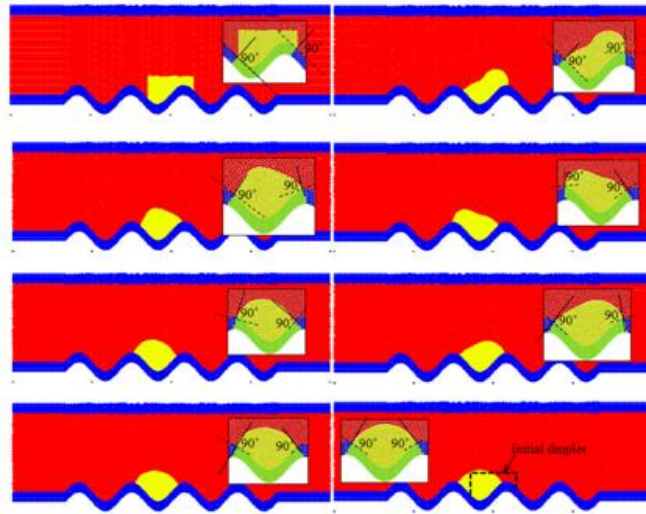


Figure 22 Evolution of liquid droplet on a non-smoothed surface ($b = 0.5\text{mm}$,

$$L = 3.63\text{mm}, \theta_{eq} = 90^\circ)$$

As shown in Figure 22, the liquid lump is initially placed offset from the center of the groove. Under the effects of several forces, the liquid lump move to the left as a whole until it arrives at the center of the groove. During this process, the droplet with smooth surface is gradually formed. The adjusting force calibrates the distribution of fluid near the triple line. Finally, the droplet is settled down in the single groove with the desired contact angle $\theta_{eq} = 90^\circ$.

Above results show that the proposed contact angle method is suitable for the rough surface whose characteristic size is of the same order as the droplet size.

5. Summary

In this paper, based on the SPH multiphase flow algorithm and the CSF surface tension model, a method of implementing contact angle is proposed to simulate the droplet morphology on smooth surface and roughness surface substrates. The results can be summarized as follows:

1. The virtual interface method is proposed to implement contact angle in SPH multiphase simulation. It requires two critical steps: (a) Identifying the triple contact points through searching the boundary particles of each fluid phase; (b) Dividing the dummy particles into two types by the desired contact line. Then, the virtual liquid-gas interface can be created from the triple contact point to the inside of solid wall.

2. The adjusting force is defined as the function of interface curvature and normal unit vector. By means of the virtual liquid-gas interface, the adjusting force which is distributed into the fluid particles near the triple line, always tends to calibrate the dynamic line towards the equilibrium contact angle.

3. The proposed contact angle method is validated for simulations of contact angles on a smooth surface and a roughness surface with sinusoidal microstructures. A initially square liquid lump can evolve towards the equilibrium droplet morphology with the desired contact angle. The predicted droplet shapes are in good agreement with the analytical solutions. In future work, the contact angle scheme for three dimensional model will be considered.

Reference

- 1 de Gans B J, Duineveld P C, Schubert U S. Inkjet printing of polymers: state of the art and future developments[J]. *Advanced materials*, 2004, 16(3): 203-213.
- 2 Goldszal A, Bousquet J. Wet agglomeration of powders: from physics toward process optimization[J]. *Powder Technology*, 2001, 117(3): 221-231.
- 3 Squires T M, Quake S R. Microfluidics: Fluid physics at the nanoliter scale[J]. *Reviews of modern physics*, 2005, 77(3): 977.
- 4 Sun M, Luo C, Xu L, et al. Artificial lotus leaf by nanocasting[J]. *Langmuir*, 2005, 21(19): 8978-8981.
- 5 Pate J S, Atkins C A, White S T, et al. Nitrogen nutrition and xylem transport of nitrogen in ureide-producing grain legumes[J]. *Plant Physiology*, 1980, 65(5): 961-965.
- 6 Gao X, Jiang L. Biophysics: water-repellent legs of water striders[J]. *Nature*, 2004, 432(7013): 36-36.
- 7 Hirt C W, Nichols B D. Volume of fluid (VOF) method for the dynamics of free boundaries[J]. *Journal of computational physics*, 1981, 39(1): 201-225.
- 8 Liu G R, Liu M B. Smoothed particle hydrodynamics: a meshfree particle method[M]. *World Scientific*, 2003.
- 9 Gingold R A, Monaghan J J. Smoothed particle hydrodynamics: theory and application to non-spherical stars[J]. *Monthly notices of the royal astronomical society*, 1977, 181(3): 375-389.
- 10 Monaghan J J. Simulating free surface flows with SPH[J]. *Journal of computational physics*, 1994, 110(2): 399-406.
- 11 Randles P W, Libersky L D. Smoothed particle hydrodynamics: some recent improvements and

- applications[J]. Computer methods in applied mechanics and engineering, 1996, 139(1-4): 375-408.
- 12 Hu X Y, Adams N A. A multi-phase SPH method for macroscopic and mesoscopic flows[J]. Journal of Computational Physics, 2006, 213(2): 844-861.
- 13 Ming F R, Sun P N, Zhang A M. Numerical investigation of rising bubbles bursting at a free surface through a multiphase SPH model[J]. Meccanica, 2017: 1-20.
- 14 Zhang A, Sun P, Ming F. An SPH modeling of bubble rising and coalescing in three dimensions[J]. Computer Methods in Applied Mechanics and Engineering, 2015, 294: 189-209.
- 15 Grenier N, Antuono M, Colagrossi A, et al. An Hamiltonian interface SPH formulation for multi-fluid and free surface flows[J]. Journal of Computational Physics, 2009, 228(22): 8380-8393.
- 16 Colagrossi A, Landrini M. Numerical simulation of interfacial flows by smoothed particle hydrodynamics[J]. Journal of computational physics, 2003, 191(2): 448-475.
- 17 Adami S, Hu X Y, Adams N A. A new surface-tension formulation for multi-phase SPH using a reproducing divergence approximation[J]. Journal of Computational Physics, 2010, 229(13): 5011-5021.
- 18 Szewc K, Pozorski J, Minier J P. Simulations of single bubbles rising through viscous liquids using smoothed particle hydrodynamics[J]. International Journal of Multiphase Flow, 2013, 50: 98-105.
- 19 Yeganehdoust F, Yaghoubi M, Emdad H, et al. Numerical study of multiphase droplet dynamics and contact angles by smoothed particle hydrodynamics[J]. Applied Mathematical Modelling, 2016, 40(19): 8493-8512.
- 20 Nugent S, Posch H A. Liquid drops and surface tension with smoothed particle applied mechanics[J]. Physical Review E, 2000, 62(4): 4968.
- 21 Tartakovsky A, Meakin P. Modeling of surface tension and contact angles with smoothed particle hydrodynamics[J]. Physical Review E, 2005, 72(2): 026301.
- 22 Brackbill J U, Kothe D B, Zemach C. A continuum method for modeling surface tension[J]. Journal of computational physics, 1992, 100(2): 335-354.
- 23 Morris J P. Simulating surface tension with smoothed particle hydrodynamics[J]. International journal for numerical methods in fluids, 2000, 33(3): 333-353.
- 24 Liu M B, Liu G R. Meshfree particle simulation of micro channel flows with surface tension[J]. Computational Mechanics, 2005, 35(5): 332-341.
- 25 Hu X Y, Adams N A. A multi-phase SPH method for macroscopic and mesoscopic flows[J]. Journal of Computational Physics, 2006, 213(2): 844-861.
- 26 Das A K, Das P K. Equilibrium shape and contact angle of sessile drops of different volumes—Computation by SPH and its further improvement by DI[J]. Chemical Engineering Science, 2010, 65(13): 4027-4037.
- 27 Wang Z B, Chen R, Wang H, et al. An overview of smoothed particle hydrodynamics for simulating multiphase flow[J]. Applied Mathematical Modelling, 2016, 40(23): 9625-9655.
- 28 Breinlinger T, Polfer P, Hashibon A, et al. Surface tension and wetting effects with smoothed particle hydrodynamics[J]. Journal of Computational Physics, 2013, 243: 14-27.
- 29 De Gennes P G, Brochard-Wyart F, Quéré D. Capillarity and wetting phenomena: drops, bubbles, pearls, waves[M]. Springer Science & Business Media, 2013.
- 30 Morris J P, Fox P J, Zhu Y. Modeling low Reynolds number incompressible flows using SPH[J]. Journal of computational physics, 1997, 136(1): 214-226.
- 31 Crespo A J C, Gómez-Gesteira M, Dalrymple R A. Boundary conditions generated by dynamic particles in SPH methods[J]. CMC-TECH SCIENCE PRESS-, 2007, 5(3): 173.
- 32 Dilts G A. Moving least-squares particle hydrodynamics II: conservation and boundaries[J]. International Journal for Numerical Methods in Engineering, 2000, 48(10): 1503-1524.
- 33 Marrone S, Colagrossi A, Antuono M, et al. An accurate SPH modeling of viscous flows around bodies at low and moderate Reynolds numbers[J]. Journal of Computational Physics, 2013, 245: 456-475.
- 34 Monaghan J J, Rafiee A. A simple SPH algorithm for multi-fluid flow with high density ratios[J]. International Journal for Numerical Methods in Fluids, 2013, 71(5): 537-561.
- 35 Liu J L, Feng X Q, Yu S W. Morphology of liquid drops and thin films on a solid surface with sinusoidal microstructures[J]. Acta Mechanica Sinica, 2006, 22(4): 315-322.
- 36 Jian-Lin, Liu. A unified analysis of a micro-beam, droplet and CNT ring adhered on a substrate: Calculation of variation with movable boundaries[J]. Acta Mechanica Sinica, 2013, 29(1): 62-72.

Acknowledgements

This project was supported by China Postdoctoral Science Foundation (Grant NO. 2017M622307). This project was also partially supported by Shandong natural science foundation (Grant NO. ZR201709210320) and Fundamental Research Funds for the Central

Universities (Grant No. 18CX02153A). I also want to give my thanks to Dr. Li-Chenxue, who inspires me to grow up, mentally and intellectually.

Relationship between Train Speed and Aural Discomfort in Tunnels Based on Tympanic Membrane Model

***Pengpeng Xie¹, †Yong Peng², Shengen Yi³, Junjiao Hu⁴**

¹Joint International Research Laboratory of Key Technology for Rail Traffic Safety, Central South University, China.

³ Research Laboratory of Hepatobiliary Diseases General Surgical Department, The Second Xiangya Hospital, Central South University, China

⁴ Department of Radiology, The Second Xiangya Hospital, Central South University, China

*Presenting author: paulxie14@csu.edu.cn

†Corresponding author: yong_peng@csu.edu.cn

Abstract

Inferior pressure changes of high-speed train have been disturbing aural feelings of passengers and staff when trains are passing through tunnels or meeting in tunnels. However, it is problematic to quantify the extent of aural discomfort according to the recorded inferior pressure data. The paper presents an original methodology to assess humans' aural discomfort based on the mechanics of a reconstructed tympanic membrane(TM) finite element model. Besides, viscoelastic coefficients of a 7-parameter Maxwell model were obtained by inverse problem-solving method. By comparison with the results derived from the dynamic simulation of TM under varying pressure amplitudes and gradients, aural discomfort was ranked into four levels from ideal, good, bad to worse. Meanwhile, displacement, velocity and stress of TM umbo were chosen as referential indicators to establish judgement rules for discomfort assessment. In addition, the pressure change history which was exerted at the lateral side of TM, was collected by field tests when trains are running in tunnels. The results reveal that the maximum displacement of umbo for the four discomfort levels are 35.35 μ m, 44.22 μ m, 63.84 μ m and 99.07 μ m respectively. Furthermore, a set of methodology was established for judgement of aural discomfort. Also, it indicates that aural discomfort begins at the pressure inversion point before which human ears feel pleasant whereas exacerbates when trains running in the middle of tunnels and alleviates when approaching the tunnel exit.

1. INTRODUCTION

Train-tunnel aerodynamic problems have long been bothersome experiences for staff and passengers. For instance, the barometric pressure outside the train fluctuates from 1kPa positive to 1kPa negative when train is passing through tunnel at 300km/h [1]. The varying pressure travels interiorly by path of openings of compartments and may incur different extents of aural discomfort which worsens with increasing running speed and long tunnel conditions. Countries like French, Germany and Japan has enacted air-tightness standards for high-speed trains from either the pressure amplitudes or pressure changing rates. Nonetheless, it poses a challenge for researchers to evaluate the discomfort feelings quantitatively from perspective of human ear biomechanics. Until recently, adequate investigations on human ears, which were mainly focused on mechanic properties, sound transmission and pathology, has been done by virtue of reconstructed finite element models [2-4]. However, using ear models to interpret barometric discomfort hasn't yet been available. TM is the first receptor tissue and

most sensitive to ambient pressure variations. Besides, it is responsible for absorbing pressure waves and converting it into vibration energy. Thus, it is closely linked with aural discomfort and chosen as a study objective in this paper. The purpose of this paper was to reveal the mechanism between the vibration characteristics of TM and aural discomfort, set judgment methodology and make assessment of the interior barometric environment under different running speeds in tunnels.

TM model was reconstructed based on CT scanning data from healthy volunteer with no history of ear diseases. Despite lack of uniformity of material properties [5], viscoelastic constitutive models were the most frequently explored via experiments on TM samples from cadavers [6-9]. Hence, a 7-parameter Maxwell viscoelastic model was utilized and the seven variables were derived from inverse problem-solving method and applied to our TM model, through which the model was validated. Furthermore, by exerting pressure loads at the lateral side of TM, the dynamic responses were simulated under varying pressure amplitudes and gradients. It deserves attention that the simulation conditions were conformed to Japan's air-tightness tests which uncovers the relationship between pressure changes and tinnitus for train cabin design [10]. Meanwhile, displacement, velocity and stress of TM umbo were selected as three indicators to represent its dynamics. Also, aural discomfort was divided into four levels ranging from easeful to awful by comparisons of the simulation results. In addition, interior pressure change history was collected by on-board field tests when train runs at a speed interval of 180~250km/h. The interior pressure then was loaded on TM surface and the aural discomfort it induced was analyzed eventually.

2. MATERIAL AND METHOD

2.1 TM model reconstruction

Three dimensional TM model was reconstructed in Mimics as shown in Fig.1 according to the MRI scanning data derived from healthy adult volunteer (Male, 24 years old) with no history of ear diseases. The modelling work was assisted by the collaboration with physician who has expert knowledge of the middle ear anatomy and radiation morphology, the reconstructed TM model was shown in Fig.2.



Fig.1 TM morphology and reconstruction in Mimics

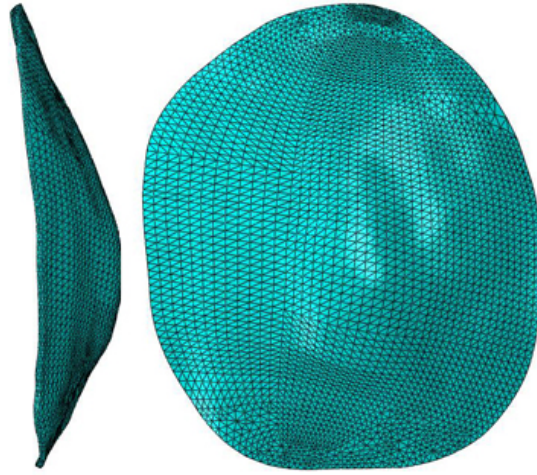


Fig.2 TM finite element model in anterior and medial view

TM is anatomically composed of three regions which are tympanic annulus (TA), pars flaccida (PF) and pars tensa (PT). PT is mostly surrounded by TA at the periphery while separated by PF at the superior of TM [11]. As is illustrated in Fig.2, TM model eliminates TA and PF on the ground that PF has slight influence on the simulation results and TA is fully clamped in simulation [6,9,12]. The thickness of TM is about 0.1mm on average and gauges 7.91mm and 9.48mm laterally and longitudinally with a cone depth of 1.54mm. Due to the non-uniform thickness distribution of TM, quadratic tetrahedral element with 10 nodes was applied and the nodes along the edge of TM was fully clamped.

2.2 Material properties for TM model

Although TM property is widely discussed [13-14], there hasn't been an agreement on which kind of constitutive model TM should be used. For biological tissue or organ, it is generally considered as both elastic and viscous. Various experiments on cadaver TMs were done to measure the viscoelastic coefficients as well as simulation methods. We utilized inverse problem-solving method to harvest the coefficients in time domain by comparison with the experimental data [9]. First, a TM strip, which gauges 1.5mm in width and 4mm in length with a uniform thickness of 0.1mm, was modeled as well as an indentation needle with a radius of 0.15mm, as shown in Fig.3.

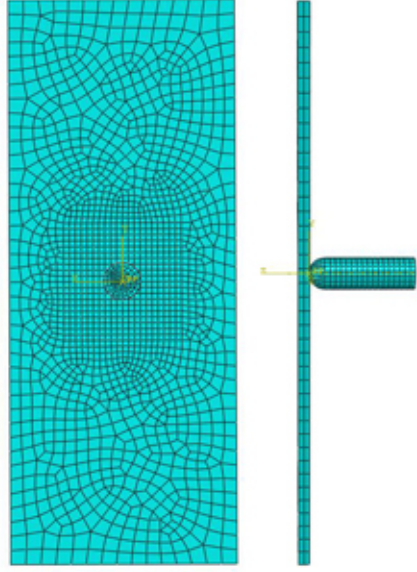


Fig.3 Indentation of needle on TM strip model

The bottom and upper surfaces of TM slice were fully clamped and the needle moves to contact TM sinusoidally at 0.2Hz. Meanwhile, a 7-parameter Maxwell model was implemented to simulate TM's viscoelasticity. Besides, genetic algorithm was employed to obtain optimal viscoelastic coefficients by controlling cost function as shown in equation 1 within 4%.

$$C = \sum_{i=1}^N \left\{ F_{sim} \left[d_i, \sum_{j=1}^3 (g_j e^{-t/\tau_j} + g_{\infty}) \right] - F_{exp}(d_i) \right\} \quad (1)$$

2.3 Simulations on TM model

In accordance with Japan's airtightness experiments for train [10], a series of pressure loads with varying amplitudes and gradients were created and loaded on TM surface. Furthermore, the aural discomfort was divided into four levels ranging from ideal, good, bad to worse, among which ideal level represents no hostile feelings on human ears, good level acceptable aural disturbance, bad level annoying and worse level awfully uncomfortable. Ideal level covers three conditions including 0.1-1, 0.4-0.5 and 0.3-0.5, of which the former number is pressure gradient(kPa/s) and the latter pressure amplitude(kPa), two for good level (0.2-1, 0.3-1), three for bad level (0.5-0.5, 0.4-1, 0.1-2) and four for worse level (0.1-3, 0.2-3, 0.3-3, 0.4-3). All the loads were uniformly distributing at TM surface and increased linearly. Besides, displacement, velocity and stress at umbo were selected as indicators on behalf of TM vibration with which aural discomfort level was linked.

2.4 Interior pressure data collection

Couples of on-board tests were done to record the interior pressure changes covering running speed levels of 180 km/h, 200 km/h, 220 km/h, 250 km/h and 350km/h together with trains' meeting in tunnels. Fig.4 illustrates the pressure change history at different speeds.

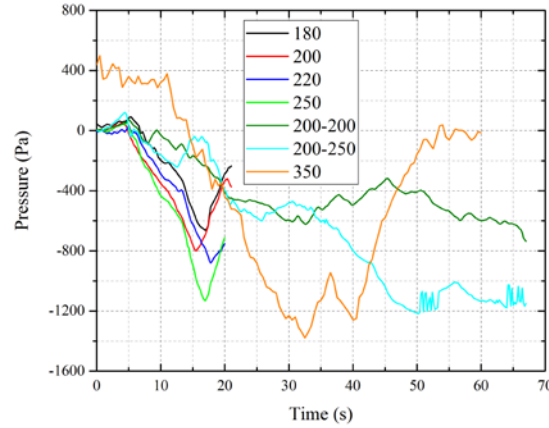


Fig.4 Interior pressure change history at different speed levels and meeting in tunnels

It should be pointed out that experiments were done in various types of tunnels but in the same train. Tests under 250km/h were completed in the same tunnel with a length of 1.05km, tests at 350km/h were done in a 5.95km-long tunnel and trains meeting tests done in 3.71km-long tunnel. For single train running in tunnel, it is visible that the interior pressure at initial stage after the train enters is positive followed by a long-term negative pressure stage and ascends when train approaches tunnel exit. With respect to trains meeting in tunnel, the interior pressure keeps negative during the whole journey. Likewise, the recorded pressure data were exerted at the lateral side of TM whose displacement, velocity and stress at umbo position were output for aural discomfort assessment.

3. RESULTS

3.1 Viscoelastic parameters for TM

Optimization achieved convergence when the cost function arrived at a value of 2.53%. The relaxation curve was displayed in Fig.5 compared with published data [9].

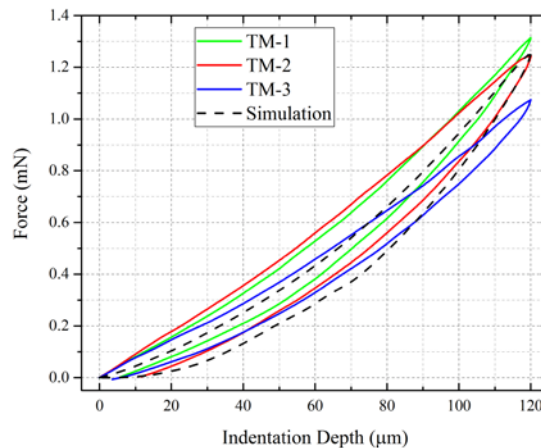


Fig.5 Comparison of relaxation curve derived from simulation with experiments

From Fig.5, it can be observed that the relaxation curve of TM shows significant viscoelastic property. The dashed simulation curve at the loading stage moves continuously within the experimental curves while at the unloading stage it goes beyond the experimental curves at the last two thirds stage. The discrepancy between simulation and experiments was considered reasonable in the light of individualities of human ears. To further justify the validity of our TM model, Fig.6 demonstrated the time-dependent relaxation modulus.

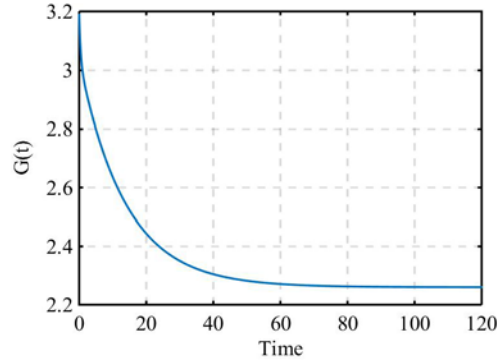


Fig.6 Reduced relaxation function of a 7-parameter Maxwell model

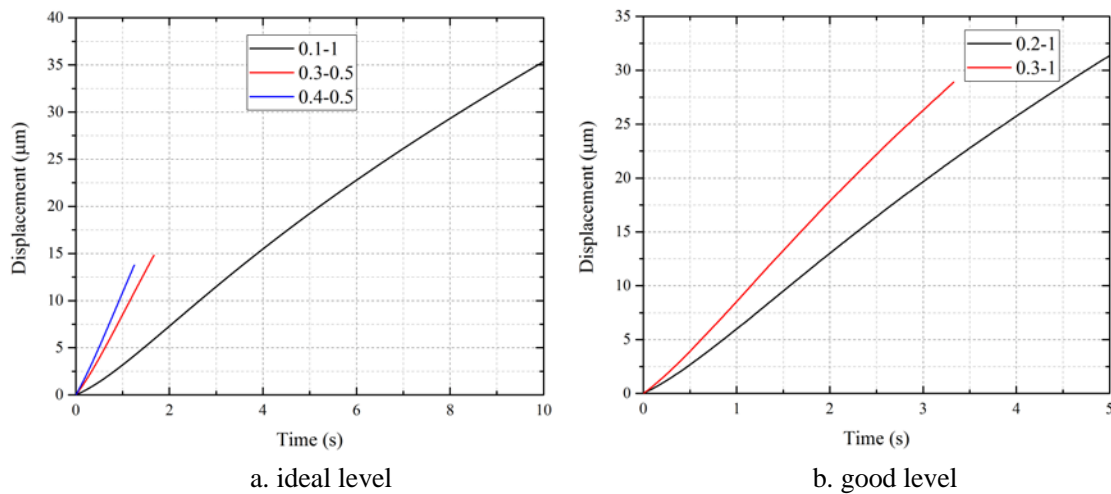
From Fig.8, the average relaxation was described by three decays. The initial decay of 4% occurs at 0.46s followed by a decay of 18% with a characteristic time 10.3s and ends up with a total relaxation of 29% with characteristic time 14.5s. Aernouts et al. has tapped into how TM thickness affects Young's modulus, which revealed that the modulus ranges from 1.74MPa to 2.87Mpa at a thickness of 0.1mm [9]. It can be seen that our steady modulus falls right in this range with a value of 2.26MPa. Hence, the simulation curve is regarded as effective despite of the disparity between simulation and experiments. Detailed viscoelastic coefficients for the model are tabulated in Table 2 below.

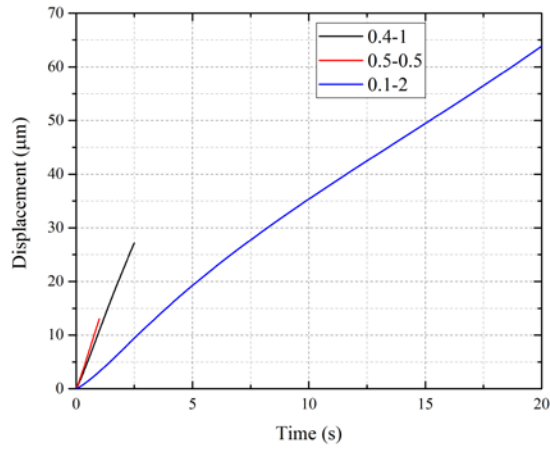
Table 2 Parameter values for the 7-parameter Maxwell model

	g_1	τ_1	g_2	τ_2	g_3	τ_3	g_∞
TM	0.154	0.46	0.103	10.3	0.672	14.5	2.26

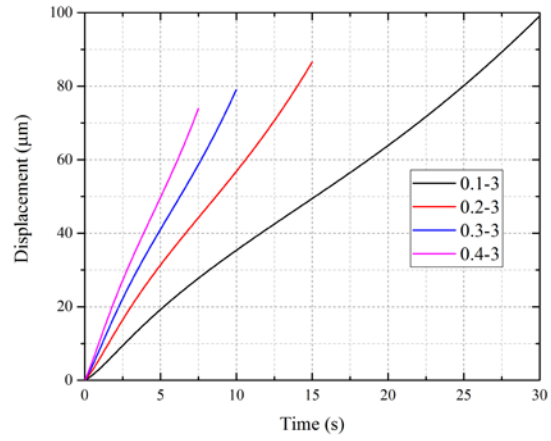
3.2 Judgement methodology for aural discomfort

Japan's airtightness tests for train has uncovered a discomfort line which divides aural discomfort into agreeable area and disturbing area. Depending on the experimental findings, we reinterpreted the discomfort line by dynamics of TM. Fig.7 and Fig.9 illustrated the output results of the three indicators under different pressure loads.



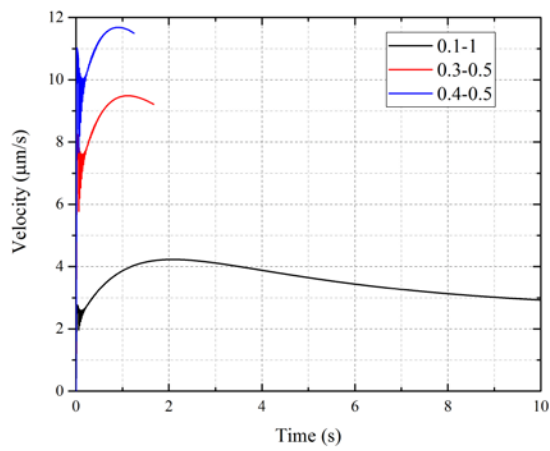


c. bad level

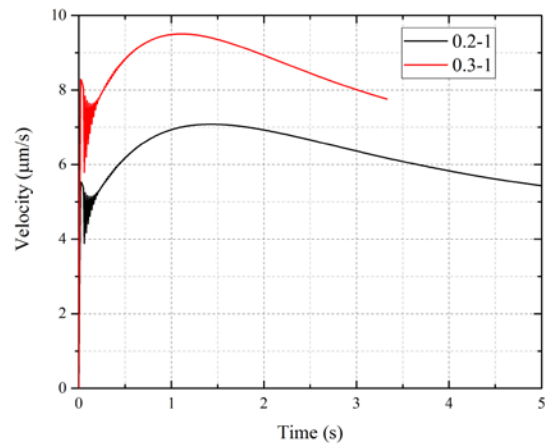


d. worse level

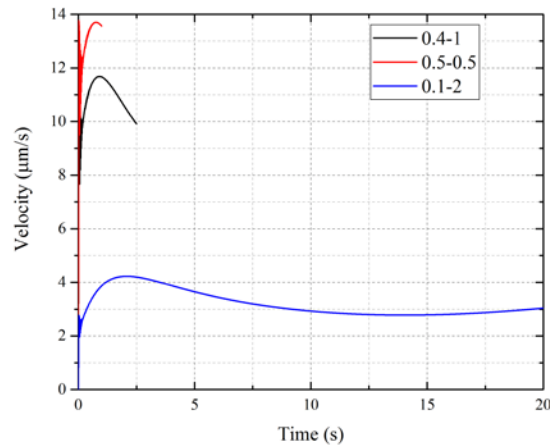
Fig.7 Displacement of umbo under pressure loads from different discomfort levels



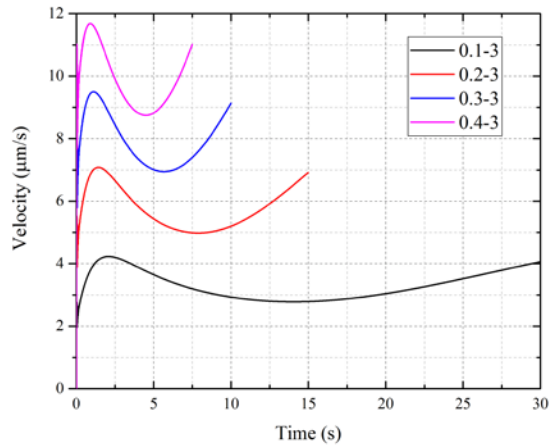
a. ideal level



b. good level



c. bad level



d. worse level

Fig.8 Velocity of umbo under pressure loads from different discomfort levels

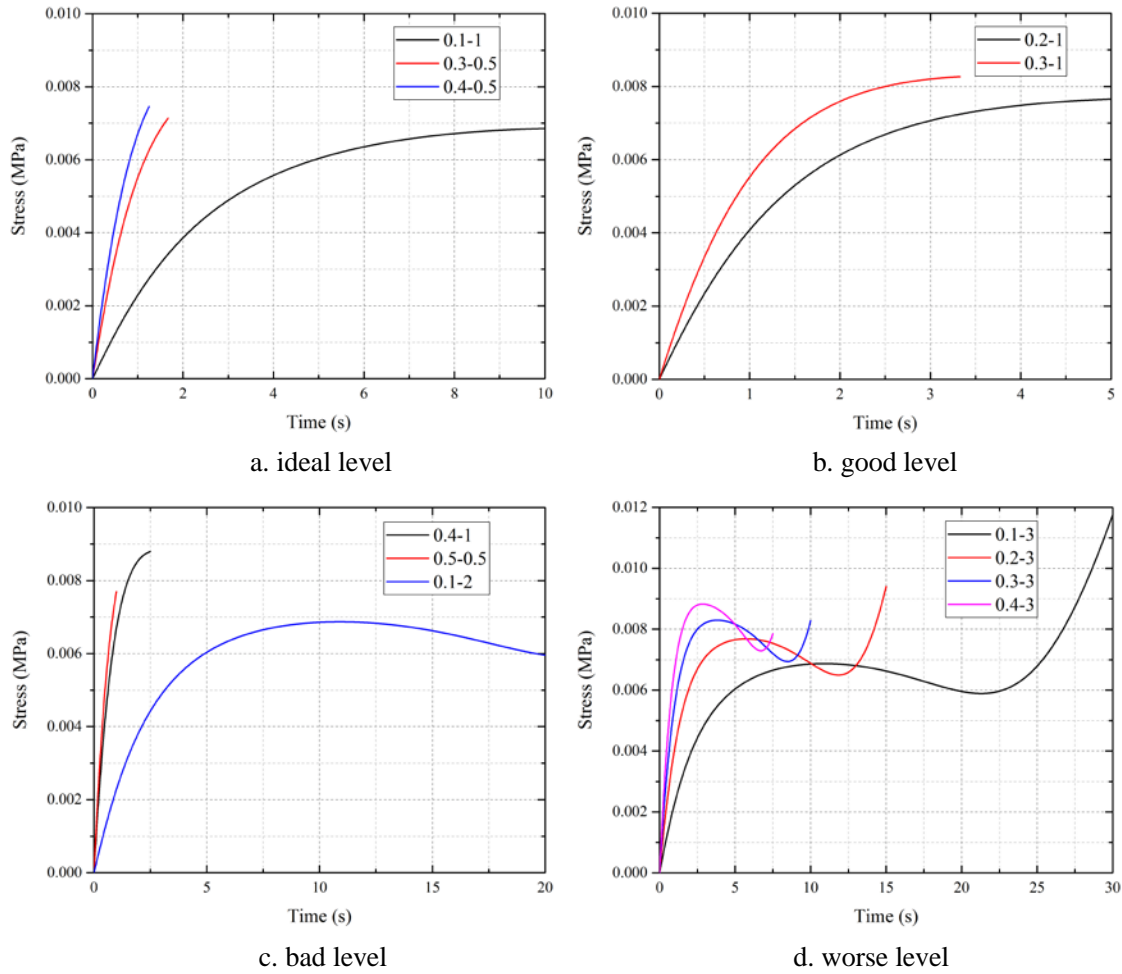


Fig.9 Stress of umbo under pressure loads from different discomfort levels

In Fig.7, it is visible that the displacement of umbo almost increases linearly and high pressure gradients generate steep ascent. The maximum values for each level are 35.35 μ m, 44.22 μ m, 63.84 μ m and 99.07 μ m respectively. In Fig.8, the velocity curves except for worse level lift sharply initially and go up moderately, then either fall or incline to stabilize after reaching amplitudes. Also, it demonstrates that high pressure gradient will produce large velocity value. Likewise, high pressure gradient also generates large stress value if the amplitudes are the same. The maximum values for each level are 0.00746MPa, 0.00827 MPa, 0.00879 MPa and 0.01177 MPa respectively. Besides, the stress of umbo tends to stabilize for long-term loading except for worse level. Moreover, the tendency of these curves justify the viscoelasticity of TM. However, it is improper to judge aural discomfort merely from indicator of velocity because it is found that the maximum velocity in ideal level is very close to that in worse level. Thus, we plotted velocity as independent variable on the X axis and displacement and stress on the Y axis respectively as shown in Fig.10 and Fig.11.

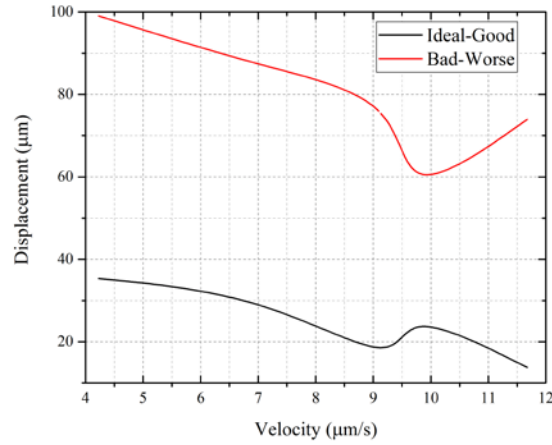


Fig.10 Velocity-Displacement curves for aural discomfort assessment

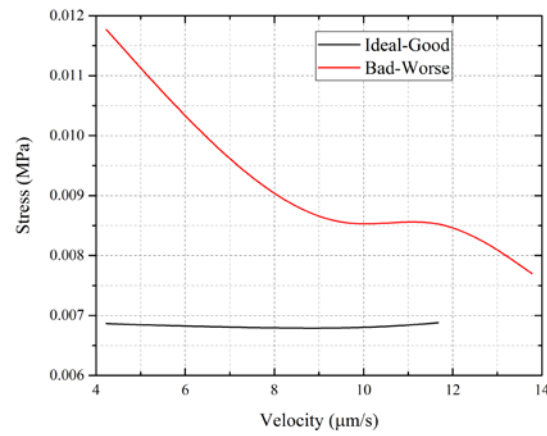
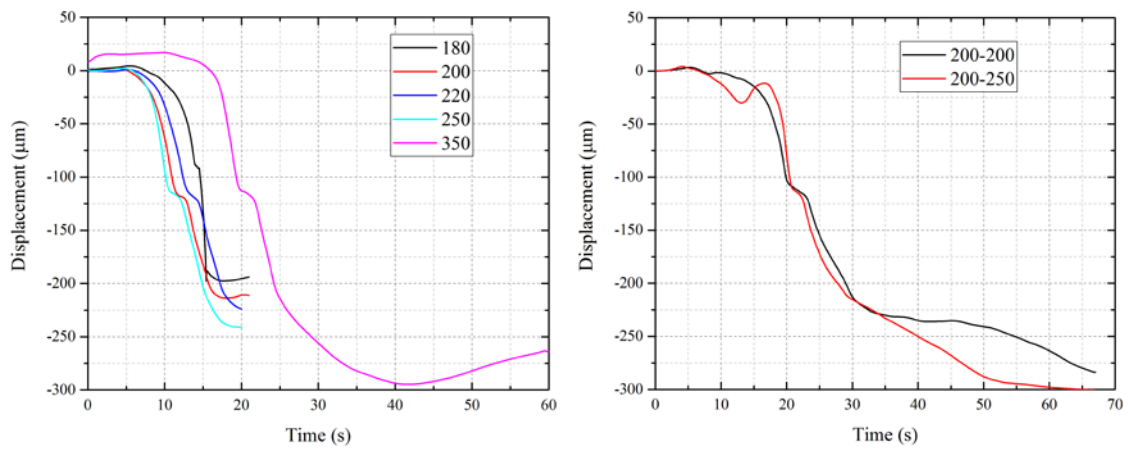


Fig.11 Velocity-Stress curves for aural discomfort assessment

As shown in Fig.10 and Fig.11, the bottom curves represent the threshold between comfort and discomfort, above which ears generate annoyed feelings. The upper curves were plotted by the values derived from bad and worse levels. The bottom curves uncover that the threshold values for indicators of displacement and stress lower when velocity increases.

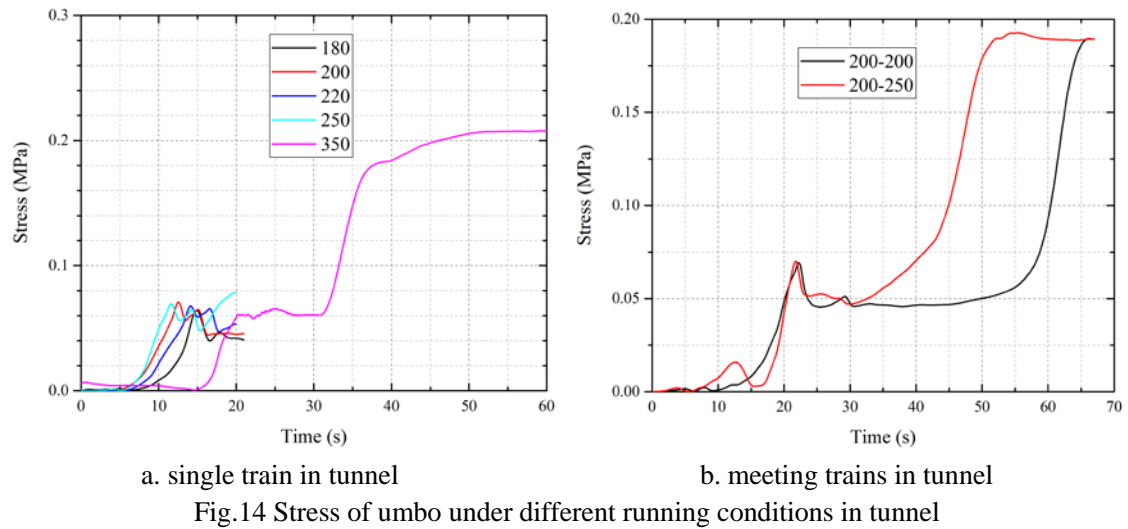
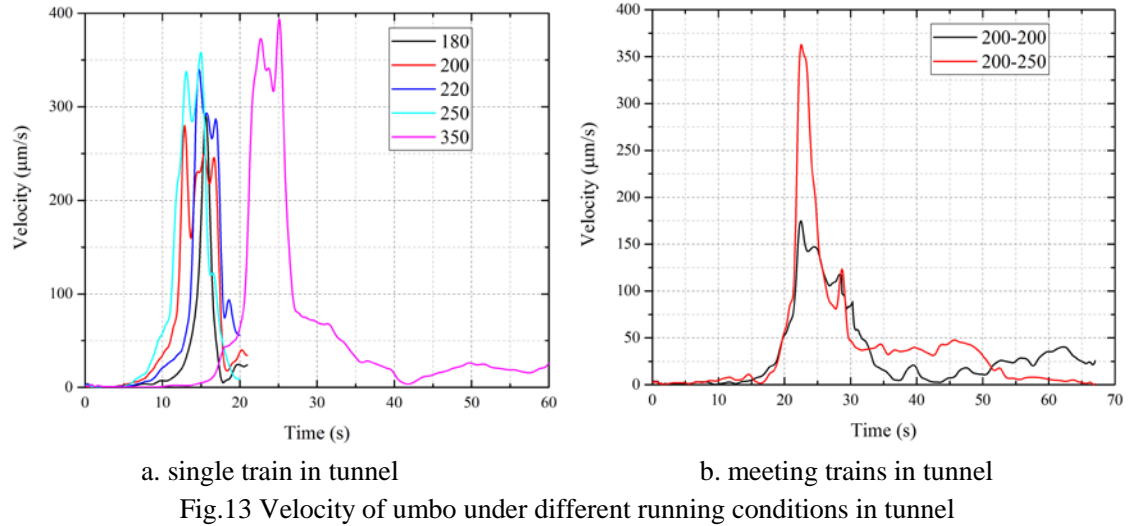
3.3 Aural discomfort assessment of interior pressure



a. single train in tunnel

b. meeting trains in tunnel

Fig.12 Displacement of umbo under different running conditions in tunnel



As is illustrated in Fig.12, it presents that the displacement curves have identical consistency throughout the whole time history. When train runs under 250km/h in tunnel, the displacement of umbo initially keeps at the range of $-5\sim 13\mu\text{m}$ and drops abruptly after 10s until arrives at its minimum and goes steady, while the displacement curve bounces to ascend after reaching the valley when train runs at 350km/h. With respect to trains meeting conditions in tunnel, the displacement curves decrease slightly before 15s and falls sharply afterwards till 35s and 50s separately, after which the curves decrease mildly. In Fig.13, the velocity curves lie under $4\mu\text{m/s}$ at the beginning followed by a sudden ascent and descent, after which they calm down. In Fig.14, the curves also go with a low value and climb steeply to about 0.065MPa. When train runs at 350km/h or trains meet in tunnel, the stress curve climbs again and stabilizes until the train departures. In general, the maximum values of the three indicators increase with the lift of running speed.

4. DISCUSSION

When train enters the tunnel, the air is compressed and propagates in the form of compression wave at sonic speed, which forces the pressure in the tunnel increasing. As the wave arrives at the tunnel exit, a new type of wave generates and returns to tunnel in the form of expansion

wave, which in turn make the pressure fall. Under the alternating effect of the two wave types, the pressure inside the tunnel fluctuates up and down. The altering pressure in the tunnel travels into the cabins by path of openings of the train. Regarding to single train running conditions as is illustrated in Fig.4, the recorded interior pressures stay positive initially followed by a negative stage and go up till the train leaves. Whereas the interior pressure continues to decline after a short stage of positive pressure for meeting conditions. In addition, it indicates that higher running speed brings larger pressure amplitudes either for single train running condition or for trains meeting condition.

Aural discomfort induced by interior pressure fluctuations when train passes through tunnels hasn't been adequately reported until recently. In generally, aural discomfort or ear trauma is mainly attributed to time-dependent transient pressure or frequency-dependent noise [15-20]. As is reported in explosion about ear injury, blast wave generates a short-lived positive overpressure initially and is prolonged by a long-term negative stage [15,21]. Virtually, the recorded interior cabin pressure is much analogical to the blast wave in pressure change characteristics. Nonetheless, blast wave differs in pressure amplitude and duration from interior pressure, whose amplitude is far lower but the time history is much more long-lasting than blast wave. Hence, blast injury is commonly manifested in TM rupture, ossicles dislocation and tinnitus. What deserves attention is that TM penetration is the most frequently diagnosed ear trauma patterns. And it is caused by overpressure in the ear canal, which was verified by otoscopic observations that patients have inverted edges of the penetration [15-16]. On the other hand, barotrauma reported in aviation is also somewhat analogical to train-tunnel effect. Cabin pressure in aircrafts alters with ascent and descent maneuvers, which is accompanied with varied degrees of ear ailments [22-23]. Researches on barotrauma reach a consensus that descent poses much higher risks on passengers than ascent, that is, positive pressure differential between ear canal and tympanic cavity is more likely to cause ear complaints than equal amount of negative pressure differential [24]. Mirza et al uncovered that ET opens passively to expel excessive pressure in the tympanic cavity when the pressure differential reaches a threshold of 15mmHg (2kPa), but ET may stops working if middle pressure is lower than ambient pressure [25]. These findings on ear complaints caused by blast and aviation offers important referential values for assessment of tunnel-induced aural discomfort.

As is illustrated in Fig.4, the lowest amplitude of the interior pressure is -1.4kPa which isn't enough to force ET opening. Hence, it is futile to balance the pressure differential between sides of TM neither by opening of ET itself nor by any physical maneuvers like swallow or yawning. It can be interpreted that human ears have to confront with the hostile barometric environment without any conscious actions of relief unless air tightness of train is improved. Regarding to single train in tunnel, the amplitudes of the three indicators lift with the increasing of running speed, which implies higher travelling speed in tunnel incurs more terrible aural feelings. Additionally, the curves of the three indicators display some common characteristics. For short tunnel to travel, there is a peaceful stage lasting for 8 seconds right after the train enters. After this stage, one abrupt change appears representing by sudden drop of displacement and peaks of velocity and stress. During this stage, human ears perceive a significant pressure shock at the first half stage and attenuate at the second half until the train leaves. Whereas for long tunnel, the peaceful stage for ears is prolonged to about 18 seconds

after which also follows with a pressure shock for ears. As the train runs, the aural discomfort continues to worsen until it departs. With respect to trains meeting in tunnel, the former 18 seconds also offers a friendly environment for ears and then ear complaints generate. It is noteworthy that meeting speeds of 200km/h versus 250km/h results in more awful aural feelings than that of 200km/h versus 200km/h. In general, it is conclusive that the whole course of train's passage can be divided into two stages. One is a pleasant process for human ears sustaining for several seconds and the other is dominated by a much unfriendly pressure environment for ears.

Judging from displacement of umbo, it indicates that the TM bulges inward initially and dents outward gradually until the train leaves. As a result of the leverage effect of ossicle bones, the energy received by inner ear is amplified for times [2]. Hence, discomfort feelings perceived by TM can also transfer to inner ear by way of ossicles and the discomfort level lifts as well. Furthermore, it is likely to infer that human ears may generate varying levels of otalgia, vertigo or tinnitus as the curves of the three indicators display. We enquired all the testers after the tests completed. Most of them complained the otalgia when train runs in tunnels, and some reported tinnitus symptoms. Nonetheless, it isn't straightforward that whether tinnitus is caused by interior pressure change or by booming noise or unified effect of both. Further investigations will be conducted to identify the true causations of tinnitus.

5. CONCLUSION

Train's entering into the tunnels will change the previous air flow field and make interior pressure fluctuate as well, which brings varying degrees of aural discomfort. For purpose of finding a way to assess aural discomfort quantitatively, a TM FE model was reconstructed based on volunteers' CT scanning data of temporal bone. To solve the problem of non-uniformity of TM material properties, TM was assumed to be viscoelastic in time domain and to satisfy a 7-parameter Maxwell model with a Poisson ratio of 0.499, because this assumption was verified to be conducive to reach convergence for cost function. Meanwhile, inverse method was employed to obtain the optimal values of the 7 parameters by genetic algorithm. Furthermore, dynamics of TM, which is represented by indicators of displacement, velocity and stress of umbo, was simulated under different barometric conditions which is in accordance with published air tightness tests. Through ranking aural discomfort into four levels and comparing the output curves of the three indicators, a set of methodology was established to judge aural discomfort. On the other hand, a series of on-board tests were conducted to collect the history of interior pressure change when trains pass through tunnels or meet in tunnels at different running speed. The recorded pressure data was exerted on the lateral surface of TM and the simulation results of the indicators were output.

The simulation results indicate that the maximum values for the four levels from ideal to worse are 35.35 μ m, 44.22 μ m, 63.84 μ m and 99.07 μ m for displacement and 0.00746MPa, 0.00817 MPa, 0.00879 MPa and 0.01177 MPa respectively. It also reveals that there exists a peaceful stage during which human ears feels pleasant. However, this sustains 15 seconds at the best and alters with variations of running speed and tunnel shape. A pressure wave shock is experienced for ears right after the peaceful stage terminates and ears begins to perceive awfully disturbed accompanied with otalgia, vertigo or tinnitus. This annoyed feeling lasts as long as the train departs the tunnel exit. It deserves attention that aural experiences differ

significantly with regard to tunnel length, because of differences of interior pressure change history.

ACKNOWLEDGEMENT

This work presented in the paper is supported by the National Natural Science Foundation of China (Grant No. 51405517, U1534210), China Postdoctoral Science Foundation (Grant No. 2015M570691).

REFERENCE

- [1] Ko, Y.Y., Chen, C.H., Hoe, T., Wang, S.T.,2012. Field measurements of aerodynamic pressures in tunnels induced by high speed trains[J]. *Journal of Wind Engineering and Industrial Aerodynamics*100, 19-29.
- [2] Sun, Q., Gan, R.Z., Chang, K.H., Dormer, K.J.,2002. Computer-integrated finite element modeling of human middle ear[J]. *Biomechanics and Modeling in Mechanobiology*1, 109-122.
- [3] Zhao, F., Koike, T., Wang, J., Meredith, R.,2009. Finite element analysis of the middle ear transfer functions and related pathologies[J]. *Medical Engineering & Physics*31, 907-916.
- [4] Aerts, J.R.M., Dirckx, J.J.J.,2010. Nonlinearity in eardrum vibration as a function of frequency and sound pressure[J]. *Hearing Research*263, 26-32.
- [5] Volandri, G., Di, Puccio.F., Forte, P., Carmignani, C.,2011. Biomechanics of the tympanic membrane[J]. *Journal of Biomechanics*44, 1219-1236.
- [6] De Greef D, Aernouts J, Aerts J, et al. Viscoelastic properties of the human tympanic membrane studied with stroboscopic holography and finite element modeling[J]. *Hearing research*, 2014, 312: 69-80.
- [7] Huang G, Daphalapurkar N P, Gan R Z, et al. A method for measuring linearly viscoelastic properties of human tympanic membrane using nanoindentation[J]. *Journal of Biomechanical Engineering*, 2008, 130(1): 014501.
- [8] Zhang X, Gan R Z. Dynamic properties of human tympanic membrane—experimental measurement and modelling analysis[J]. *International Journal of Experimental and Computational Biomechanics*, 2010, 1(3): 252-270.
- [9] Aernouts J, Aerts J R M, Dirckx J J J. Mechanical properties of human tympanic membrane in the quasi-static regime from in situ point indentation measurements[J]. *Hearing research*, 2012, 290(1-2): 45-54.
- [10] Suzuki, H., Gao, K.Y.,1999. A review of research trends on passengers' aural discomfort caused by rail tunnel pressure change[J]. *Foreign Rolling Stock*5, 15-18.
- [11] Daphalapurkar, N.P., Dai, C.K., Gan, R.Z., Lu, H.,2009. Characterization of the linearly viscoelastic behavior of human tympanic membrane by nanoindentation[J]. *Journal of the Mechanical Behavior of Biomedical Materials*2, 82-92.
- [12] Liang J, Luo H, Yokell Z, et al. Characterization of the nonlinear elastic behavior of chinchilla tympanic membrane using micro-fringe projection[J]. *Hearing research*, 2016, 339: 1-11.
- [13] Aernouts, J., Soons, J.A.M., Dirckx, J.J.J.,2010. Quantification of tympanic membrane elasticity parameters from in situ point indentation measurements: validation and preliminary study[J]. *Hearing Research*263, 177-182.
- [14] Fay, J., Puria, S., Decraemer, W.F., Steele, C.,2005. Three approaches for estimating the elastic modulus of the tympanic membrane[J]. *Journal of Biomechanics*38, 1807-1815.
- [15] Kerr,A.G.,1980. The effects of blast on the ear. *The Journal of Laryngology & Otology*94, 107-110.
- [16] Šprem,N., Branica,S., Dawidowsky,K.,2001. Tympanoplasty after war blast lesions of the eardrum: retrospective study. *Croat Med J*42, 642-645.
- [17] Shupak,A., Doweck,I., Nachtigal,D., Spitzer,O., Gordon,C.R.,1993. Vestibular and audiometric consequences of blast injury to the ear. *Archives of Otolaryngology–Head & Neck Surgery*119, 1362-1367.
- [18] Raphael,Y.,1992. Evidence for supporting cell mitosis in response to acoustic trauma in the avian inner ear. *Journal of Neurocytology*21, 663-671.
- [19] Yamane,H., Nakai,Y., Takayama,M., Iguchi,H., Nakagawa,T.,1995. Appearance of free radicals in the guinea pig inner ear after noise-induced acoustic trauma. *European Archives of Oto-rhino-laryngology*252, 504-508.

- [20] Park,B., Jeon,J.Y., Choi,S., Park,J.,2015. Short-term noise annoyance assessment in passenger compartments of high-speed trains under sudden variation. *Applied Acoustics*97, 46-53.
- [21] Garth,R.J.N.,1994. Blast injury of the auditory system: a review of the mechanisms and pathology. *The Journal of Laryngology & Otology*108, 925-929.
- [22] Love,J.T., Caruso,V.G.,1978. Civilian air travel and the otolaryngologist. *The Laryngoscope*88, 1732-1742.
- [23] Harding,R.,1994. Aeromedical aspects of commercial air travel. *Journal of Travel Medicine*1, 211-215.
- [24] Kraus,R.N.,1959. IV Treatment of Sinus Barotrauma. *Annals of Otology, Rhinology & Laryngology*68, 80-89.
- [25] Mirza,S., Richardson,H.,2005. Otic barotrauma from air travel. *Journal of Laryngology & Otology*119, 366-370.

Numerical Study on Debonding Failure between FRP and Concrete

†Lihua Huang¹, Zhiquan Yang¹, and Yuefang Wang²

¹School of Civil Engineering, Dalian University of Technology, China.

²Department of Engineering Mechanics, Dalian University of Technology, China.

*Presenting author: lhhang@dlut.edu.cn

†Corresponding author: lhhang@dlut.edu.cn

Abstract

Fiber reinforced polymer (FRP) has been widely employed in retrofitting concrete structures. Debonding of FRP from concrete is a typical failure mode in this technique. Cohesive zone model(CZM) of fracture energy-based criteria is demonstrated to be a well-founded numerical approach to characterize the brittle behavior of interfacial debonding failure. A simple but robust finite-element (FE) model of CZM for simulating the debonding procedure induced by the intermediate concrete crack(IC) and discontinuous FRP edge is presented in this paper. The bilinear bond-slip relationship in the interface is applied in the numerical model. Ten FRP strengthened beams of IC debonding and edge debonding failures are simulated by the suggested FE model and verified by the experimental results. For FRP strengthened plain concrete beams of pre-cracks, the variations of applied loads with concrete crack mouth open displacements (CMOD) can be accurately revealed, and when the cohesive strength in the interfacial bilinear model is reduced to 40%, the edge debonding failures of reinforced concrete beams can be accurately illustrated. The variations of stress and strain in FRP and concrete as well as in the interface with the increase of CMODs and deflections are expediently and expressly indicated through the proposed FE model.

Keywords: FRP; Concrete beam; CZM; IC debonding, Edge debonding

Introduction

External bonding of fiber reinforce polymer (FRP) sheet or plate has been widely accepted as an effective and convenient technique for the strengthening of concrete structures. One of the key factors that influence the strengthening is the interfacial debonding failure of FRP from substrate [1]-[4]. The behavior of FRP-to-concrete interfaces has been deeply investigated, and well expressed by bond-slip models [5][6], which make the prediction of interfacial debonding possible. Among the debonding failure modes, one is caused by the opening up of concrete flexural cracks and the debonding initiates from those places. This kind of debonding failure is commonly referred to as the intermediate concrete crack(IC)-induced debonding[7]. Another main failure mode is named as the edge debonding because the interfacial debonding initiates from the edge of FRP sheet and then propagates towards the middle of the interface.

The sound analytical studies on interfacial debonding have been conducted by many researchers, where the self-programming finite element(FE) models are developed to predict the load-deflection behavior. Although the proposed models are demonstrated to be valid, most of them are lack of accessibility. The primitive versatile numerical models are commonly based on the strength of interfacial shear stress represented by the bridging elements such as link and spring[8]-[10], which is actually inconsistent with the interfacial brittleness of debonding process. Fracture energy approach is recognized to be more appropriate to capture the interfacial fracture behavior and account for the possible failure

modes [6]. Among the FE models of fracture energy-based criteria, cohesive zone modeling(CZM) approach is well known for its simplicity and accurate kinematics representation of the quasi-brittle fracture process zone, which arises prior to complete fracture in, e.g., concrete materials and macro-molecular based polymer materials[11][12]. The bond-slipping of the interface can be governed by a fundamental fracture energy that is also the energy required to break apart the interface surfaces in this model. The whole debonding process from crack initiation to growth and further slipping can be unified into one model and easily formulated and implemented by CZM. According to this idea, Wang(2006) has established the closed-form solution of CZM for IC induced debonding[13]. De Lorenzis(2009) and Cornetti(2015) put forward an analytical cohesive crack modeling approach to the edge debonding failure of FRP-plated beam[14] [15]. Chen(2009) conducted debonding analysis of adhesively bonded interface between two balanced adjacent flexural cracks by CZM[16]. As stressed in many studies, much further research is needed to develop a simple and accurate simulating model that has a rational theoretical basis and a practical engineering application.

In the present work, the most common debonding failures of FRP strengthened concrete beams are simulated by applying the proposed FE model of CZM based on the FE package ABAUSE/standard. Two rational models are developed to easily reveal the processes of IC-induced debonding and FRP edge debonding. Exponential degradation of concrete cracking and bilinear relationship between interfacial shear stresses and slips proposed by Lu(2005) are assigned as to the properties of the interfaces[5]. To demonstrate the validity of the numerical models, the corresponding experimental data from the literatures are taken as the verification of the numerical results. As the governing parameters in CZM, the values of the maximum interfacial shear stress and the fracture energy have been discussed. The simple FEM and suggested parameters are provided to be useful for predicting the propagation of debonding failures between FRP/concrete interfaces.

Material Modeling

Concrete, Steel and FRP Composite

The reduction of the strength and stiffness of concrete is represented by the plastic-damage model introduced in ABAQUS. The uniaxial compressive stress-strain curve for concrete is determined by the equation of Hognestad[17][18], in which the maximum strain is taken as $\epsilon_{cu} = 0.0038$. For the finite-element implementation, the values of the tensile strength, f_t , and elastic modulus, E_0 , if not given, are approximated based on the following ACI-318-05[19] equations:

$$E_0 = 4730\sqrt{f_c'} \quad (1)$$

$$f_t = 0.53\sqrt{f_c'} \quad (2)$$

Concrete is assumed to be elastic before cracking. The tensile degradation of concrete is expressed by the exponential equation of Reinhardt [18][20] in fracture mechanics as follow:

$$\frac{\sigma_t}{f_t} = \left[1 + \left(\frac{c_1 \omega}{\omega_0} \right)^3 \right] \exp\left(-\frac{c_2 \omega}{\omega_0}\right) - \frac{\omega}{\omega_0} (1 + c_1^3) \exp(-c_2) \quad (3)$$

where cohesive stress, σ_t , is corresponding to the open distance of crack, ω . Factors, $c_1=3$, $c_2=7$ for normal concrete. $\omega_0=0.16$ mm as $\sigma_t=0$.

Steel is represented by an elastic-plastic constitutive relationship with linear strain hardening, and FRP is assumed to be linear elastic constitutive relationship.

FRP/Concrete Interface

In the current study, the bilinear bond-slip model proposed by Lu(2005), as shown in figure 1, is adopted for its simplicity and easy incorporation into the FE analysis [5]. The behavior of the FRP/concrete interface is modified as a relationship between the local shear stress, τ , and the relative displacement, s , as follow.

$$\beta_\omega = \sqrt{\frac{2.25 - b_f / b_c}{1.25 + b_f / b_c}} \quad (4)$$

$$\tau_{\max} = 1.5 \beta_\omega f_t \quad (5)$$

$$s_0 = 0.0195 \beta_\omega f_t \quad (6)$$

$$G_f = 0.308 \beta_\omega^2 \sqrt{f_t} \quad (7)$$

in which b_f / b_c and f_t are the ratio of the width of FRP to concrete and the splitting tension strength of concrete, respectively.

The maximum nominal stress criterion is applied to determine the initiation of interfacial cracking, namely shear crack develops at the point τ_{\max} . Interfacial performance of bond-slip is governed by the cohesive strength τ_{\max} and the fracture energy G_f , which are essentially identical in this model.

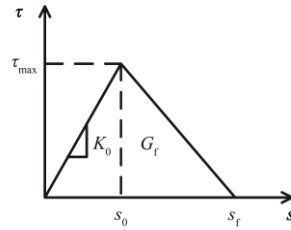


Figure 1. Bilinear bond-slip model

FE Modeling of IC Debonding

Profile of the Strengthened Beams with IC Debonding

IC debonding failure has been experimentally investigated by many researchers through testing FRP strengthened plain concrete beams of various seam-height ratios. Three-point bending beams of different heights, lengths, and seam height ratios from the literature[21] are simulated by using the FE model of CZM in this paper. FRP is of the same length, but one-third width of concrete beam. To prevent the conical shear failure around intermediate crack, unbonded segment is set up near the crack.

Parameters such as the splitting tension strength of concrete, the flexural strengths of beams with different heights, tensile strength of CFRP, modulus of elasticity and Poisson's ratios of concrete and CFRP are taken from the experimental data in the literature[21].

Based on the experimental data, the maximum shear stress and the fracture energy are determined according to the equations (4)-(7), namely $\tau_{\max}=5.94$ MPa, $G_f=0.7$ N/mm, and the maximum slip $s_f=2G_f / \tau_{\max} = 0.236$ mm as the interfacial debonding occurs.

Constitutive Properties of Concrete Cracking

Since IC debonding is induced by the propagation of intermediate concrete crack, the modeling of concrete cracking with CZM is essentially identified first. It is assumed that concrete is a linearly elastic material when the equivalent stress in plain concrete is less than $0.3f_c$ [22]. Fictitious cracking model is adopted to avoid the singularity of crack tip in FEA. According to the exponential concrete virtual crack model proposed by Reinhardt[17][20], equivalent crack opening displacement $\omega_0=0.16\text{mm}$ when cohesive force decreases to zero. The maximum nominal stress criterion is applied for modeling the initiation of concrete crack, in which fictitious crack develops when tensile stress reaches to the flexural strength of concrete. Equations of cohesive force with crack opening distance is expressed by

$$f(\delta) = f_t \left[1 - \frac{1 - e^{-\alpha \left(\frac{\delta - \delta_0}{\delta_f - \delta_0} \right)}}{1 - e^{-\alpha}} \right] \quad (8)$$

$$\delta_0 = 1 \times \frac{f_t}{E_c} \quad (9)$$

where f_t is the flexural strength of concrete; default value is one for zero constitutive thickness of cohesive element; δ_f represents the open distance of crack when cohesive force equals zero, namely $\delta_f = \omega_0$. Coefficient, α , is used to govern the curves of exponential degradation of concrete, which is determined by the experimental data in literature [21]. When $\alpha=10$, as listed in table 1, the peak loads of five beams determined by numerical solution, P_{Num} , are close to the experimental values, P_{Exp} . Based on the equations (8), (9) and the parameter α , the constitutive relationships between loads and crack mouth open displacements (CMOD) of the concrete beams are constructed as shown in figure 2, which are also the curves for governing IC propagation.

Table 1. Comparison of numerical and experimental peak loads

Beams	P_{Exp}/kN	P_{Num}/kN	Error /%
C202	9.4	10.07	7.13
C203	6.93	8.08	16.60
C204	5.65	6.05	7.08
C253	8.69	9.42	8.40
C303	9.92	10.02	1.01

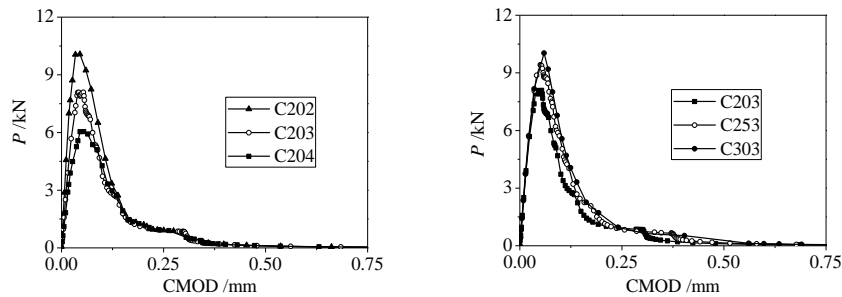


Figure 2. Curves of load-CMOD of the concrete beams

Numerical and Experimental Results Comparison

In the FE model, the concrete and the FRP are modeled by solid C3D8R and shell S4R in the package ABAQUS. The unbonded segment around intermediate crack is treated as frictionless contact. Cohesive elements COH3D8 are embedded in the middle of concrete beam and the FRP/concrete interface for identifying the IC propagation and FRP debonding. The five plain beams bonded with CFRP from literature[21] are simulated by the suggested FE model of CZM, and the corresponding load-CMOD curves are constructed, as shown in figure 3. It is shown that, close to the experimental results, there are apparently two peak points of loads, P_{1max} and P_{2max} . The applied load is linearly proportional to CMOD when it is less than P_{1max} . After the first peak point, the load decreases with the IC propagation. Then the load rises again until to the second peak point due to the activation of CFRP. Interfacial slip starts at this time. The numerical results obtained from the suggested FE model are well agreeable with the experimental results, as shown in table 2 and figure 3.

Table 2. Comparison of the peak loads obtained from the FEM and experiments

Beams	P_{1Num}/kN	P_{1Exp}/kN	Error/%	P_{2Num}/kN	P_{2Exp}/kN	Error/%
P202	11.30	10.97	3.00	11.98	11.95	0.25
P203	8.71	9.07	3.97	11.97	11.66	2.66
P204	6.73	7.72	12.82	11.97	12.33	2.92
P253	9.88	11.18	11.63	11.96	12.80	6.56
P303	10.34	12.93	20.03	12.02	13.55	11.29

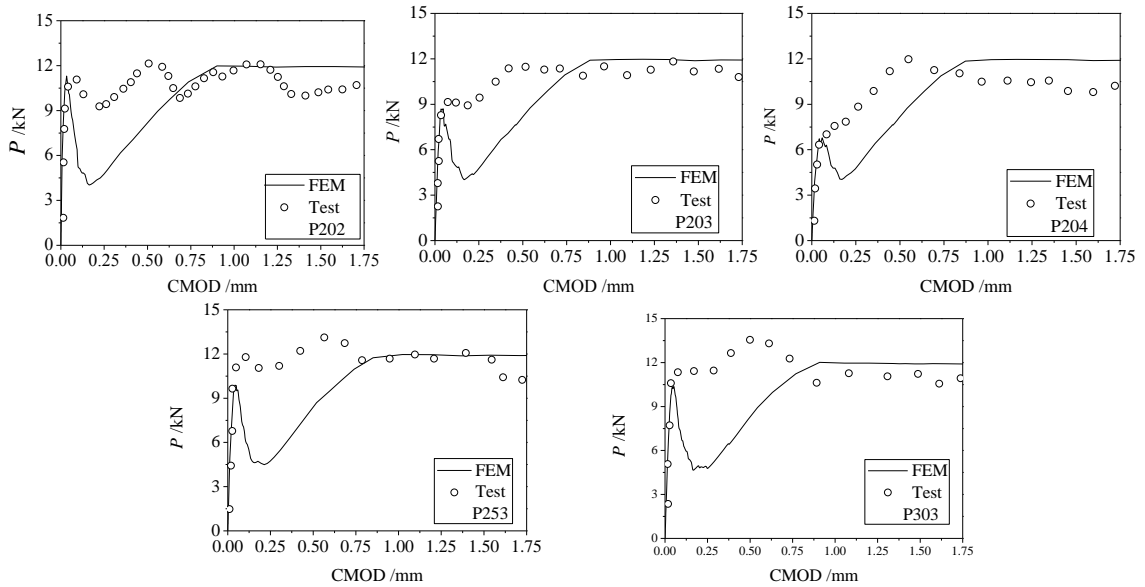


Figure 3. Comparison of load-CMOD curves obtained from the FEM and experiments

Identification of IC-induced Debonding

Stress distribution of CFRP of beam 203 under different CMOD is constructed in figure 4. It is shown that the stress at the middle of CFRP linearly increases with CMOD before the applied load reaches to P_{2max} . When the load equals to P_{1max} , where CMOD equals to 0.046 mm as shown in figure 4, stress in CFRP remains in a low level, namely 88MPa. While, when the applied load equals to P_{2max} , where CMOD equals to 1.183 mm, stress in CFRP reaches to

the highest value of 1430 MPa. Then the stresses remain in a high level with the continuous increase of CMOD, and gradually transfer from the middle to the edge of CFRP. It illustrates that the stress in CFRP is mainly caused by the IC propagation and interfacial slip. When CMOD reaches to 2.968 mm, most of the CFRP stays in a high stress level of 1430 MPa, as shown in figure 4.

Stresses in concrete indicate that most of the concrete stay in a low stress level because of concrete cracking and interfacial slipping. Stress concentrations are obvious at the tips of concrete crack and interfacial shear crack, and the maximum stress is close to the flexural strength of concrete. Cohesive stress distribution of concrete crack under specific CMOD is shown in figure 5. When the applied load reaches to P_{1max} , fictitious crack is of 55 mm length and the cohesive tensile stress near the tip of concrete crack is 0.870 MPa. As the load reaches to P_{2max} , the fictitious crack extends to 135 mm and the macro-crack to 105 mm at the same time. Because of the extension of concrete crack, the loading capacity of FRP-bonded concrete beam gradually decreases after the applied load exceeds P_{1max} , and most of the load is sustained by CFRP with little contribution of concrete after the applied load exceeds P_{2max} .

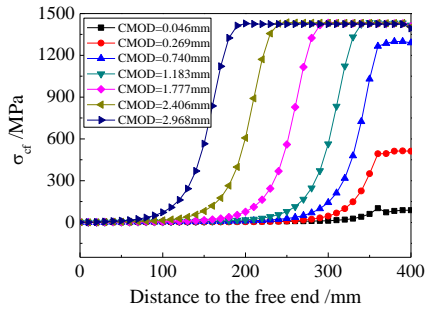


Figure 4. Stresses of CFRP

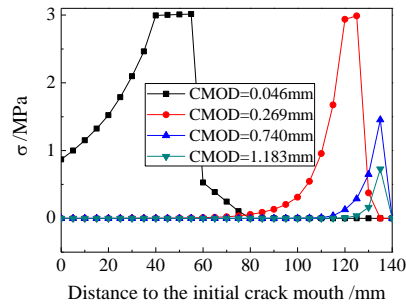


Figure 5. Adhesive stress of concrete crack

Along with the increase of CMOD, the interfacial bonding stress moves from the middle toward the edge, as shown in figure 6. When the applied load equals to P_{1max} , where CMOD equals 0.046mm, the maximum interfacial shear stress is 0.5MPa. Interface behaves in elastic with a low level of shear stress. Interfacial shear stress near the intermediate crack decreases to zero when the applied load increases to the second peak load P_{2max} , namely CMOD getting to 1.183 mm. Macro shear crack of 20mm length occurs at this time. The interfacial shear stress gradually moves from the middle to the edge due to the enlargement of CMOD.

Variation of interfacial slip with CMOD is plotted in figure 7. Analogous to the interfacial shear stress, interfacial slip extends from the middle to the edge of the interface following the enlargement of CMOD. The inflection points of the curves in figure 7 are corresponding to the maximum shear stresses presented in figure 6.

FE Modeling of Edge Debonding

Edge Debonding Failure in FRP Strengthened RC Beams

FRP strengthened RC beams of four-point bending are taken from the experiments in literature[23][24]. Five beams with the edge debonding are simulated by the proposed FE model of bilinear bond-slip interfacial property. Based on the experimental data, the maximum shear stress and the fracture energy are determined according to the equations (4)-(7), and the maximum slip $s_f = 2F_f / \tau_{max}$ as the interfacial debonding occurs.

Concrete, steel bar and CFRP are modeled with solid C3D8R, truss T3D2 and shell S4R in the package ABAQUS, respectively. Interfacial bond-slip is represented by COH3D8. Slipping between concrete and steel bars is neglected. The numerical load-displacement curves determined by the supposed FE model are compared with the experimental results, as shown in table 3. It is shown that the peak loads determined by the numerical model are well agreeable with the experimental results.

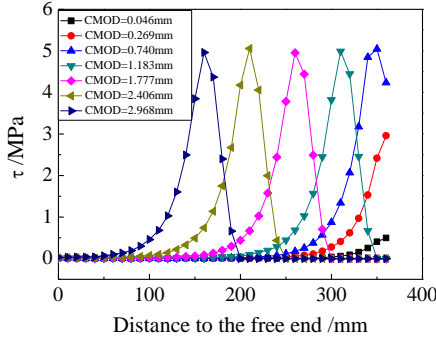


Figure 6. Variation of interfacial stress

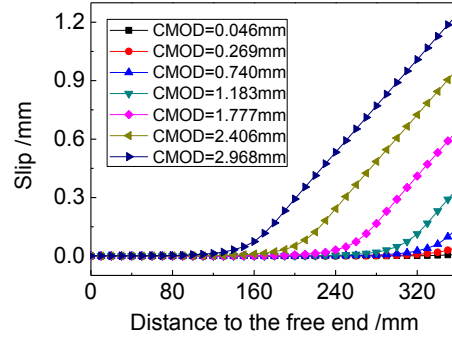


Figure 7. Variation of interfacial slip

Table 3. Numerical and experimental peak loads

Literature	Beams	P_{Num}/kN	P_{Exp}/kN	Error/%
Obaidat[23]	RF1	166.2	166	0.1
	RF2	146.5	142	3.2
	RF3	129.9	128	1.5
Quantrill[24]	B2	35.4	34	4.1
	B3	26.8	24.6	8.9

The numerical curves of load-displacement are compared with the experimental data, and plotted in Figure 8. It is shown that beams analyzed by the FE model exhibit greater ductility than the specimen. The deflections of the beams are bigger than the real beams when the applied loads reach or exceed the peak loads. The reasons that lead the exaggeration of ductility in FE analysis lie in that concrete is assumed to be an elastic-plastic property excluding the effect of concrete crack.

Parameters in Bond-slip Relationship

If the fracture energy G_f determined by equation (7) remains constant, and the maximum cohesive shear stress is assumed to be $0.3\tau_{max}$, $0.4\tau_{max}$, $0.5\tau_{max}$, $0.6\tau_{max}$ or $1.0\tau_{max}$, where τ_{max} is determined by equations (5), the curves of load-displacement from the FE model are constructed and compared with the experimental results, as shown in figure 9. Beam RF3 is not discussed since there is no obvious strengthening effect for short strengthening length. It can be seen that when the maximum shear stress is greater than $0.4\tau_{max}$, the values of the maximum shear stress have little effect on the peak loads, but greatly increase deflections. If the maximum shear stress is smaller than $0.3\tau_{max}$, the peak loads are apparently lower than the real values. The curves of load-displacement obtained from the FE model are the closest to the experimental results when the maximum interfacial shear stress equals to $0.4\tau_{max}$. There are apparent downward jumps to the loading strength of non-strengthened beams when

debonding occurs in the curves. The loading capacities of the strengthened beams are obviously controlled by the debonding of FRP, and the region of damage in the interface narrows down with the increase of the maximum shear stress. Therefore, the rational maximum shear stress in CZM is suggested to be $0.4\tau_{\max}$ in the FE analysis of FRP debonding from RC beams.

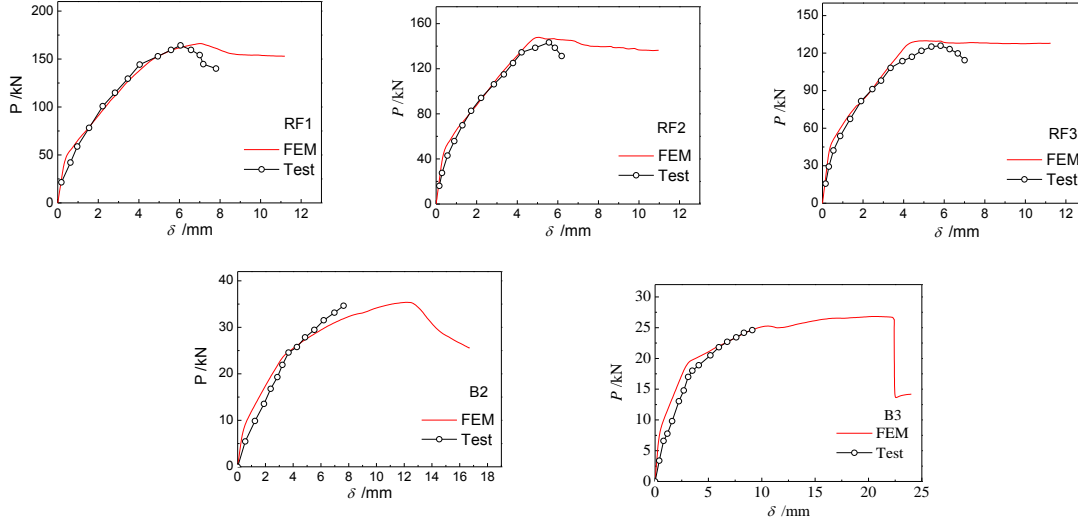


Figure 8. Load-displacement curves from FEM and experiments

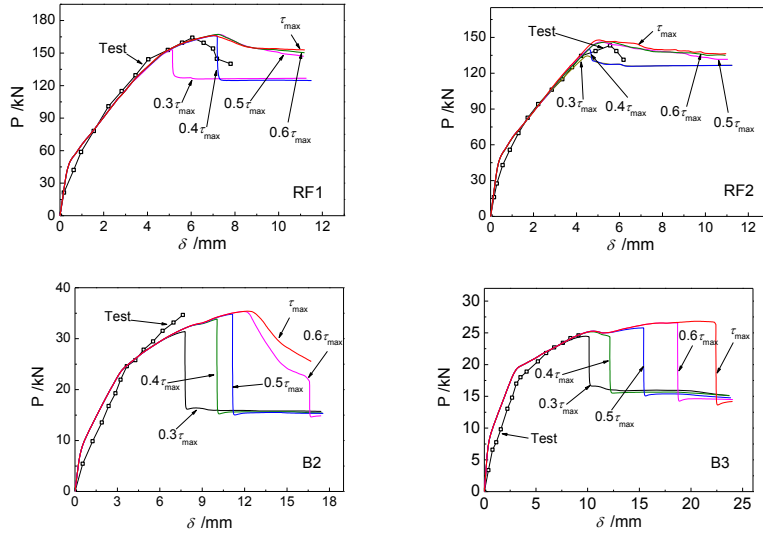


Figure 9. Load-displacement curves of different cohesive strength

The effect of fracture energy in CZM on numerical results is studied below. If the maximum shear stress is identified as $0.4\tau_{\max}$, and the fracture energy is taken as $0.5G_f$, $0.75G_f$, G_f , $1.25G_f$ or $1.5G_f$, where G_f is determined based on the equation (7), the corresponding load-displacement curves are constructed in figure 10. It can be seen that the ductility is exaggerated with the increase of the fracture energy. Large values of fracture energy result the failure of concrete crushing rather than FRP debonding as shown in the beams RF1 and RF2 in the case of fracture energy greater than $1.5G_f$. Small values of fracture energy result lower prediction of loading strength as shown in the curves of $0.5G_f$. The loading capacities of beams are obviously improved with the increase of fracture energy as it is less than G_f , and is

little upgraded when the fracture energy is greater than G_f . The curves of load-displacement are the closest to the experimental results when fracture energy equals to G_f .

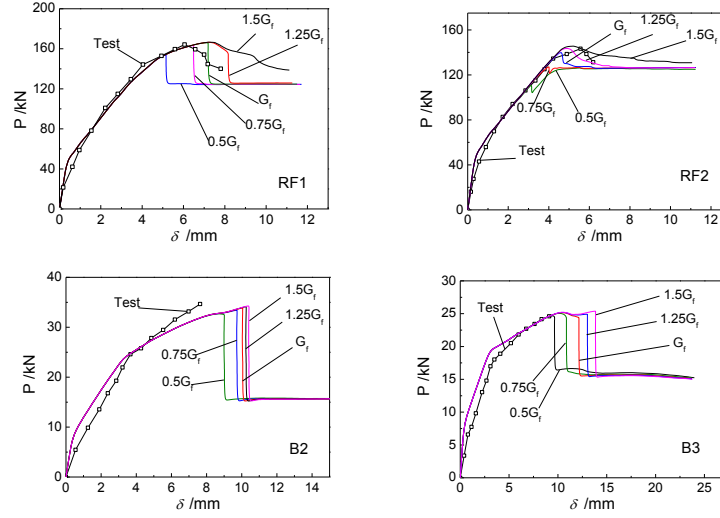


Figure 10. Influence of the fracture energy in CZM on FE analysis

Identification of Edge Debonding

Based on the forgoing discussion, the fracture energy determined by the equation (7) and the maximum shear stress modified to $0.4\tau_{\max}$ are applied in the FEA of beam RF1. Stress distribution in FRP shows that the tensile stress at the edge of FRP reaches to the maximum, 35MPa, when the deflection at the middle of the beam reaches to 6.2mm. Later, the stress decreases quickly with the increase of deflection until to zero when deflection d equals 7.2mm, as shown in figure 11. Stress at the middle of FRP nonlinearly increases to 667MPa as $d=7.2$ mm, and then drops rapidly to a very low level due to the edge debonding, as illustrated in figure 12. The curves of stress distributions along FRP with respect to the enlargement of deflections clearly indicate that the stresses gradually increase with deflection until to the peak load, where $d=7.1$ mm, and then rapidly drops to very low level as $d=9$ mm shown in figure 13.

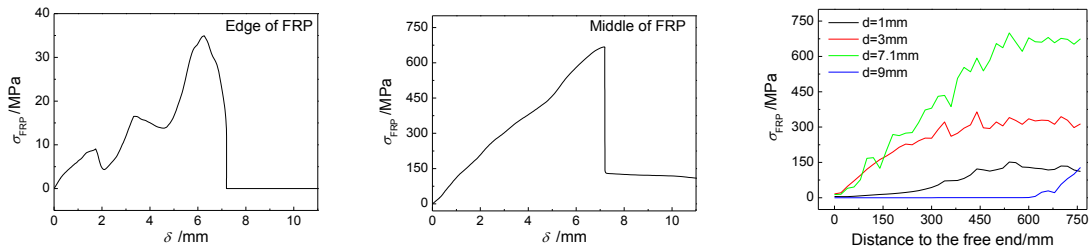


Figure 11. Stress at the edge Figure 12. Stress in the middle Figure 13. Stress along FRP

The behavior of interface between FRP and concrete is illustrated in figure 14-17. The stress at the edge of interface increases with deflection until to the maximum 3.5 MPa, where $d=3.4$ mm, and then gradually decreases to zero as deflection equals 7.2 mm in figure 14. At the middle of the interface in figure 15, the stress stays in a very low level about 1MPa before

edge debonding occurs. Then it rapidly increases to the maximum corresponding to the peak load at the moment of edge debonding. Afterward, stress stays in a very high level as debonding propagates from the edge to the middle of the beam. The distributions of stresses and slips along the interface under various deflections are shown in figure 16 and figure 17. Before the peak load, the interfacial stresses increase with the deflections, while slips stay in low levels. After the peak load, most of the interfacial stresses drop to zero and the slips greatly upgrade in the relevant region of the interface.

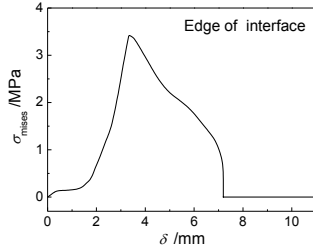


Figure 14. Stress at the edge of interface

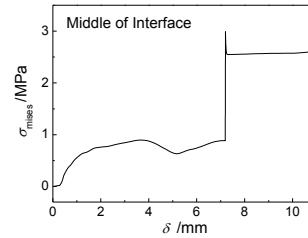


Figure 15. Stress at the middle of interface

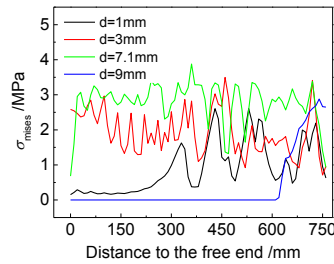


Figure 16. Interfacial stress distributions

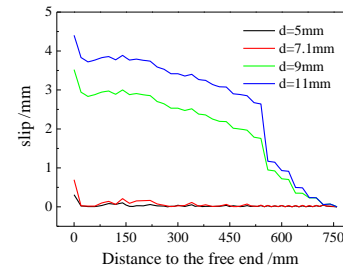


Figure 17. Interfacial slip distributions

The performance of concrete is revealed in figure 18 and figure 19. At the moment of the edge debonding, there is a downward jump in the curves of concrete stresses at the middle of the beam with respect to deflection. Afterwards, stresses continue to upgrade with the increase of deflection. When most of the FRP has debonded at the moment $d=11.5\text{mm}$, the stress and strain contours indicate that the biggest equivalent plastic strain occurs at the bottom of concrete, which is agreeable with the peeling of concrete surface happened in many experiments.

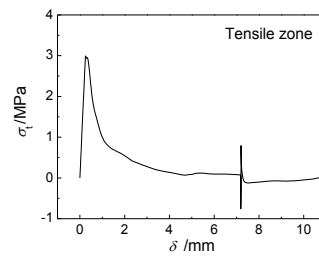
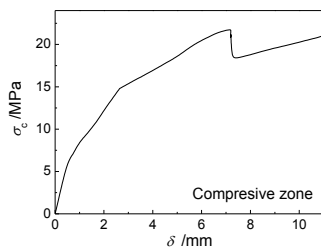


Figure 18. Curves of concrete stresses at the middle of beam to deflections

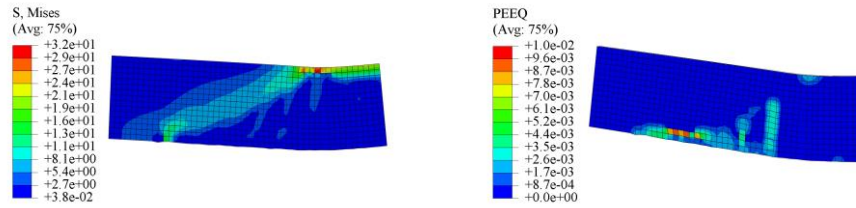


Figure 19. Stress and strain contour of concrete

Conclusions

A simple but robust FE model of fracture criteria-based CZM is set up for simulating the debonding failures of FRP strengthened concrete beams in this paper. Two types of debonding processes are analyzed: IC debonding and edge debonding. In the FEA of IC debonding, cohesive elements of concrete fracture properties are embedded in the middle of beam as to indicate the effects of flexural cracking on interfacial debonding. Cohesive elements in the interface between FRP and concrete are assigned bilinear bond-slip properties. The interfacial debonding failure revealed by the FE model is verified to be well consistent with the experimental phenomenon. When the suggested FE model is applied to simulate the edge debonding of FRP strengthened RC beam, parameters in CZM must be modified because concrete is assumed to be of elastic-plastic properties, which leads the exaggeration of the ductility when Lu's bond-slip relation is employed in the CZM. When the fracture energy and the maximum interfacial bonding stress are assigned G_f and $0.4\tau_{\max}$, the edge debonding process of FRP strengthened RC beam is reasonably predicted and well captured.

Acknowledgements

The research was funded by the Liaoning Science and Technology Project 2014020008.

References

- [1] Teng J. G., Chen J. F. (2008) Mechanics of debonding in FRP-plated RC beams, *Structures and Granular Solids: From Scientific Principles to Engineering Application*, 313.
- [2] Saxena P., Toutanji H., Noumowe A. (2008) Failure analysis of FRP-strengthened RC beams, *Journal of Composites for Construction*, **12**(1), 2–14.
- [3] Buyukozturk O., Gunes O., Karaca E. (2004) Progress on understanding debonding problems in reinforced concrete and steel members strengthened using FRP composites, *Construction and Building Materials*, **18**(1), 9–19.
- [4] Yao J., Teng J. G., Chen J. F. (2005) Experimental study on FRP-to-concrete bonded joints, *Composites Part B: Engineering*, **36**(2), 99–113.
- [5] Lu X. Z., Teng J. G., Ye L. P. (2005) Bond-slip models for FRP sheets/plates bonded to concrete, *Engineering Structures*, **27**(6), 920–937.
- [6] Toutanji H., Han M., Ghorbe El. (2012) Interfacial bond strength characteristics of FRP and RC substrate, *Journal of Composites for Construction*, **16**(1), 35–46.
- [7] Teng J. G., Smith S. T., Yao J. (2003) Intermediate crack-induced debonding in RC beams and slabs, *Construction and Building Materials*, **17**(6), 447–462.
- [8] Baký H. A., Ebead U. A., Neale K. W. (2007) Flexural and interfacial behavior of FRP-strengthened reinforced concrete beams, *Journal of Composites for Construction*, **11**(6), 629–639.
- [9] Lu X. Z., Ye L. P., Teng J. G. (2005) Meso-scale finite element model for FRP sheets/plates bonded to concrete, *Engineering Structures*, **27**(4), 564–575.
- [10] Lu X. Z., Teng J. G., Ye L. P., Jiang J. J. (2007) Intermediate crack debonding in FRP-strengthened RC beams: FE analysis and strength model, *Journal of Composites for Construction*, **11**(2), 161–74.
- [11] Xu R., Liu C. (2011) CZM-based debonding simulation of cracked beams strengthened by FRP sheets, *Journal of Engineering Mechanics*, **138**(2), 210–220.
- [12] Yang Q. D., Thouless M. D., Ward S. M. (1999) Numerical simulations of adhesively-bonded beams failing

- with extensive plastic deformation, *Journal of the Mechanics and Physics of Solids*, **47(6)**, 1337–1353.
- [13] Wang J. (2006) Cohesive zone model of intermediate crack-induced debonding of FRP-plated concrete beam, *International Journal of Solids and Structures*, **43(21)**, 6630–6648.
 - [14] De Lorenzis L., Zavarise G. (2009) Cohesive zone modeling of interfacial stresses in plated beams, *International Journal of Solids and Structures*, **46(24)**, 4181–4191.
 - [15] Cornetti P., Corrado M., De Lorenzis L. (2015) An analytical cohesive crack modeling approach to the edge debonding failure of FRP-plated beams, *International Journal of Solids and Structures*, **53**, 92–106.
 - [16] Chen F., Qiao P. (2009) Debonding analysis of FRP-concrete interface between two balanced adjacent flexural cracks in plated beams, *International Journal of Solids and Structures*, **46(13)**, 2618–2628.
 - [17] Reinhardt H. W., Cornelissen H. A. W., Hordijk D. A. (1986) Tensile tests and failure analysis of concrete, *Journal of Structural Engineering*, **112(11)**, 2462–2477.
 - [18] Hillerborg A., Modér M., Petersson P. E. (1976) Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements, *Cement and Concrete Research*, **6(6)**, 773–781.
 - [19] American Concrete Institute. (2004) *Building Code Requirements for Structural Concrete (ACI 318-05) and Commentary (ACI 318R-05)*, American Concrete Inst.
 - [20] Reinhardt H. W., Cornelissen H. A. W., Hordijk D. A. (1986) Tensile tests and failure analysis of concrete, *Journal of Structural Engineering*, **112(11)**, 2462–2477.
 - [21] Yi F. M. (2010) The fracture properties of CFRP strengthened Cracked concrete beam, *Doctoral Dissertation*, Dalian University of Technology.
 - [22] Buyukozturk O. (1997) Nonlinear analysis of reinforced concrete structures, *Computers and Structures*, **7(1)**, 149–156.
 - [23] Obaidat Y. T., Heyden S., Dahlblom O. (2011) Retrofitting of reinforced concrete beams using composite laminates, *Construction and Building Materials*, **25(2)**, 591–597.
 - [24] Quantrill R. J., Hollaway L. C., Thorne A. M. (1996) Experimental and analytical investigation of FRP strengthened beam response: Part I, *Magazine of Concrete Research*, **48(177)**, 331–342.

Simulation of liquid cargo – vehicle interaction under lateral and longitudinal accelerations

†José A. Romero¹, Frank Otremba², and Alejandro A. Lozano Guzmán³

¹Faculty of Engineering, Queretaro Autonomous University, Mexico.

²Federal Institute of Materials Research and Testing (BAM), Berlin.

³Applied Science and Advanced Technology (CICATA), Queretaro, Mexico.

†Presenting and corresponding author: jaromero@uaq.mx

Abstract

Amongst the vehicle parameters influencing road safety, the carried cargo plays a critical role in the case of a liquid cargo, posing rollover risk and affecting the available friction forces for braking. While the lateral sloshing of the cargo within the vehicle's compartments, can be excited when the vehicle negotiates a turn, the longitudinal motion of the cargo derives from changes of speed. The combination of both types of perturbations occurs when the vehicle brakes while negotiating a turn. In this paper, a two-pendulum formulation is used to simulate the lateral and longitudinal behavior of a vehicle when negotiating a braking in a turn maneuver. The suspension forces are thus calculated as the linear superposition of both models. Results suggest that the vehicle roll stability is affected by the cargo sloshing, with increments on the order of 100% in the lateral load transfer, for a 50% filled tank. On the other hand, the dispersion of the travelling speed also affects the lateral stability of such type of vehicles, as a function of the dispersion of the vehicle's travelling speed.

Keywords: Sloshing cargo, road tankers, pendulum analogy, braking in a turn, lateral load transfer ratio

Introduction

Causality for a road crash derives from an unfortunate combination of events, that potentially involves the vehicle, the driver, the infrastructure and the environment. For a road crash, different levels of contributing factors and situations are identified, including a critical event, a critical reason and a critical source [1]. While the human factors are recognized as the main contributing factors for road crashes, the factors associated to the vehicle represent a major road crashes contribution. Within the vehicle, however, different influencing elements can be identified, including the failure of the mechanical parts of the vehicle, and the condition of the cargo. The cargo contributes to road crashes in different ways, as its configuration and nature can affect the performance of the vehicle, in different ways. While the cargo configuration refers to the dimensions of the cargo or its container; its nature refers to its physical state (liquid, solid) as well its level of hazardousness. While the height of the center of gravity position of a solid cargo can affect the lateral stability of a vehicle, the shifting of such center of gravity, which is an intrinsic characteristic of a liquid cargo, can pose a major risk to vehicle's lateral stability and longitudinal behavior. The mobility of the cargo within the carrying vehicle produces two main effects. On the one hand, the shifting of the center of gravity of the cargo, whether laterally or longitudinally, affects the balance of forces on the different ends of the vehicle, further affecting the rollover trend and the available braking force. The other potential effect of a liquid, sloshing, cargo, derives from the vibration frequency the liquid cargo. That is, the mobility of the cargo generates a coupled dynamic system with the mass-spring system, represented by the sprung and un-sprung masses of the vehicle.

The behavior of a partially filled road tanker while performing steering and braking maneuvers has been studied under the principles of several mathematical approaches, including computer fluid dynamics formulations (CFD), and some mechanical analogous formulations. Such models have been dynamic or quasi-static. Kolei et al. [2], consider the linear slosh theory for developing a hybrid multimodal and boundary-element model, to simulate the natural sloshing modes of the fluid within the tank. It is reported that the natural modes of vibration of the liquid inside the tank, can be excited as a function of the perturbation input, whether it is roll- or pitch- related. The authors assess the accuracy and computer effort to simulate the free response of the liquid inside the tank, when subjected simultaneously to longitudinal and lateral accelerations. However, the simulated tank is not suspended, that is, the contributing factors associated to the vibration of the tanker chassis, are not taken into account. Dasgupta [3], reports that the superposition of longitudinal and lateral accelerations strongly affects the longitudinal load shift, but that such superposition reduces the lateral load shift because of the cornering of the fluid within the tank. In the case of straight braking, the use of mechanical analogous models has been reported in [4] of partially filled road tankers, with high levels of validation.

In this paper, several formulations are set together to simulate the suspension forces of a 4-wheel straight road tanker carrying a liquid cargo at a partial fill level. The different models include a simplified analogous model to simulate the sprung tank containing a liquid cargo at a partial fill level, for both the roll plane and the pitch plane.

Model description

A multi-body formulation is proposed for the simulation of the response of a road tanker to perturbations derived from simultaneously changing both, direction and speed. Two similar un-coupled pendulum models are used to simulate the response of the vehicle in the roll and pitch planes. The model considers only the vibration of the sprung mass of the vehicle. That is, the effect of the un-sprung dynamics is not considered, and the flexibility provided by the tires stiffness, is assumed to be included in the equivalent torsional stiffness of the sprung mass [5].

Part (a) of Figure 1 illustrates a schematic representation of the two-axle straight road tanker, and parts (b) and (c) of this figure describe the analogous pendulum models representing the roll and pitch – plane vibration modes of the road tanker, respectively. As it can be seen in this figure, the individual suspension stiffness at the end of each vibration plane, was substituted by an equivalent torsional stiffness, producing a two-degree of freedom system, involving a simple pendulum and an inverted pendulum. A cylindrical shape for the tank is assumed in this paper. According to this model, a simple pendulum represents the sloshing cargo, while the chassis vibration is characterized in terms of an inverted pendulum.

In these models, the length of the simple pendulum is calculated in terms of a validated methodology to determine the natural sloshing frequency of the liquid inside the tank. The suspension springs and dampers of the 4-wheel road tanker, are assumed as having a linear behavior.

The equations of motion for the resultant mechanical system described in parts (b) and (c) of Figure 1, are derived from an Isaac Newton's approach.

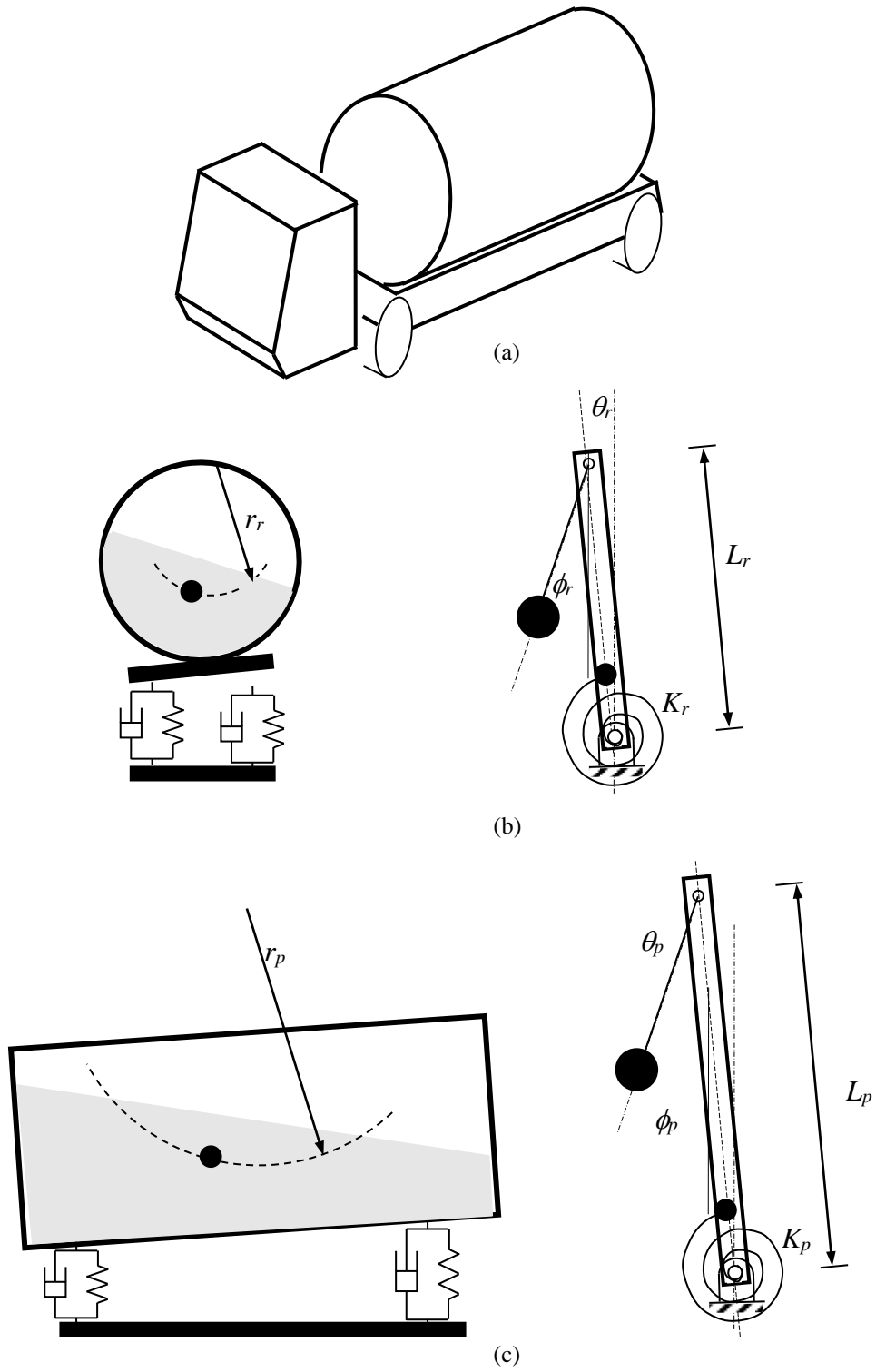
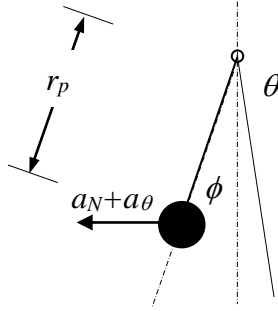


Figure 1. (a) Schematic representation of the vehicle; (b) roll vibration model; (c) pitch vibration model.

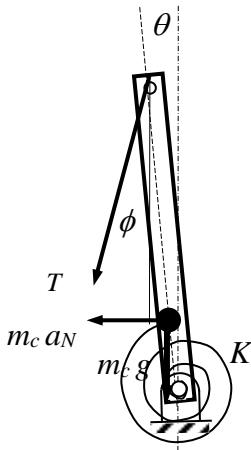
For the simple pendulum, its mass is subjected to lateral accelerations, as a result of the string tangential acceleration ($r_p \ddot{\phi}$), the inverted pendulum centripetal acceleration $L \left(\dot{\theta} \right)^2 (\phi + \theta)$;

and the tangential acceleration of the inverted pendulum $\ddot{\theta} L \cos(\theta)$. In this regard, the oscillation angles are considered small enough, to consider that $\sin(\phi + \theta) = \phi + \theta$; and $L \cos(\theta) = L$. The development of the equations of motion for the simple pendulum, is presented in Figure 2, and in Figure 3 the derivation of the equation of motion of the inverted pendulum, is presented. The tension of the simple pendulum that represents the sloshing cargo, is exerted on the inverted pendulum modeling the vehicle chassis. It should be noted that the center of rotation for the pendulum, is obtained from the summation of the center of gravity position of the sloshing cargo, and the length of the simple pendulum representing such cargo, r_p , whose length derivation is presented below in this paper.



$$\begin{aligned}
 a_\theta &= r_p \ddot{\phi} - L \left(\dot{\theta} \right)^2 (\phi + \theta) + \ddot{\theta} L \cos(\theta) \quad (1) \\
 -mg\phi &= ma_\theta = m \left[r_p \ddot{\phi} - L \left(\dot{\theta} \right)^2 (\phi + \theta) + \ddot{\theta} L \cos(\theta) \right] \\
 -mg\phi &= ma_\theta = m \left[r_p \ddot{\phi} - L \left(\dot{\theta} \right)^2 (\phi + \theta) + \ddot{\theta} L \cos(\theta) \right] \\
 \ddot{\phi} &= \phi \left[-\frac{g}{r_p} + \frac{L}{r_p} \left(\dot{\theta} \right)^2 \right] + \frac{L}{r_p} \left(\dot{\theta} \right)^2 \theta - \frac{\ddot{\theta}}{r_p} L \cos(\theta)
 \end{aligned}$$

Figure 2. Equations of motion derivation for the simple pendulum representing the sloshing cargo (similar models for pitch and roll vibration)



$$\ddot{\theta} = \frac{1}{I_p} \left\{ -K\theta - C\dot{\theta} + T(\cos(\phi)L\theta + \sin(\phi)L\cos(\theta)) \right\} \quad (2)$$

where:

$$T = mg \cos(\phi) + ma_r$$

where:

$$a_r = \ddot{\theta} L \cos[(\pi/2) - \phi - \theta] + L \left(\dot{\theta} \right)^2 \cos(\phi + \theta) - L \left(\dot{\theta} \right)^2$$

Figure 3. Equations of motion derivation for the inverted pendulum representing the vehicle chassis cargo (similar models for pitch and roll vibration)

The coupled equations of motion (1) and (2) are solved through the Transition Matrix Methodology (TMM), which represents a computationally efficient method for obtaining the dynamic response of linear mechanical systems in the time dominium. In this regard, some variables in Eq. (1) are also the outputs from the model, so that the time response will be obtained on the basis of previous outputs from the model.

According to the TMM scheme, the coupled equations of motion (1) and (2) are expressed as a first order system, on the basis of State Vector variables, as follows:

$$\left\{ \dot{y}(t) \right\} = [A] \{y(t)\} + [B]_1 \{Y_1(t)\} + [B]_2 \{Y_2(t)\} \quad (3)$$

where: $\left\{ \dot{y}(t) \right\} = \left\{ \dot{\phi} \quad \ddot{\phi} \quad \dot{\theta} \quad \ddot{\theta} \right\}^T$; $\{Y_1(t)\} = \left\{ 0 \quad \ddot{\theta} \quad 0 \quad \ddot{\phi} \right\}^T$; $\{Y_2(t)\} = \{0 \quad a_N \quad 0 \quad 0\}^T$

$$[A] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{9.81}{r_p} & -\frac{C_A}{m} & -\frac{L_B \left(\dot{\theta} \right)^2}{r_p} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{T\phi L_B \cos\theta}{I_p} & 0 & \frac{T\phi L_B \cos\theta - K_\theta}{I_p} & -\frac{C_\theta}{I_p} \end{bmatrix}$$

$$[B]_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -\frac{L_B}{r_p} \cos\theta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[B]_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{r_p} 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{m w L_B}{I_p} & 0 & \frac{1}{I_p} \end{bmatrix}$$

The solution of Equation (3), proceeds according to the Transition Matrix Approach [6], where the discrete time response is expressed as follows:

$$\{y(t + \Delta t)\} = [\Phi] \{y(t)\} + [\Gamma]_1 \{Y_1(t)\} + [\Gamma]_2 \{Y_2(t)\} \quad (4)$$

where: $[\Phi] = e^{[A]\Delta t} = [1] + \Delta t [A] + \frac{\Delta t^2}{2!} [A]^2 + \frac{\Delta t^3}{3!} [A]^3 + \dots \dots \dots \frac{\Delta t^n}{n!} [A]^n$

and: $\Gamma_1 = [A]^{-1}(\Phi - I)[B_1]; \quad \Gamma_2 = [A]^{-1}(\Phi - I)[B_2]$

On the other hand, when incorporating into a curve that has a constant radio, the vehicle is subjected to transient accelerations before it is subject to the acceleration derived from the constant radius turn. The transition from an infinite, straight road segment, into a constant radius turn, represents different transition radiuses, as it is illustrated in Figure 4, where the representation is simplified into a bicycle model. The details of the model are presented in [7].

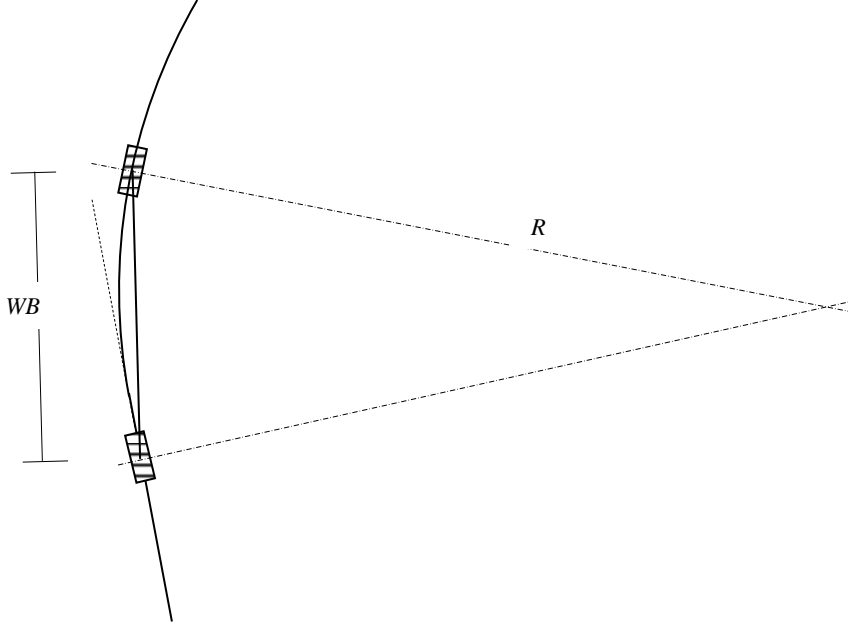


Figure 4. Transition from a straight into a constant radius curve.

The pendulum length r_p in the above equations, is calculated on the basis of a validated methodology reported in [8], based upon the sloshing motion in a rectangular tank of length L as follows:

$$2Lf = \left(\frac{g}{\kappa} \tanh \kappa H \right)^{1/2} \quad (5)$$

where H is the depth of the rectangular tank, L is the free length of the fluid surface, and κ is the wave number $= 2\pi / \lambda$, where $\lambda = 2L$. This formula has been successfully used for non-rectangular tank situations, if an equivalent depth H_e is used, as a function of the cross sectional area of the tank, as follows [9]:

$$H_e = \frac{Area}{L_f} \quad (6)$$

where L_f is the length of the free surface.

Results

The described model is used to simulate the behavior of a half-filled road tanker under two circumstances: when making a brake in a turn maneuver, and when travelling along a winding road. For the braking in a turn maneuver, two different turn radiuses are considered, for an initial vehicle speed (30 km/h). Table 1 lists the values for the different vehicle properties, including the lengths of the pendulums for the roll and pitch modes of vibration. For half-

filled circular tank, the center of gravity, is located at 0.5 m from the free surface. For the calculation of the natural frequency, a tank diameter is considered of 2.4 m, while the length is determined as 6.6 m for a 30 cubic meters capacity straight road tanker. The corresponding natural sloshing frequencies for the roll and pitch sloshing, according to the formulation described above, are 0.52 Hz and 0.24 Hz, respectively, resulting in pendulum lengths of 0.9 m, and 4.06 m, respectively,

Table 1. Properties of the vehicle

Equivalent torsional stiffness (pitch)	118300000 N-m/rad
Equivalent torsional stiffness (roll)	7056000 Nm/rad
Simple pendulum length (pitch)	4.06 m
Simple pendulum length (roll)	0.9 m
Chassis mass moment of inertia (pitch)	652 km-m ²
Chassis mass moment of inertia (roll)	1305 km-m ²
Chassis mass	4000 kg
Liquid mass	15000 kg
Working fluid	Water
Vehicle wheelbase	6.5 m

Performance measure

The load transfer that affects the vehicle when performing a braking in a turn maneuver, is characterized in terms of the lateral load transfer ratio (*LLTR*), as follows:

$$LLTR = \text{abs}(F_L - F_R) / (F_L + F_R) \quad (7)$$

where F_L is the suspension force on the left side, and F_R is the right suspension force. According to this expression, *LLTR* can have a value from 0 to 1, where 1 means that one side of the vehicle is losing contact with the road. While the standard *LLTR* is calculated based on the tire forces, in this paper such forces are the suspension forces. Furthermore, in the case of the three-dimensional situation considered in this paper, two *LLTR* shall be calculated, on the front and rear sides.

Braking in a turn maneuver

Figures 5 and 6 illustrate the simulation results when the vehicle performs an emergency braking in a turn maneuver at 3.5 m/s², for an initial speed of 30 km/h. Parts (a) and (c) of these figures describe the time history of the four suspension forces involved: front and rear; left and right. While part (a) corresponds to a braking maneuver on a 500-meter radio curve, part (c) describes the results for a tighter turning maneuver, corresponding to a 287 m curve. Parts (b) and (d) describe the corresponding values of the *LLTR*, for the front and rear suspension positions. According to these results, reducing from 500-meter to 287 meters the radius of the turn (42% reduction), causes an increase in the maximum *LLTR* from 0.4 to 0.7 (75%). Consequently, there is a non-linear effect of the turning radius on the rollover trend of the vehicle. It can also be noted that, regardless of the turning radius involved, the maximum value for *LLTR* occurs after about 5 cycles of chassis – liquid cargo interaction. Additionally, it can be observed that the greater rollover trend occurs in the case of the front axle.

Figure 7 describes a summary of the effect of both cargo condition and curve radius, on the maximum values obtained for the *LLTR*.

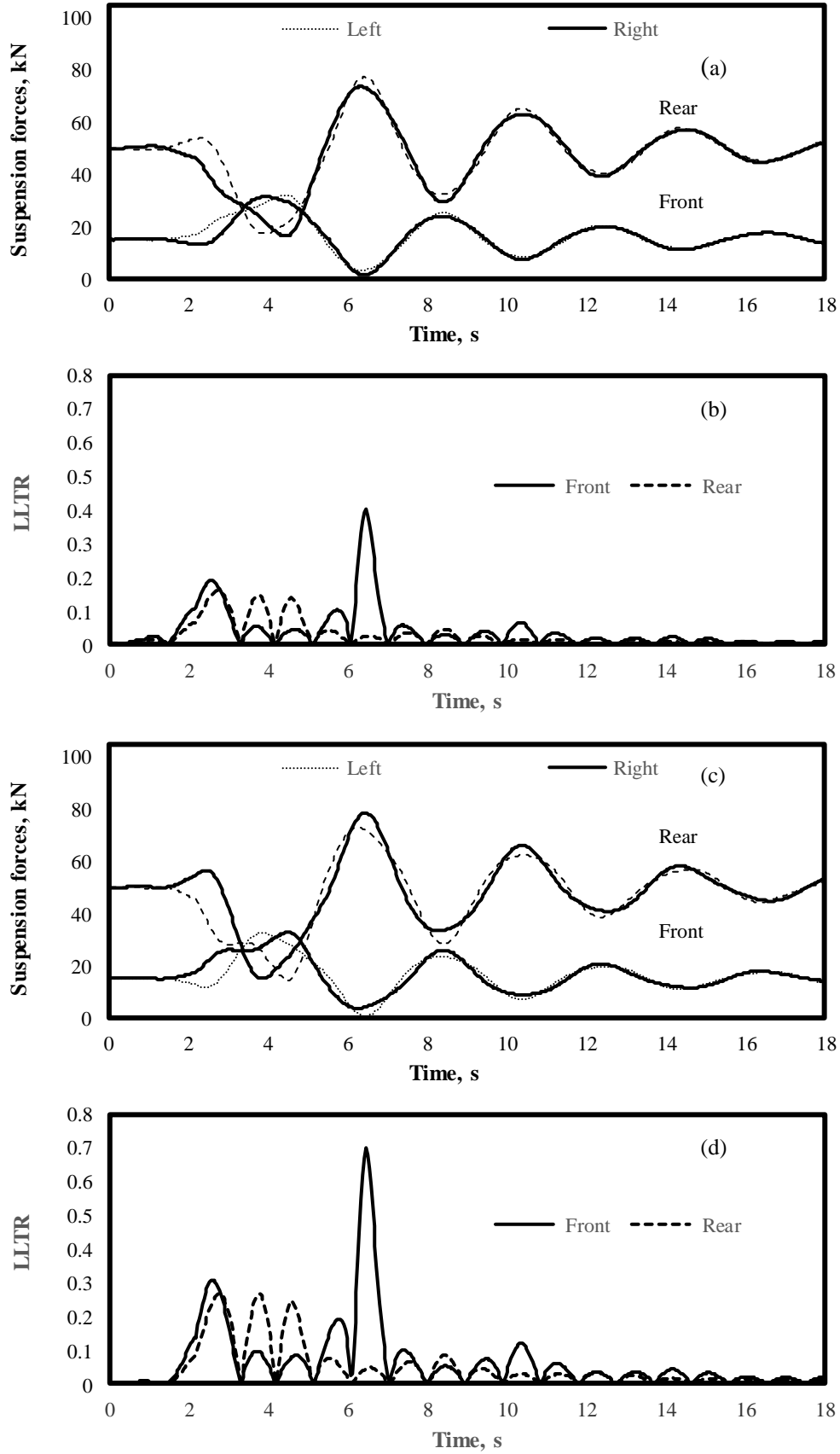


Figure 5. Braking in a turn, two different radiuses, from 30 km/h: (a) and (b) 500 m radius curve; (c) and (d) 325 m radius curve (Sloshing cargo)

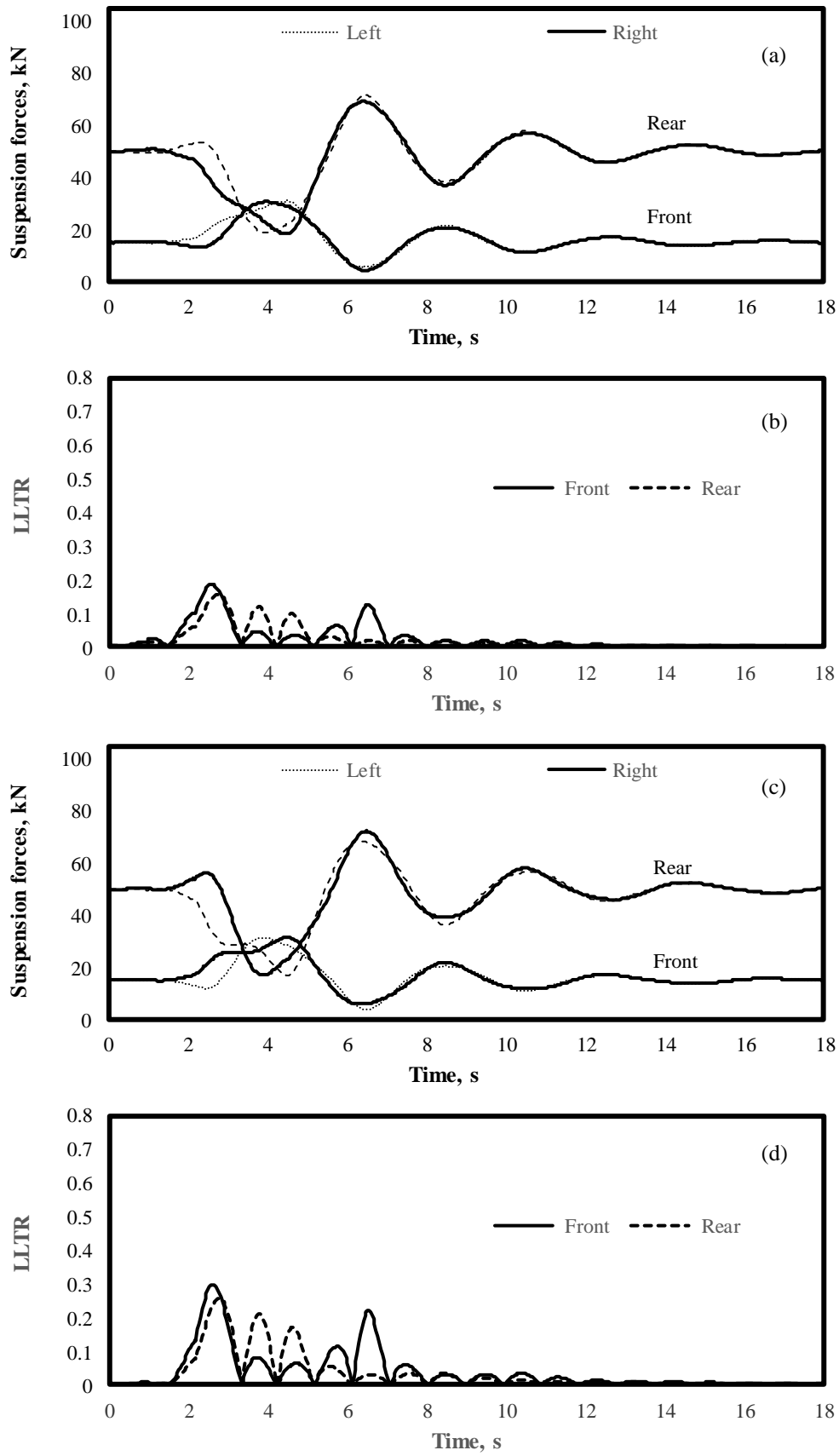


Figure 6. Braking in a turn, several radiuses, from 30 km/h: (a) and (b) 500 m radius curve; (c) and (d) 325 m radius curve (Non-sloshing cargo)

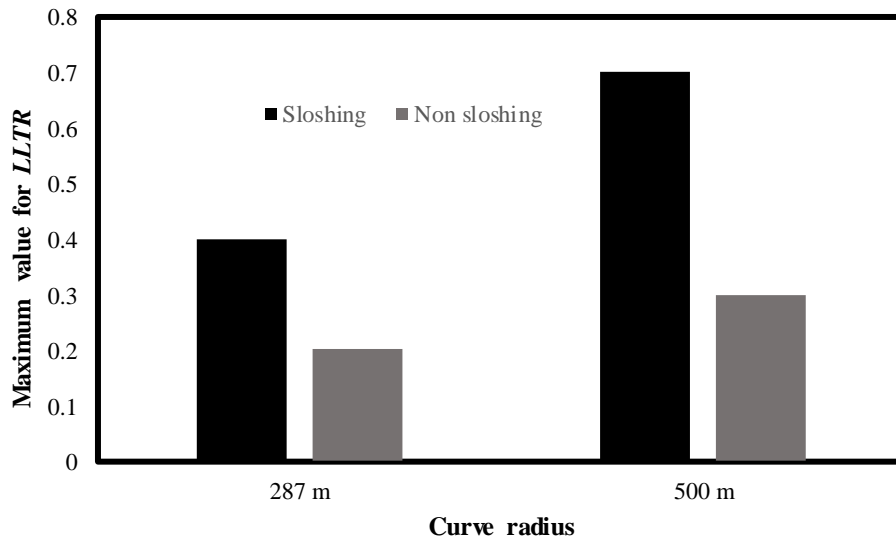


Figure 7. Load transfers during braking in a turn maneuver, for two curve radiuses and cargo condition.

Winding road

The journey of the road tanker along a 4-curve road is now presented. Figure 8 describes the road geometry, involving the merging of actual road curves in a 70 second journey involving a 583 m length road, where the second and third curves have different directions. The formulation described above was used to calculate the instantaneous curving radius for the transition from straight to the constant radius curve. A sample of the instantaneous turning radius is presented in Figure 9, where the maximum values are assumed as straight segments. To assess the effect of the driving style on the lateral stability of the vehicle, two levels of dispersion for the driving speed, are presented. Figures 10 to 12 illustrate the input and performance results for three levels of speed dispersion: great (dispersed), medium and zero. The coefficient of variation (COV) of the rough dispersed speed input, is 12.49%, while the corresponding value for the medium dispersed speed, is 5%.

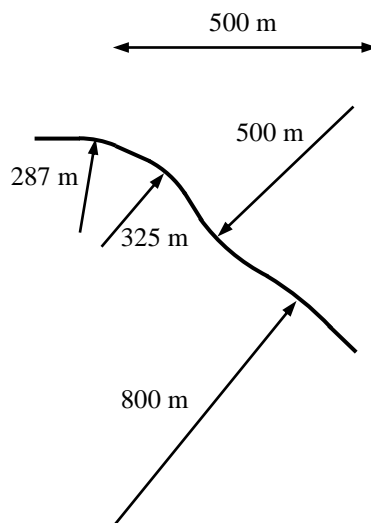


Figure 8. Road geometry for assessing roll stability of a partially filled straight road tanker

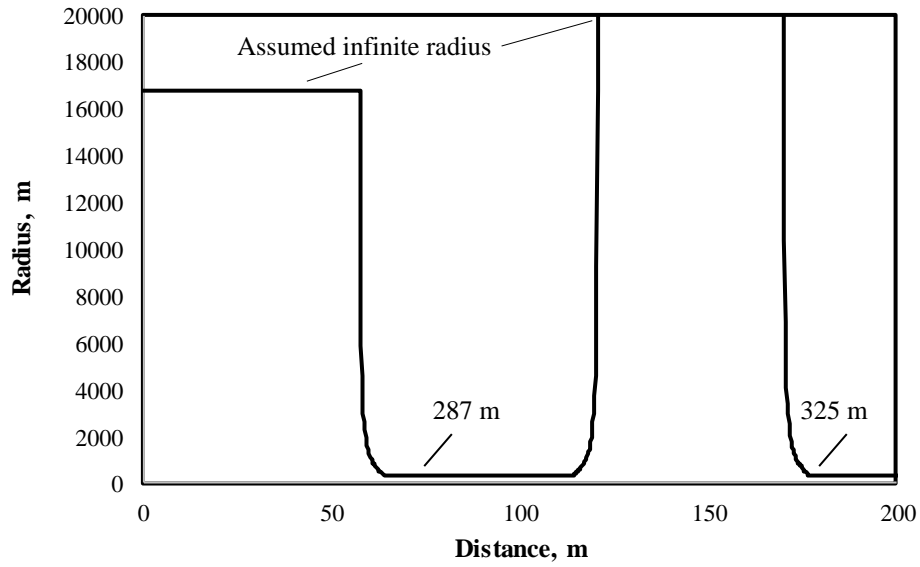


Figure 9. Sample transitional radius for the 287-m and 325-m turns.

According to these results, the dispersion in speed has a significant effect on the lateral stability of the vehicle, as an increase from the medium to the greater dispersion (150% increase) represents an increase in the maximum lateral load transfer from 0.59 to 0.88, that is, an increase of 151%. While this is a linear relationship, it should be noted that the peaks in the *LLTR*, occur in very different positions along the road. On the other hand, eliminating any dispersion of the speed, involves a reduction in the maximum *LLTR* (0.4). Figure 13 illustrates the effect of speed dispersion (COV) on *LLTR* average value. According to these results, increasing from zero to a medium speed dispersion, involves only a small effect on the average *LLTR*, while changing from medium to rough speed dispersion, involves a significant increase of the average *LLTR*.

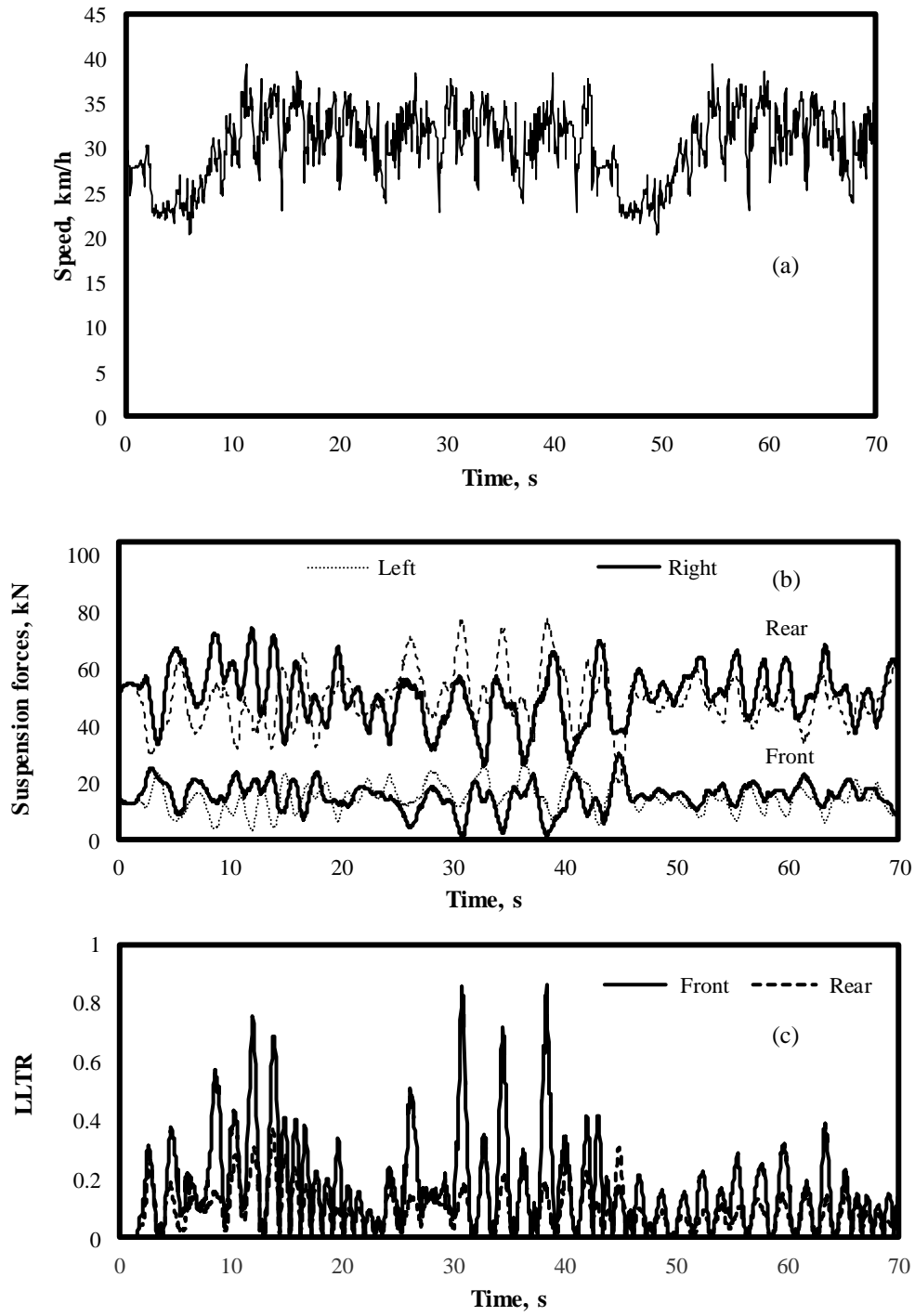


Figure 10. Performance measures for the road tanker on the winding road of Fig. 7, for a disperse speed situation

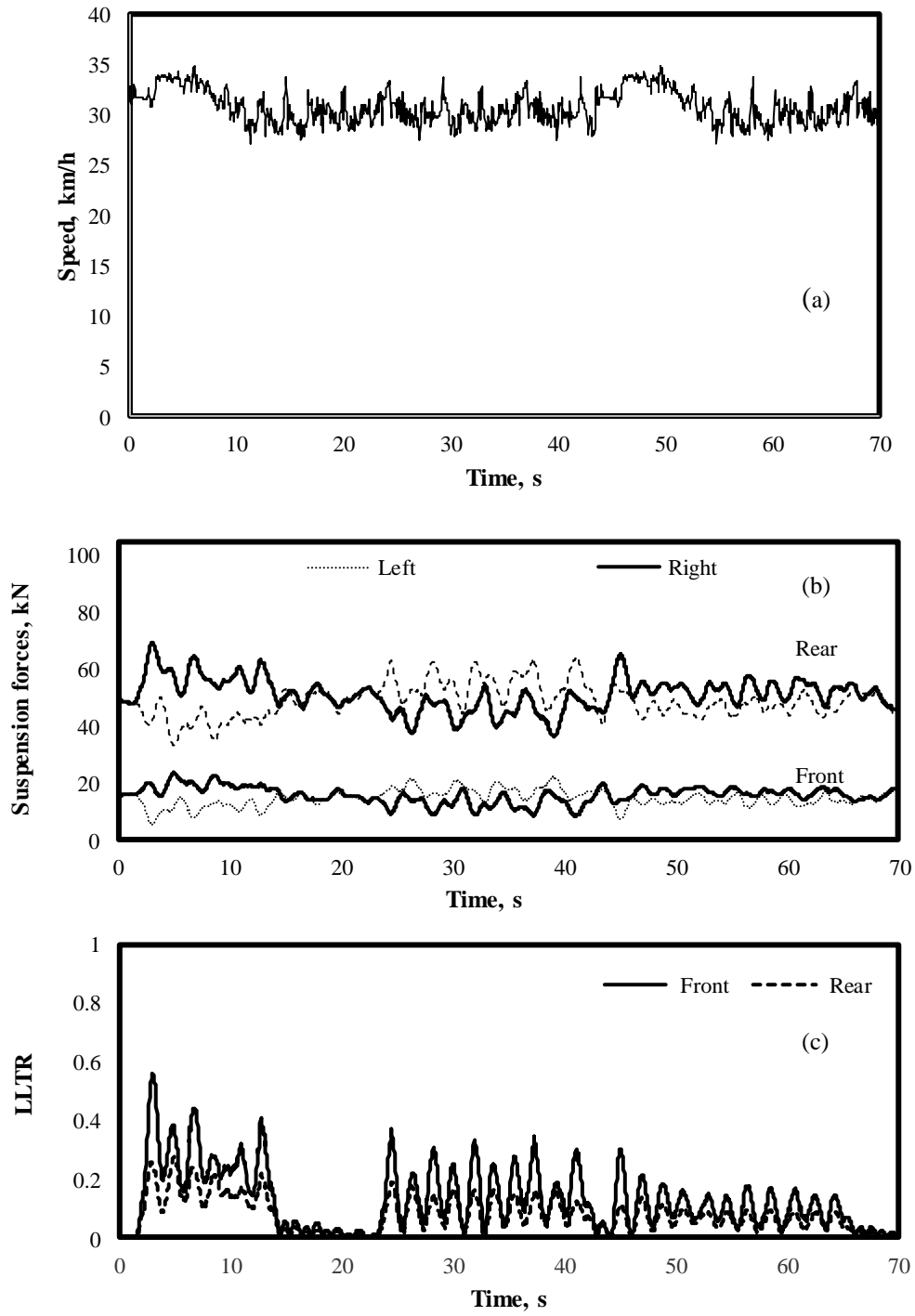


Figure 11. Performance measures for the road tanker on the winding road of Fig. 7, for a medium speed dispersion situation

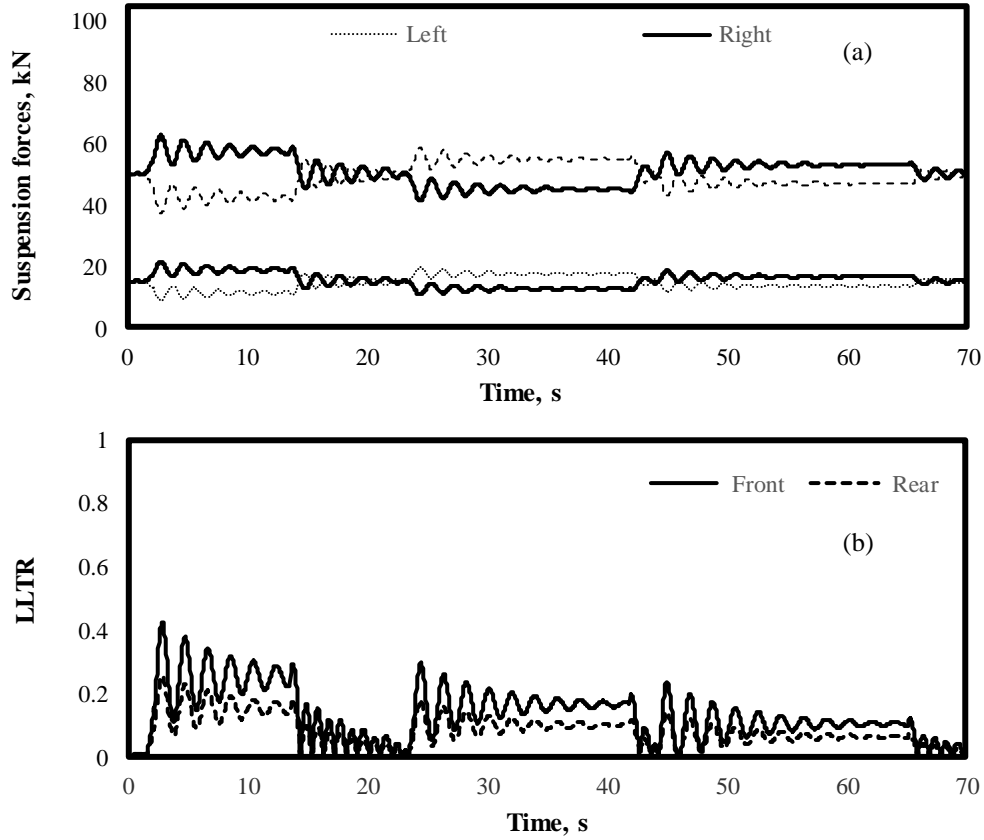


Figure 12. Performance measures for the road tanker on the winding road of Fig. 7, for a constant speed situation (30.64 km/h)

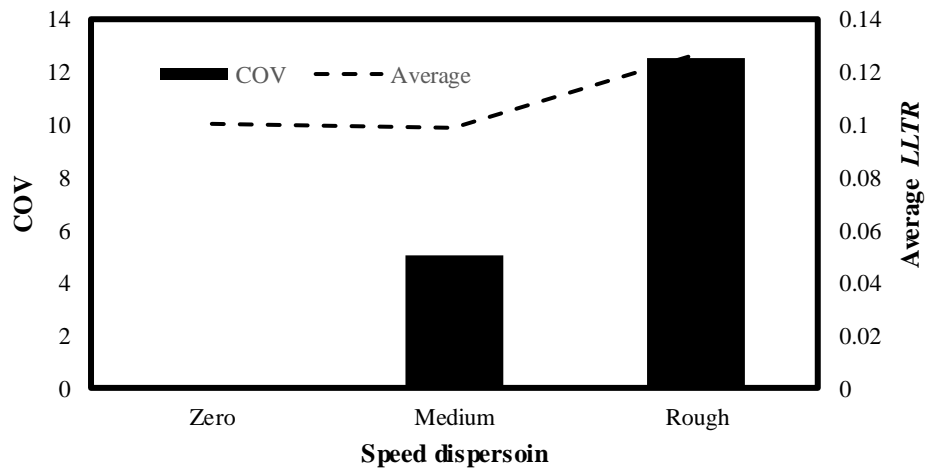


Figure 13. Effect of speed dispersion on COV and average LLTR

Conclusions

A simplified model is being proposed to simulate a braking in a turn maneuver of a straight road tanker. The model involves two uncoupled sets of inverted and simple pendulum models, representing the sprung mass and the sloshing cargo in the vehicle, respectively, for the longitudinal and lateral planes. Several formulations are assembled for this purpose, including the dimensioning of the simple pendulum on the basis of a validated methodology, and the transient radius when the vehicle transits from a straight to a constant radius curve. The results illustrate the longitudinal and lateral load transfer due to the braking in a turn maneuver, suggesting a nonlinear relationship between the magnitude of the radius and the maximum value reached for the selected performance measure (lateral load transfer ratio). On the other hand, the dispersion in the travelling speed also represents significant variations in the vehicle performance, when it runs along a winding road. While the main simplification of the composed model consists of the uncoupling between the longitudinal and the lateral response models, the overall response obtained is congruent with what has been reported in the literature, about the risk associated to braking while turning. Some validation elements should be provided as a continuation this research effort.

References

- [1] NHTSA (2015) Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey. National Highway Traffic Safety Administration. DOT HS 812 115.
- [2] Kolei, A., Rakheja, S., and Richard, R. (2015) Three-dimensional dynamic liquid slosh in partially-filled horizontal tanks subject to simultaneous longitudinal and lateral excitations.
- [3] Dasgupta, A. (2011) Effect of tank cross section and longitudinal baffles on transient liquid slosh in partly-filled road tankers. Master of Science Thesis. Concordia University, Montreal.
- [4] Romero, J.A., and Otremba, F. (2016) Experimental and theoretical modeling of cargo sloshing during braking. IMECE2016-65698.
- [5] Romero, J.A., Hildebrand, R., Ortiz, W., and Gomez, J.C. (2005) Study of roll motions in high-filled tank trucks using minimum instrumentation. SAE paper 2005-01-2264.
- [6] Meirovitch, L. (1986) Elements of vibration analysis. Mc Graw Hill Int. 2nd ed. 345 pp.
- [7] Romero, J.A., Lozano-Guzmán, A.A., and Arroyo-Contreras, G.M. (2015) Proceedings, 2015 IFToMM World Congress. Taipei, Taiwan.
- [8] Elmore, C.W. and Heald, M.A. (1989) Physics of waves, Dover Publications Inc. New York.
- [9] Romero, J.A., Hildebrand, R., Ortiz, W., and Gomez, J.C. (2005B) Natural sloshing frequencies of liquid cargo road tankers, International Journal of Heavy Vehicle Systems 12(2): 121-138.

Fractional Order Derivative Computation with a Small Number of Discrete Input Values

Dariusz W. Brzeziński

Institute of Applied Computer Science, Lodz University of Technology, Poland.

Corresponding & presenting author: dbrzezinski@iis.p.lodz.pl

Abstract

High-accurate computer approximation of fractional derivatives and integrals by applying the Grünwald-Letnikov formula requires generally a high number of input values. If required amount can not be supplied, the accuracy of approximation drops drastically. In the paper we solve a difficult and crucial problem in this scope, i.e. when input data consist only of a small number of discrete values. Furthermore, some of the values may be unusable for computational purposes. Our problem solution include appropriate method of input data preprocessing, an interpolation algorithm with extrapolation abilities, central point function discretization schema, recurrent computational method of coefficients and the application of Horner's schema for the core of the Grünwald-Letnikov method: coefficients and function's values multiplication. Numerical method presented in the paper enables computing fractional derivatives and integrals of complicated functions with much higher accuracy than it is possible when the default approach to the Grünwald-Letnikov method computer implementation is applied. This new method usually takes only 10% of function's values required by the default approach for the same computations and it is much less restrictive for their quality. The general novelty of the method is an efficient configuration of existing numerical methods and an enhancement of their abilities by applying modern programming language - Python and arbitrary precision for computations.

Keywords: Numerical Methods, Finite Differences, Fractional Order Derivatives and Integrals, Accuracy of Numerical Calculations, Arbitrary Precision.

1 Introduction

Fractional calculus (FC) or more aptly calculus of any order has been successfully applied for many areas of technical sciences including electrical engineering, electronics and control systems as well as signals analysis and processing. The application for close-loop control systems require computations to be conducted with high accuracy and in precisely provided time. Otherwise, the system control fails.

The time factor of the computations requires the application of mathematical formulas enabling developing fast and compact computer programs. Despite the existence of numerous formulas for numerical approximation of fractional order derivatives and integrals (FOD/I) [1–6], only the popular Grünwald-Letnikov formula (GL) [7–11] fulfills this requirement. Therefore it became the first choice for the purpose of systems control.

The algorithm of this method consists of multiplication of some coefficients (weights) and function's values.

Figure 1 presents first few subsequent coefficients values of the two utmost fractional orders of derivatives/integrals: 0.2 and 0.8.

The remaining orders between 0 and 1 can be deduced from this figure: the first coefficient has always value 1, the second one has a value of an fractional order (negative for derivatives, positive for integrals). The rest of the coefficients decrease their sum to 0.

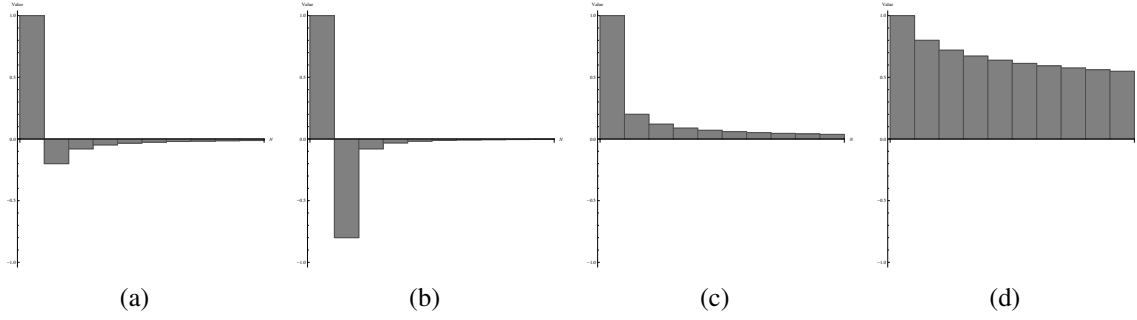


Figure 1: Ten first coefficients for FOD of order: 0.2 (a), 0.8 (b) and FOI 0.2 (c), 0.8 (d).

The values of subsequent coefficients depend on the fractional order value, i.e. higher orders' coefficients assume higher values than lower orders do. Their amount applied for computation have a direct impact on approximation accuracy.

Our previous research on GL method [12] proved that coefficient amount required for certain level of FOD/I approximation accuracy is determined by “the shape of function”. This term describes a behavior of a function in terms of values of its derivatives: (1^{st} and 2^{nd} and higher). If they assume high values, coefficients requirements increase enormously. In this scope: figure 2 presents an indicative number of coefficients (in thousands) required for FOD/I computation of monotonically increasing, monotonically decreasing and constant function in the range $(0, 1)$ with accuracy up to $1.0e^{-04}$ (measured as relative error) by applying a common approach to GL method implementation.

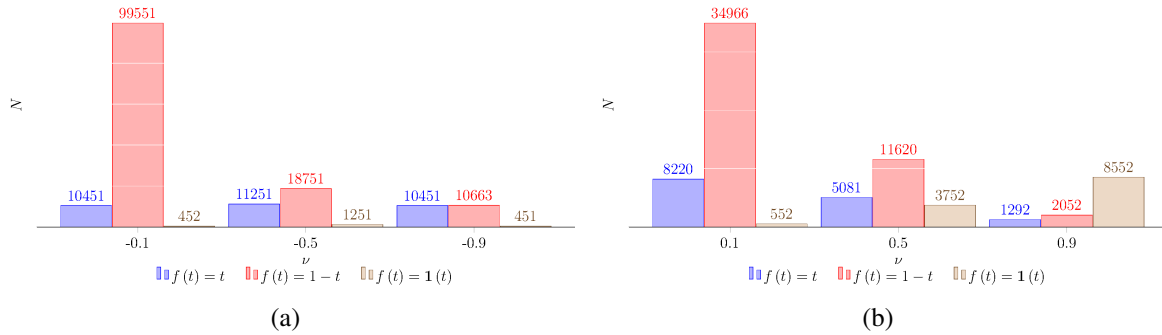


Figure 2: An indicative number of coefficients N required by respective functions and orders ν for FOI (a) and FOD (b) computation with accuracy up to $1.0e^{-04}$.

The selection of functions in figure 2 is on purpose: it shows that computation of lower order FOD/I of simple monotonically increasing functions with accuracy (expressed as relative error calculated in respect to the exact value) as high as four significant decimal places requires application of 600 of coefficients. For higher accuracy of computation and more complicated functions, there are required many hundreds of thousands of them. In fact, calculation with high accuracy (i.e. with more than four significant decimal places) of a low fractional order of FOD for a high frequency periodic function which bounding box is either constant or decreasing, requires over 2 billion of coefficients. This can become a difficult task even for a state-of-the-art computer. This is pictured in figure 3 with the coefficient requirements for exponential and periodical functions and their combination for an arbitrary selected fractional order 0.45.

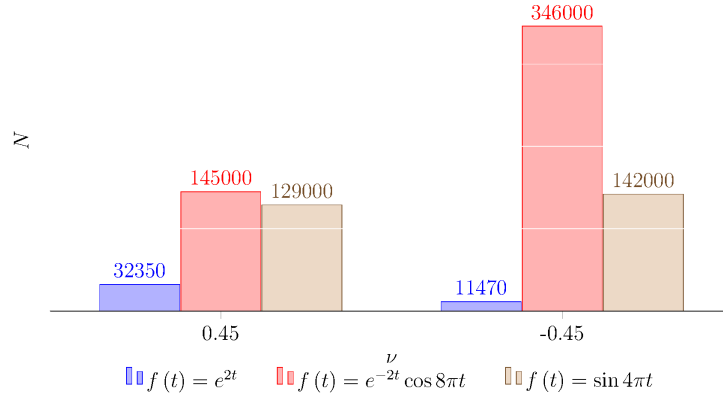


Figure 3: An indicative number of coefficients N required by respective functions and orders ν for FOD and FOI computation with accuracy up to $1.0e^{-04}$.

Long computation time associated with processing of high number of coefficients and multiplying them by function's values makes real-time application difficult. Therefore, our research on GL method was associated from the very beginning with the increase of its efficiency by generally reducing requirements for a number of coefficients for requested accuracy of FOD/I computations. In papers [13, 14] we presented successful attempts achieved by modifying commonly applied form of GL formula with the well known Horner's schema of polynomial's value calculation. Application for real-time computation of an equivalent form of GL formula (Horner's form) enabled us removing from computation up to 20% of coefficients required by application of default approach while still maintaining target accuracy. We refer to it as *calculation tail* of variable length technique.

Furthermore, in paper [15] we presented another successful approach to the research on general accuracy and efficiency increase of GL formula by evaluating some alternative to the commonly applied formulas for coefficients computation and some other forms of GL formula. We also assessed application usefulness for efficiency increase of forward and central point discretization schemas as well as the use of three-point discretization schema.

All the efforts resulted in significant efficiency increase of GL method.

Additional results presented in the same paper showed that the magnitude of errors which influence negatively computational accuracy of scientific calculations can be mitigated or often even eliminated by applying a right form of a mathematical formula and by careful selection of a programming language for its implementation. It usually led to the replacement of the commonly used double precision computer arithmetic with "arbitrary precision" for computation (this term is explained in Section 3).

The following paper presents some interesting results of our latest research on GL formula aimed at solving a practical problem - computation of FOD/I with a small number (often not sufficient for target accuracy) of discrete input data. The number may decrease if some of the data are unusable for calculations with computer. This includes infinite and NaN (not a number) values. In such situation computational accuracy can drop below two digit error expressed in percent. Therefore, for the following research we aimed at increasing computational accuracy to at least two decimal places with the same data amount supplied.

The paper is divided into the following sections: At first there are presented several forms of GL formula which are applied for FOD/I computation. This section also includes a brief description of Horner's form of the GL formula and its application algorithm of "calculation tail" of variable length. Next, there is explained importance of arbitrary precision over standard dou-

ble precision for computations and its positive impact on their accuracy and reliability. Section 4 contains description of computing tools. Section 5 includes details of conducted numerical experiment. Last sections 6 and 7 presents concisely results and conclusions.

2 Mathematical Background

The Grünwald-Letnikov FOD/I approximation method can be represented using the following formula

$${}_{t_0}D_t^{(\nu)} f(t) = \lim_{h \rightarrow 0^+} \frac{1}{h^\nu} \sum_{i=0}^{N-1} (-1)^i \binom{\nu}{i} f(t - ih), \quad (1)$$

in which $\nu \in \mathbf{R}$ is the order of derivative; fractional integral is defined as derivative evaluated for negative order $-\nu < 0$, N denotes an amount of steps in summation, t_0, t is the interval and $h = \frac{t-t_0}{N}$ is the subinterval width.

Formula (1) includes ∞ limitation, which thwarts its computer application. The next formula has it removed, which makes it useful for computational purposes. The formula will be referred as the commonly used form of GL formula for FOD/I computations, in which Γ denotes Euler's Gamma function

$${}_{t_0}D_t^{(\nu)} f(t) \simeq \frac{h^{-\nu}}{\Gamma(-\nu)} \sum_{k=0}^{N-1} \frac{\Gamma(k-\nu)}{\Gamma(k+1)} f(t - kh). \quad (2)$$

The formulas (1) and (2) represent backward-difference. GL formula can be applied with central point discretization schema [16]

$${}_{t_0}D_t^{(\nu)} f(t) \simeq \frac{h^{-\nu}}{\Gamma(-\nu)} \sum_{k=0}^{N-1} \frac{\Gamma(k-\nu)}{\Gamma(k+1)} f\left(t - \left(k - \frac{\nu}{2}\right)h\right) \quad (3)$$

and with the forward point discretization schema

$${}_{t_0}D_t^{(\nu)} f(t) \simeq \frac{h^{-\nu}}{\Gamma(-\nu)} \sum_{k=0}^{N-1} \frac{\Gamma(k-\nu)}{\Gamma(k+1)} f(t - (k - \nu)h). \quad (4)$$

The application of formulas (1)-(4) for computation has some serious restriction: it requires that $t_0 = 0$.

In paper [17] we can find a formula, which removes these restrictions and it has higher accuracy order (second instead of first one)

$${}_{t_0}D_t^{(\nu)} f(t) \simeq \frac{h^{-\nu}}{\Gamma(-\nu)} \sum_{k=0}^{N-1} \frac{\Gamma(k-\nu)}{\Gamma(k+1)} f\left(t - \left(k - \frac{\nu}{2}\right)h\right) + \frac{h^{-\nu}}{\Gamma(-\nu)} \frac{(1+\nu)}{2} f(t_0) N^{-1-\nu}. \quad (5)$$

Before presenting Horner's form of the the Grünwald-Letnikov formula we introduce a discrete version of the formula (2) for $h = 1$ and $t = ih$.

For a given discrete-time, real bounded function $f(k) = f_0, f_1, \dots, f_{k-1}, f_k$ GL formula of

Fractional Order Backward Difference (FOBD) is defined as

$${}_0^{GL}\Delta_k^{(\nu)} f_k = \sum_{i=0}^k a_i^{(\nu)} f_{k-i}, \quad (6)$$

where ν is the FOBD order. Fractional Order Backward Sum (FOBS) is defined as the FOBD evaluated for negative order, f_k is a discrete time function and $a_i^{(\nu)}$ are the coefficients for $i = 0, 1, 2, 3, \dots, k-1, k$.

The coefficients $a_i^{(\nu)}$ can be calculated by applying several formulas, which include the formula involving factorial function calculation

$$a_i^{(\nu)} = \begin{cases} 0 & \text{for } i < 0 \\ 1 & \text{for } i = 0 \\ (-1)^i \frac{\nu(\nu-1)\dots(\nu-i+1)}{i!} & \text{for } i > 0. \end{cases} \quad (7)$$

This formula presents a serious limitation for computational accuracy due to the use of factorial function. It causes the overflow which limits a number coefficients that can be computed to 170. Therefore the following recurrent formula should be applied instead. It is derived from the relation between coefficients $a_i^{(\nu)}$

$$a_i^{(\nu)} = a_{i-1}^{(\nu)} \left(1 - \frac{1+\nu}{i} \right) \text{ for } i > 0. \quad (8)$$

The next formula presents the algorithm (7) expressed in a matrix-vector form

$${}_0^{GL}\Delta_k^{(\nu)} f(k) = \begin{bmatrix} a_0^{(\nu)} & a_1^{(\nu)} & \dots & a_k^{(\nu)} \end{bmatrix} \begin{bmatrix} f_k \\ f_{k-1} \\ \vdots \\ f_0 \end{bmatrix}. \quad (9)$$

The Horner form of GL formula is a formula to which the well known Horner's schema of polynomial's value calculation is applied. Horner's schema possesses some significant computational advantages, e.g. lower computational complexity and a natural method of data input for computation.

By applying the same assumptions as in case of the Grünwald-Letnikov definition, introducing new coefficients $c_i^{(\nu)}$

$$c_i^{(\nu)} = \begin{cases} 0 & \text{for } i < 0 \\ 1 & \text{for } i = 0 \\ 1 - \frac{1+\nu}{i} & \text{for } i > 0 \end{cases} \quad (10)$$

we apply Horner's schema to (6). We obtain [18]

$${}_0^H\Delta_k^{(\nu)} f(k) = c_0^{(\nu)} \left[f_k + c_1^{(\nu)} \left[f_{k-1} + c_2^{(\nu)} \left[f_{k-2} + \dots + c_{k-1}^{(\nu)} \left[f_1 + c_k^{(\nu)} [f_0] \right] \right] \right] \right]. \quad (11)$$

The formulas (6) and (11) are equivalent, i.e.

$${}_0^{GL}\Delta_k^{(\nu)} f_k = {}_0^H\Delta_k^{(\nu)} f_k.$$

However,

$$\begin{aligned}\lim_{i \rightarrow \infty} a_i^{(\nu)} &= 0, \\ \lim_{i \rightarrow \infty} c_i^{(\nu)} &= 1,\end{aligned}\tag{12}$$

and for some $i > k - L$

$$\begin{aligned}a_i^{(\nu)} &\approx 0, \\ c_i^{(\nu)} &\approx 1\end{aligned}$$

We can use the property of $c_i^{(\nu)}$ coefficients to modify the formula (11)

$$\begin{aligned}{}_0^{Hs} \Delta_{k,L}^{(\nu)} f(k) \\ = \begin{cases} c_0^{(\nu)} \left[f_k + c_1^{(\nu)} \left[f_{k-1} + c_2^{(\nu)} \left[f_{k-2} + \dots + c_{k-1}^{(\nu)} \left[f_1 + c_k^{(\nu)} [f_0] \right] \right] \right] \right] & \text{for } k \leq L \\ c_0^{(\nu)} \left[f_k + c_1^{(\nu)} \left[f_{k-1} + \dots + c_{k-L}^{(\nu)} \left[\sum_{i=0}^L f_i \right] \right] \right] & \text{for } k > L. \end{cases}\end{aligned}\tag{13}$$

Application of formula (13) enables reducing up to 20% requirements for an amount of coefficients during FOD/I computation, i.e. for samples $k > L$ we assume $c_i^{(\nu)} = 1$ and replace multiplication with summation of the remaining function values.

The L value is an arbitrary number and is to be set empirically. However, its value must be selected carefully, i.e. the removal of L coefficients is not to decrease accuracy of calculations over a permissible error.

Based on our past experience, we apply $L = 20\%$ less coefficients required for FOD/I computation with target accuracy by the default approach (see as examples numbers of coefficients presented in Figures 2 and 3), e.g. if a function requires 600 coefficients for the accuracy $1.0e^{-04}$, then we calculate 20% of this number (which in the case is 120). From now on, $L = 480$ for the use with formula (13). After this operation, there are applied $600 - 120 = 480$ coefficients.

For full reasoning and procedure description how to set L , please refer to the papers mentioned in introduction.

The next table presents computational accuracy decrease over $1.0e^{-04}$ after removal of $L = 20\%$ of former number of coefficients required for this accuracy (see again Figures 2 and 3) by applying formula (13).

Table 1: Relative error increase over $1.0e^{-04}$ after removal of $L = 20\%$ of coefficients number required for this accuracy for selected fractional orders ν (positive - FOD, negative - FOI).

	Function $f(t)$				
	t	$1-t$	$\mathbf{1}(t)$	$e^{-2t} \sin(8\pi t)$	$\cos(4\pi t)$
0.1	0.28	0.02	2.49	4.39	0.1
0.5	2.15	0.08	2	1.05	0.25
0.9	5.25	0.04	0.53	0.09	0.17
-0.1	0.22	0	2.44	8.84	0.17
-0.5	0.6	0.04	1.74	7.32	1.79
-0.9	0.19	0.01	0.42	3.16	7.96

3 The Importance of Arbitrary Precision

Limiting factor to the accuracy of computation by applying the formulas (2)-(5) is a precision which uses the computer to store data that are being supplied to it.

Consider calculating the value of $\Gamma(z)$ for large values of z . At $z = 171$ the approximate value provided by the computer programmed to use standard double precision arithmetic is $7.26e306$. At $z = 171$ the computer begins to refer to the value of $\Gamma(z)$ as 'Inf'. Hence, for the formulas (1)-(2) there is no practical use of calculating coefficients beyond the 171st.

The above example of overflow occurrence shows that the selection of uniform C++ equipped with the standard mathematical library as a main programming tool is not enough nowadays to take full advantage of available hardware. The application of arbitrary precision computing for increasing the accuracy and the correctness of numerical calculations and Nvidia CUDA parallelization technology for their effectiveness, are the best examples in this context [19].

Application of arbitrary precision makes it possible for the user to choose precision for calculation and for each variable storing a value. It is not machine-dependent or IEEE standard types. With its help we can - among the others - increase general accuracy of mathematical computations. However, its application purpose is above all to increase accuracy of numerical calculations, e.g. by eliminating under- and overflows, increasing accuracy of a polynomial zeros finding and derivative and integral calculating.

The importance of elimination of limited precision in computer calculations was aptly presented by Toshio Fukushima in *The Astronomical Journal* in 2001 by giving the following example: "In the days of powerful computers, the errors of numerical integration are the main limitation in the research of complex dynamical systems, such as the long-term stability of our solar system and of some exoplanets [...]" and gives an example where using double precision leads to an accumulated round-off error of more than 1 radian for angular position of planets [20].

Double precision computer arithmetic is optimized for speed and has many flaws which negatively influence the accuracy of computations, e.g. limitations of number values which double precision variables can hold or no programmer influence on mathematical operations rounding.

However, it is the lack of clarity in handling of intermediate results which troubles the most, i.e. the floating-point standard [21] only defines that the results must be rounded correctly to the destination's precision and fails to define the precision of destination variable. This choice is commonly made by a system or a programming language. The user can not influence it in any way. Therefore, the same program can return significantly different results depending on the implementation of the IEEE standard.

Arbitrary precision application is applied in conjunction with special libraries which include their own data structures and mathematical functions.

There are many programming languages, which can be used with arbitrary precision. They include Python. Python is an object oriented script language, which achieves a higher abstract level than for example C++, i.e. an individual programmer can achieve the same results in a much shorter time and with far fewer lines of code. It also has especially clean and straightforward syntax. It can lead to programs' shorter executing time.

An important advantage of programming using Python is availability of ready to use libraries. They enable solving a scientific problem by focusing rather on selecting the right tools and by adopting them if necessary instead of designing a new algorithm from ground up. Therefore, we selected it as the main programming language for our research.

4 Tools for the Problem Solving

The requirement of an efficient solution to the problem presented in the introduction includes constructing an interpolation polynomial using supplied values. It is crucial for arbitrary selection of a discretization point schema, i.e. central (10) or forward (11) one or the one with three points.

We applied SciPy - an open source Python library for scientific and technical computing. SciPy provides a module for interpolation based on the FITPACK library of FORTRAN functions, which is assumed as reliable.

Central point discretization formula (3) requires at least one point beyond t to be accessible for the interpolation. Unfortunately standard routines in SciPy do not allow to include any points from outside the interpolation range. At first our solution to make `scipy.interpolate` give an extrapolated result beyond the input range included modifying an interpolating algorithm based on spline interpolation by adding:

- Constant extrapolation: extrapolating left and right values as constant beyond the range
- Linear extrapolation: writing a wrapper around an interpolation function, which does linear extrapolation
- Manually inserted points and values to the initial array
- `scipy.interpolate.splrep` (with degree 1 and no smoothing)

Due to unsatisfactory accuracy and speed concerns, we have decided to apply `InterpolatedUnivariateSpline` from the same library `scipy.interpolate` instead. It does interpolation and extrapolation and can be applied in conjunction with `mpmath`. `mpmath` is a free (BSD licensed) Python library for real and complex floating-point arithmetic with arbitrary precision. It is based on GNU GMP and GNU MPFR libraries. It enables switching from double to arbitrary precision computation by applying Python programming language.

The mathematical library `mapmath` is required not only to increase overall exactness of computations, but also due to accessibility of excellent implementations of gamma and reciprocal gamma functions required for FOD/I calculations by applying formulas (2)-(5).

5 Details to the Numerical Experiment

In our previous numerical experiments with GL method the number of coefficients (and function's values at the same time) was commonly limited to 600, because we applied such an amount for real-time calculations (in our case the maximum number of coefficients was determined by amount of available memory in a test DSP-system).

However, to present high efficiency of FOD/I computing method presented in the paper, our experiment was conducted with only 60 values as an input. Additionally, up to 20% of them were randomly assigned ∞ and NaN values to mimic unsuitability for computation, which often occur if experiment data were collected without the knowledge of computing input data requirements.

Unlike simple, monotonically increasing, decreasing or constant functions, which require relatively moderate number of coefficients (up to 600) for accuracy up to four significant decimal places, high-frequency, exponential and periodical functions, which have "dramatical shape changes" require millions or more of them. For this reason, general efficiency and accuracy of the method was assessed against two sets of functions.

Set one included constant, monotonically increasing and decreasing functions:

$$\mathbf{f}_1 \quad f(t) = \mathbf{1}(t) \in (0, 1)$$

$$\mathbf{f}_2 \quad f(t) = tt - t, t \in (0, 1)$$

$$\mathbf{f}_3 \quad f(t) = t^{0.1}, t \in (0, 1).$$

Set two included more complicated functions:

$$\mathbf{f}_4 \quad f(t) = te^{-t}, t \in (0, 10)$$

$$\mathbf{f}_5 \quad f(t) = e^{-2t} \cos 8\pi, t \in (0, 1)$$

$$\mathbf{f}_6 \quad f(t) = e^{-2t}, t \in (0, 1)$$

$$\mathbf{f}_7 \quad f(t) = e^t \sin t, t \in (0, \pi)$$

$$\mathbf{f}_8 \quad f(t) = 1.5 \cos 2t + 2.2 \cos 4t, t \in (0, 2\pi)$$

$$\mathbf{f}_9 \quad f(t) = \sin 2\pi \cos t, t \in (0, 5).$$

The functions were tabulated before supplying them to the program, i.e. the actual input to a program was in form of a vector with discrete values.

At first we applied interpolation method described in section 4 to construct an interpolation polynomial using the vector with the supplied values. Next, we interpolated 600 values for FOD/I calculations using central (3) and forward (4) discretization schemas. Finally FOD/I were computed.

For comparison purposes, FOD/I were computed by applying the commonly used reference formula (6) with 60 values as well to present the real-life accuracy.

The FOD/I exact values required for accuracy assessment were computed by applying analytical formulas (if available) or by using high-accuracy integration method [22, 23]. This method involves Gauss-Jacobi Quadrature application for integration and is reliable for computation of FOD/I using Riemann-Liouville and Caputo formulas [24] with accuracy up to 120 significant decimal places.

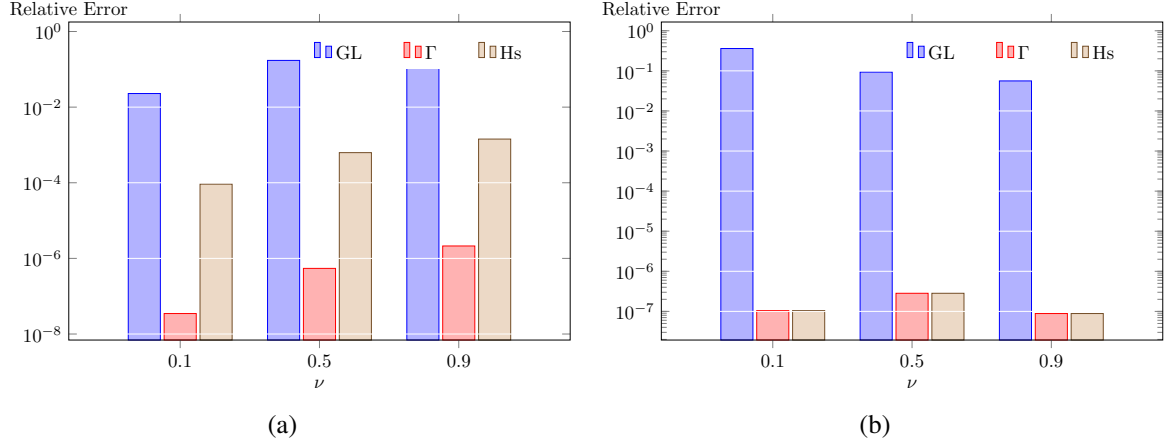
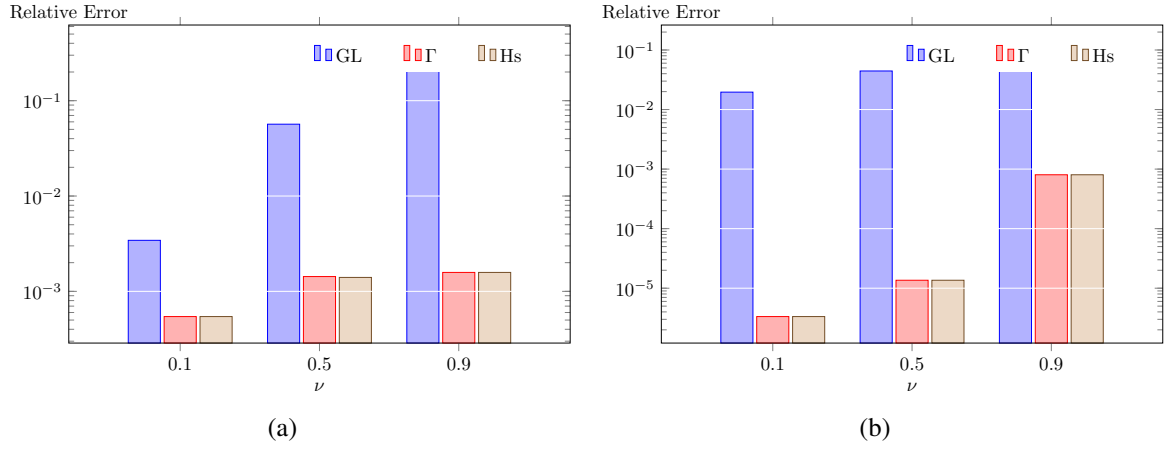
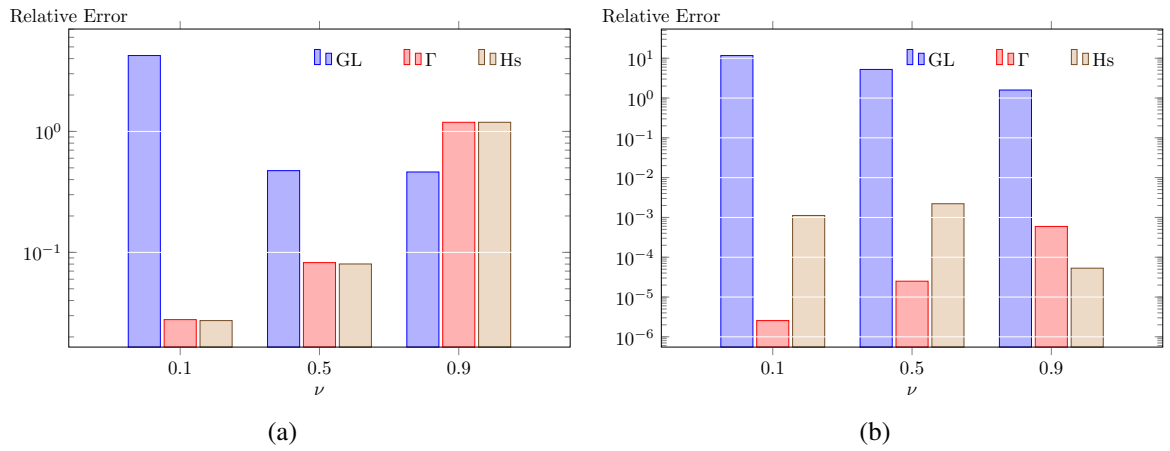
6 Results

FOD/I computational accuracy is assessed as the relative error

$$e_r(m) = \left| 1 - \frac{v_c}{v_e} \right|, \quad (14)$$

in which: v_c is a calculated value, v_e is a value assumed as exact and m denotes a number of source input values for FOD/I computation.

Figures 4-8 present accuracy of FOD/I computation denoted as relative error (14) for: GL - a classical Grünwald-Letnikov formula (2) applied with $m = 60$ function's values (to present the accuracy of computations, which can be expected by applying only real supplied functions' values); Γ denotes computations by applying the formula (5) and Hs - the formula (13) - both with 600 interpolated values, which are computed by applying the combination of techniques described in the paper (to present the accuracy which can be obtained despite the only 60 supplied real function's values).


 Figure 4: Computational accuracy of FD, order ν for: (a) f_1 and (b) f_2 .

 Figure 5: Computational accuracy of FD, order ν for: (a) f_3 and (b) f_4 .

 Figure 6: Computational accuracy of FD, order ν for: (a) f_5 and (b) f_6 .

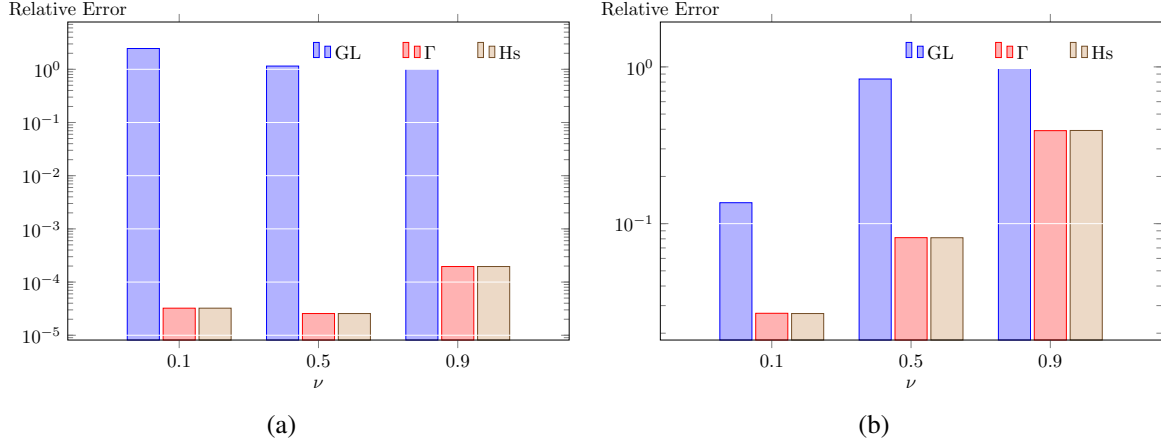


Figure 7: Computational accuracy of FD, order ν for: (a) f_7 and (b) f_8 .

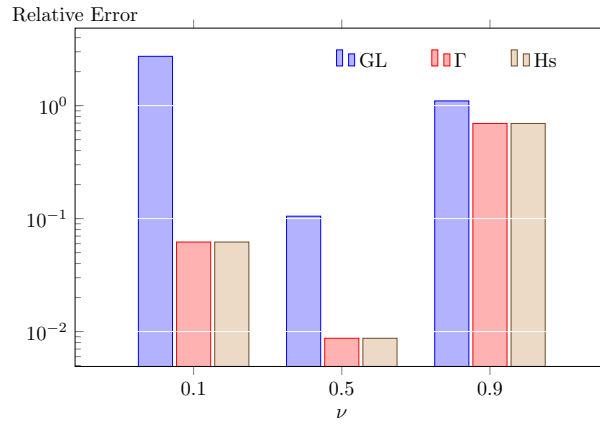


Figure 8: Computational accuracy of FD, order ν for f_9 .

7 Conclusions

The Grünwald-Letnikov method requires a high number input values for high accuracy FOD/I computation. If required number of values can not be supplied, the accuracy of computation drops drastically, particularly for complicated functions.

To solve this deficiency problem we proposed a new numerical method combining existing numerical techniques, arbitrary precision computation and modern programming language.

The method presented in the paper enables computation of FOD/I with similar or higher accuracy with only 10% of input values required by the traditional approach to FOD/I computation by applying the Grünwald-Letnikov method.

Application of the simplified Horner's form of the Grünwald-Letnikov formula for the core of the method (multiplication of the coefficients and function's values) decreases again by up to 20% of those 10% the requirements for input values during computation without noticeable accuracy drop over the assumed level.

Additionally, the developed method enables "repairing" input values unusable for computational purposes.

The computational method described in the paper combines several programming techniques that include application of Python programming language and accompanying mathematical library (mpmath) for arbitrary precision of computations. This combination enables

elimination of common errors associated with double precision computer mathematics and increases significantly accuracy and reliability of scientific computation.

Acknowledgement

The work was created as a result of the research project no. 2016/23/D/ST6/01709 financed from the funds of the National Science Center, Poland.

References

- [1] R. Herrmann. *Fractional Calculus. An Introduction for Physicists*, 2nd. ed. World Scientific, Singapore, 2014.
- [2] D. Baleanu, K. Diethlem, E. Scalas, and J.J. Trujillo. *Fractional Calculus. Models and Numerical Methods*. World Scientific, Singapore, 2012.
- [3] J. T. Machado. Numerical calculation of the left and right fractional derivatives. *Journal of Computational Physics*, 293:96–103, 2015.
- [4] L. Debnath and D. Bhatta. *Integral Transforms and Their Applications, Third Edition*. CRC Press, Taylor & Francis Group, Boca Raton London New York, 2015.
- [5] P. J. Nahin. *Inside Interesting Integrals*. Springer-Verlag, NY, 2015.
- [6] C. Li and F. Zeng. *Numerical Methods for Fractional Calculus*. Chapman & Hall, 2015.
- [7] I. Podlubny. *Fractional Differential Equations*. Academic Press, INC, San Diego Ca, 1999.
- [8] M. D. Ortigueira, J. A. T. Machado, and J. Sa da Costa. Which differintegration? *Proceedings - Vision, Image and Signal Processing*, 152(6), 2005.
- [9] M. D. Ortigueira. Fractional central differences and derivatives. *Journal of Vibration and Control*, 14(9-10):1255–1266, 2008.
- [10] M.D. Ortigueira. *Fractional Calculus for Scientists and Engineers*. Springer-Verlag, NY, 2011.
- [11] M. D. Ortigueira and J. A. T. Machado. What is a fractional derivative. *J.Comput. Phys*, 2014.
- [12] D. W. Brzeziński and P. Ostalczyk. About utility of the simplified grünwald-letnikov formula equivalent horner form. *Discontinuity, Nonlinearity, and Complexity. Special Issue: "Fractional Dynamics and Systems with Power-Law Memory"* (Eds. M.Edelman and J. Tenreiro Machado), 4(4), 2015.
- [13] D. W. Brzeziński and P. Ostalczyk. The grünwald-letnikov formula and its horner's equivalent form accuracy comparison and evaluation for application to fractional order pid controller. In *IEEE Explore Digital Library: IEEE Conference Publications-17th International Conference On Methods and Models In Automation and Robotics (MMAR)*, pages 579–584, 2012.

- [14] P. Ostalczyk, D. W. Brzeziński, P. Duch, M. Łaski, and D. Sankowski. The variable, fractional-order discrete-time pd controller in the iisv1.3 robot arm control. *Central European Journal of Physics*, 11(6):750–759, 2013.
- [15] D. W. Brzeziński and P. Ostalczyk. About accuracy increase of fractional order derivative and integral computations by applying the grünwald-letnikov formula. *Communications in Nonlinear Science and Numerical Simulation*, 40:151–162, 2016.
- [16] K. Oldham and J. Spanier. *The Fractional Calculus. Theory and Applications of Differentiation and Integration to Arbitrary Order*. Academic Press, INC, San Diego Ca, 1974.
- [17] Y. Takeuchi and R. Suda. New numerical computation formula and error analysis of some existing formulae in fractional derivatives and integrals. In *Proceedings to the Fifth Symposium on Fractional Differentiation and its Applications*, Hohai University, Nanjing, China, 2012.
- [18] P. Ostalczyk. Fractional-order backward difference equivalent forms. In *Fractional Differentiation and Its Applications. Systems Analysis, Implementation and Simulation, System Identification and Control*. 1995.
- [19] J. M. Müller, N. Brisebarre, F. De Dinechin, C. P. Jeannerod, V. Lefevre, G. Melquiond, N. Revol, D. Stehle, and S. Torres. *Handbook of Floating-Point Arithmetic*. Birkhauser, New York, NY, 2010.
- [20] K. R. Ghazi, V. Lefevre, P. Theveny, and P. Zimmermann. Why and how to use arbitrary precision. *IEEE Computer Society*, 12(3):1–5, 2001.
- [21] Microprocessor Standards Committee. *IEEE Standard for Floating-Point Arithmetic*, 2008. <http://dox.doi.org/10.1109/IEEESTD.2008.4610935>.
- [22] D. W. Brzeziński and P. Ostalczyk. High-accuracy numerical integration methods for fractional order derivatives and integrals computations. *Bulletin of the Polish Academy of Sciences Technical Sciences*, 62(4):723–733, 2014.
- [23] D. W. Brzeziński. Accuracy problems of numerical calculation of fractional order derivatives and integrals applying the riemann-liouville/caputo formulas. *Applied Mathematics and Nonlinear Sciences*, 1(1):23–43, 2016.
- [24] D. W. Brzeziński. Comparison of fractional order derivatives computational accuracy - right hand vs left hand definition. *Applied Mathematics and Nonlinear Sciences*, 2(1):237–248, 2017.

Interface effect on failure of ceramic coating/alloy substrate systems

Hao Long^{1,3}, *†Lihong Liang^{1,3, a}, Yueguang Wei²

¹LNM, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190

²College of Engineering, Peking University, Beijing 100871, China

³University of Chinese Academy of Sciences, Beijing 101408, China

*Presenting author: lianglh@lnm.imech.ac.cn

†Corresponding author: lianglh@lnm.imech.ac.cn

Abstract

The interface cohesive zone model is usually used in the finite element method to describe the damage and fracture of interfaces between two layers, two kinds of materials, or two segments in one kind of material or a layer. An introducing of interface cohesive elements in one layer should not affect the layer's properties, that how to take the thickness and stiffness of the interface cohesive elements is studied firstly and the related criterion is given. A finite element model of ceramic coating/alloy substrates under three-point bending loading with the interface cohesive elements inserted into coatings is developed, and the transverse crack evolution of coatings is studied. The simulation results indicate that the coating cracking is later and the crack length decreases with increasing interface toughness, i.e., the damage, defined by a total crack length, is slower with increasing fracture toughness. It can explain the experimental results that damage rate of nanostructured thin coatings with smaller cohesive energy is larger than that of conventional coatings with microscale microstructure, because the fracture toughness is proportional to the cohesive energy of coatings. The effect of cohesive strength on coating damage changes at a critical strength, when the cohesive strength is larger than the critical value, the crack length and damage rate increase with decreasing interface strength.

Keywords: Interface cohesive element, coatings, cohesive strength, fracture toughness.

Introduction

Ceramic coatings are widely used in mechanical, electronic, chemical engineering fields due to its better properties such as wear resistance, erosion resistance, and thermal protection. Once ceramic coatings crack, their function will lose. Therefore, the study on cracking behavior and mechanism of coating/substrate systems attracts great attention [1-5]. Crack density of thin films under tensile stress was predicted based on fracture mechanics model [2] or by developing an elastic-plastic shear-lag model [3]. Crack distribution of ceramic coatings was observed in the in-situ bending experiments by scanning electron microscope [4]. In order to study systematically crack and damage evolution of coating systems, finite element method (FEM) is a good choice. Interface cohesive zone model (CZM) is an effective tool to characterize cracking and is often introduced in FEM [5]. However, the introduction of CZMs should not affect the original mechanical properties of materials before cracking.

In this paper, the stiffness criterion of CZM with finite thickness is given firstly, then which is used to simulate cracking of ceramic coatings bonded on alloy substrates under three-point bending loading. Interface cohesive strength and fracture toughness effects on cracking

^a The work is supported by NSFC of China (Grant Nos. 11672296, 11372318).

damage and fracture behavior of coatings are characterized and the related mechanism is revealed.

Selection of cohesive element

In order to ensure that the insertion of cohesive elements does not affect original mechanical properties, the elastic constant E_{eq} of the equivalent continuous medium should be equal to that of the matrix material E_m correspondingly. A representative volume element (RVE) can be selected for the system is in tension, as shown in Fig. 1(a). Due to the introduction of cohesive elements, the length of the RVE increases from L to $L' = L + t_0$, t_0 is thickness of the cohesive element and interface stress-displacement relation is showed in Fig. 1(b). The elongation of the RVE can be divided into two parts: elongation of the matrix material and elongation of cohesive element. The normal strain of the RVE in the x direction and thus the equivalent Young's modulus E_{eq} can be expressed. Since it is required that $E_{eq} = E_m$, we obtain

$$k_n t_0 = E_m \quad (1)$$

where k_n is the stiffness of interface cohesive model as shown in Fig. 1(b).

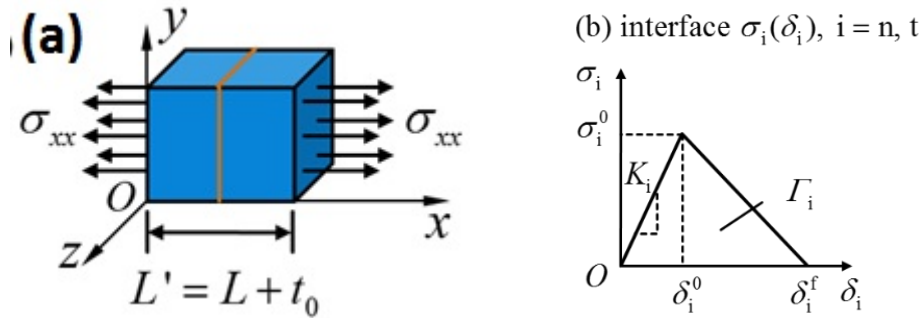


Figure 1. (a) Representative volume element of the system composed of matrix material and cohesive elements in uniaxial tension; (b) Interface cohesive zone model as cohesive elements.

It should be noted that Eq. (1) can be rewritten in the form of the ratio of two lengths: $t_0 / \delta_n^0 = E_m / \sigma_n^0$ with interface cohesive strength σ_n^0 and the corresponding critical displacement δ_n^0 as shown in Fig. 1(b), Γ_n is interface fracture toughness with the subscript n denoting normal direction and t tangential direction. Since the ratio of strength and Young's modulus is about $E_m / \sigma_n^0 \approx 10^2 - 10^3$, the thickness of cohesive elements is two or three orders of magnitude larger than the critical displacement. When thickness of cohesive elements is very small, a large stiffness should be selected based on Eq. (1).

Finite element model of ceramic coating/alloy substrate systems

Ceramic coating/substrate systems are assumed to be under the plane strain condition and the 2D FEM analysis is carried out using the commercial software ABAQUS. Due to symmetry, only the left half of the model is considered, as shown in Fig. 2. The model includes two layers: substrate with thickness h_s of 1.2 mm and ceramic coating with thickness h_c of varying a range compared with the experimental samples. The span length is 16 mm. The vertical loading displacement w is applied on the indenter. Ceramic coating is considered as linear elastic material with Young's modulus of E_c and Poisson's ratio of ν_c [4]. Superalloy substrate is assumed to be elastic-plastic material with Young's modulus of E_s and Poisson's ratio of ν_s ,

and its constitutive relation can be referred to Ref. [4]. The four-node plane strain reduced integration elements (CPE4R) are selected to mesh substrate and coating.

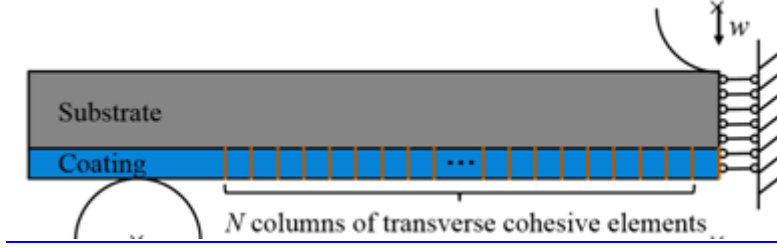


Figure 2. Coating/substrate model under three-point bending loading.

Only transverse cohesive elements with thickness of $t_0^{(T)}$ are inserted into the coating, as shown in Fig. 2. The four-node cohesive elements (COH2D4) are inserted into the coating. The strength and fracture toughness of transverse cohesive elements actually refer to coating strength and coating fracture toughness respectively as shown in Fig. 1(b). For simplicity, values of strength and fracture toughness of cohesive elements in normal and tangential directions are assumed to be the same [6], i.e., $\sigma_n^{0(T)} = \sigma_t^{0(T)}$, $\Gamma_n^{(T)} = \Gamma_t^{(T)}$. The thickness of cohesive elements is selected as $t_0^{(T)} / h_s = 1 \times 10^{-4}$. According to Eq. (1), dimensionless stiffness of transverse cohesive elements is selected as follows:

$$\begin{cases} \frac{K_n^{(T)} h_s}{\sigma_Y} = \frac{E_c / \sigma_Y}{(1 - \nu_c^2) t_0^{(T)} / h_s} = 2.34 \times 10^5 \\ \frac{K_t^{(T)} h_s}{\sigma_Y} = \frac{1 - \nu_c}{2} \frac{K_n^{(T)} h_s}{\sigma_Y} = 9.38 \times 10^4 \end{cases} \quad (2)$$

Dimensionless coating strength and coating fracture toughness are $\sigma_n^{0(T)} / \sigma_Y = 0.03-0.34$ and $\Gamma_n^{(T)} / (\sigma_Y h_s) = (1.0-5.1) \times 10^{-5}$, respectively, $\sigma_Y = 800$ MPa [4]. Each of the interface parameters of CZMs varies in the range while others remain unchanged to consider influence of corresponding interface parameter.

Simulation results

The cracking mode of ceramic coating is showed in Fig. 3. The interface cohesive strength and fracture toughness effects on fracture behavior of coatings are obtained as shown in Fig. 4. It can be seen that crack length decreases with increasing coating toughness as shown in Fig. 4(a) and cracking occurs later, too, i.e., damage defined by a total crack length is slower for coatings with higher toughness, which is consistent with the previous energy analysis [7]. For coating strength effect, when the strength is larger than a critical value, the crack length also decreases with increasing strength as shown in Fig. 4(b), i.e., there exists a critical value of strength for changing damage rate of coatings. It was found that the damage rate of nanostructured thin coatings was higher compared with that of corresponding microscale microstructure coatings in the previous experimental measurement [8], which implies the lower fracture toughness of nanostructured coatings based on the above simulation. The study also shows the cohesive energy decreases for nanostructured materials compared to corresponding bulk materials [9], the fracture toughness of materials should be proportional to the cohesive energy, therefore, the present simulation results can explain the experimental results.

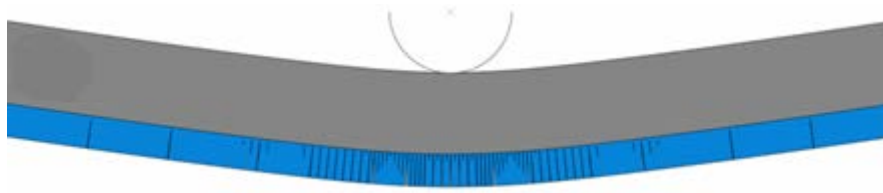


Figure 3. Cracking of coatings.

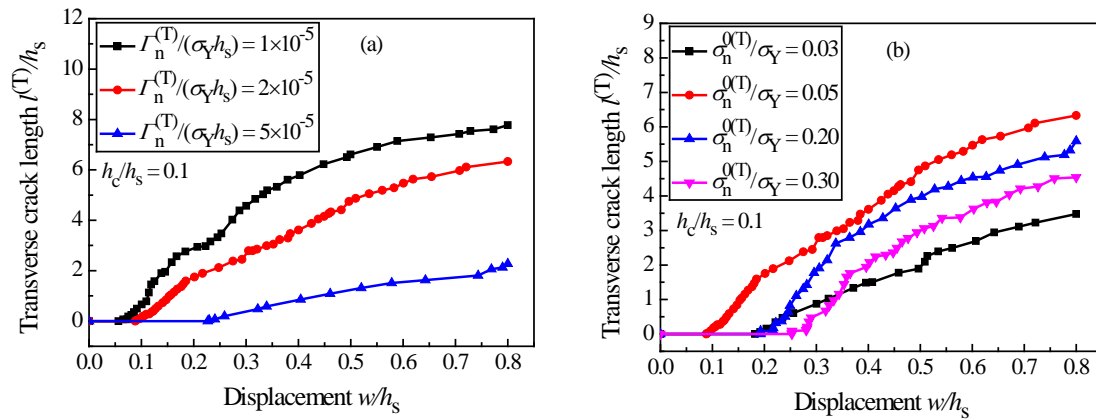


Figure 4. (a) Fracture toughness effect and (b) Strength effect of coating cracking.

Conclusions

In summary, interface effects on failure of ceramic coating/substrate systems are studied by using finite element method combining with interface cohesive zone model. The selection method of stiffness of interface cohesive elements is firstly proposed. Then cracking of coating systems under three-point bending loading is simulated. The results indicate that cracking is easier for the coating with lower interface toughness. For interface strength, there exists a critical value of changing damage rate of coatings.

References

- [1] Barenblatt, G. I. (1959) The formation of equilibrium cracks during brittle fracture. General ideas and hypotheses. Axially-symmetric cracks, *J Appl Math Mech.* **23**, 622-636.
- [2] Thouless, M. D. (1990) Crack spacing in brittle films on elastic substrates, *J. Am. Ceram. Soc.* **73**, 2144-2146.
- [3] McGuigan, A. P., Briggs, G. A. D., Burlakov, V. M., Yanaka, M., Tsukahara, Y. (2003) An elastic-plastic shear lag model for fracture of layered coatings, *Thin Solid Films* **424**, 219-223.
- [4] Li, X. N., Liang, L. H., Xie, J. J., Chen, L., Wei, Y. G. (2014) Thickness-dependent fracture characteristics of ceramic coatings bonded on the alloy substrates, *Surf. Coat. Technol.* **258**, 1039-1047.
- [5] Turon, A., Dávila, C. G., Camanho, P. P., Costa, J. (2007) An engineering solution for mesh size effects in the simulation of delamination using cohesive zone models. *Eng Fract Mech.* **74**, 1665-1682.
- [6] Blal, N., Daridon, L., Monerie, Y., Pagano, S. (2012) Artificial compliance inherent to the intrinsic cohesive zone models: criteria and application to planar meshes. *Int J Fracture.* **178**, 71-83.
- [7] Liu, H. Y., Wei, Y. G., Liang, L. H., Liu, X. H., Wang, Y. B., Ma, H. S. (2018) Damage characterization model of ceramic coating systems based on energy analysis and bending tests, *Ceram. Int.* **44**, 4807-4813.
- [8] Liang, L. H., Li, X. N., Liu, H. Y., Wang, Y. B., Wei, Y. G. (2016) Power-law characteristics of damage and failure of ceramic coating systems under three-point bending, *Surf. Coat. Technol.* **285**, 113-119.
- [9] Liang, L.H., Li, M.Z., Qin, F.Q., Wei, Y.G. (2013) Temperature effect on elastic modulus of thin films and nanocrystals, *Philos. Mag.* **93**, 574-583.

Numerical Analysis and Experimental Verification of Ti/APC-2/Kevlar Hybrid Composite Laminates due to Low-Velocity Impact

†,* Ming-Hwa R. Jen¹, Dong-Yi Cai¹, Che-Kai Chang¹, and Feng-Chi Hsu²

¹ Dept. of Mechanical and Electro-Mechanical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan 80424, ROC

² Language Center and Dept. of Nursing, Fooyin University, 151 Jinxue Rd., Kaohsiung 83102, Taiwan, ROC

*Presenting author: jmhr@mail.nsysu.edu.tw

†Corresponding author: jmhr@mail.nsysu.edu.tw

Abstract

The residual mechanical properties of Ti/APC-2/Kevlar/epoxy hybrid composite laminates after low velocity impact were investigated at room temperature. There were three types of samples tested, including three layered $[\text{Ti}/(0/90)_s/\text{Ti}]$, five layered $[\text{Ti}/(0/90)_2/\text{Ti}]_s$ and nine layered $[\text{Ti}/\text{Kevlar}/\text{Ti}/(0/90)_2/\text{Ti}]_s$. The lay-ups of APC-2 were stacked in a way of cross-ply sequence, while Ti layer was anodized with chromic acid anodic method. Ti and APC-2 were combined together to fabricate the composite laminates via hot press curing process. Kevlar layers were added to cover five-layer composite laminates to fabricate nine-layered composite laminates via vacuum assisted resin transfer molding.

The drop-weight tests were conducted with a hemispherical nosed projectile in 10 mm diameter. The impact loads were 5kg and 10kg according to the simulated results by ANSYS. The impact heights were increased until the samples were penetrated or the height reached the maximum height, 1.50 m, of our instrument. The static tensile tests were conducted to measure the composite laminate residual mechanical properties after the impact testing.

The results showed that the bottom Ti layer absorbed more internal energy than the top Ti layer, so that the cracks were found on the bottom Ti layer more often. The crack shape was the opening that resembling petal after the penetration. Also, the ultimate tensile strength reduced significantly after the impact, and it raised slightly after the samples fully penetrated. The initial longitudinal compliance increased with the impact height increasing and decreased after the samples penetrated. Comparing the experimental data with the numerical simulation results, we found the latter was more serious than the former. On the conservative side, the results of simulation can be adopted for applications in the case of no testing data available.

Keywords: Titanium, APC-2, Kevlar, low-velocity, impact, tensile test.

Introduction

Fiber metal laminates (FMLs) are hybrid composite structures constructed from thin sheets of metal alloys and plies of fiber-reinforced polymeric materials. The first FMLs, called aramid-reinforced aluminum laminates (ARALLs), were introduced in 1978 at the Faculty of Aerospace Engineering at Delft University of Technology in the Netherlands [1]. In 1990, an improved type of ARALL called glass laminate aluminum reinforced epoxy (GLARE), or ARALL with glass fibers, was successfully developed[2]. Furthermore, Lin et al. [3] developed carbon-reinforced aluminum laminates, which contain carbon fibers (CFs) rather than aramid fibers. FMLs have the advantages of metal alloys and fiber-reinforced plastic (FRP) composites. Castrodeza et al. [4] demonstrated that GLARE and ARALL possess superior fracture toughness and crack tolerance to those of their constituent alloys. Vlot [5, 6]

investigated GLARE and ARALL, finding that the impact resistance of FMLs was superior to that of the studied FRP. Gunnink [7] showed that ARALL retains excellent durability even after very long exposure to highly aggressive environments. Additionally, FMLs containing various types of metal alloys and FRPs have been developed by researchers for wider application. Khalili et al. [8] studied the mechanical properties of steel–aluminum–FRP laminates. Furthermore, Zhou et al. [9] investigated the tensile behavior of Kevlar fiber-reinforced aluminum laminates. Jen et al. developed magnesium/CF/PEEK nanocomposite laminates [10] as well as titanium (Ti)/CF/PEEK nanocomposite laminates [11] and obtained their mechanical properties at elevated temperatures.

Owing to the brittleness of thermosetting matrix resulted by cross-linking, the epoxy resins and epoxy-based fiber composites are susceptible to impact damage. Thermoplastics, having greater toughness, are considered to be potential for alleviating this problem [12]. Although damage inflicted by low-velocity impact appears quite complicated, the major failure modes include only matrix cracking, delamination, and fiber breakage [13]. The delamination mode of failure is induced by matrix cracks which occur prior to other failure modes. Thus, suppression of matrix cracking will suppress delamination. It is conceivable that the use of tougher matrices will yield composites that are more resistant to impact damage.

Except for the degree of damage, the plate specimens did not differ from beam specimens in failure modes or impact tolerance properties [14], i.e., no plate size effect. The postimpact load-carrying capability of a composite laminate is of prime concern to the design engineer. After a tool-drop type accident where no damage is visible from the surface, the structure is still expected to carry the full spectrum of loading. However, it may be wrong of overestimation. In all cases the residual strength decreased as the impact velocity increased. From the results [15] the tough matrix composites may provide excellent impact resistance properties at low-impact velocities. However, beyond a certain threshold velocity, i.e., $v \geq 25 \text{ m/s}$, the use of tough matrix materials may result in more laminate tensile and flexural strength reduction than that of brittle matrix materials. Additionally, the PEEK composites have significantly lower contact rigidity, i.e., for a given contact force the resulting indentation in the PEEK composites would be larger, yielding a larger contact area, and, therefore, a low contact pressure. A larger contact area with lower pressure will reduce the transverse shear stress concentration and thus minimize local matrix cracking.

Anodic method is a commonly used surface treatment, however, the bonding capability of polymer composites to titanium thin plates is still a problem. In order to improve the interfacial bonding capability, Ramani et al. [16] found the chromic acid anodic method was excellent. Chromic acid anodic oxidation produced an oxide layer of thickness 40~80 nm for the 5V and 10V treatments [17]. In recent years, inorganic nanoparticles filled polymer composites have attracted attention because the filler/matrix interface in these composites might constitute a great area and influence the properties of composites at rather low filler concentration [18]. Based on above-mentioned statements we fabricated Ti/APC-2 FLMs to investigate their resistance to impact loads, measure residual mechanical properties and compare the data with the results of numerical simulation by using software LS DYNA-3D.

Specimen Fabrication

The twelve-inch wide prepregs of Carbon/PEEK (Cytac Industries Inc., USA) unidirectional plies were cut and stacked into cross-ply [0/90]_s laminates. The grade 1 (H: 0.015%, O:0.18%, N:0.03%, Fe:0.2%, C:0.08%) Ti sheets, supplied by Kobe Steel Ltd (Japan), were 0.5mm

thick after rolled, heated and flattened with scratch brushing. The ultimate tensile strength of Ti is 353MPa, and modulus of elasticity 109GPa.

After a series of tests, the surface treatment by chromic acid anodic method of electro-plating was found better as demonstrated by the results of tensile tests. The anodic oxide coating film was observed uniform by Scanning Electron Microscope, and the composition of coating consisting of TiO_2 by Energy-Dispersive X-ray spectroscopy.

The APC-2 prepregs were sandwiched with the Ti alloy sheets to produce Ti/APC-2 hybrid $[\text{Ti}/(0/90)_s/\text{Ti}]$ three-layered and $[\text{Ti}/(0/90)_2/\text{Ti}]_s$ five-layered laminated composites. The hot press and modified diaphragm curing process were adopted to fabricate laminates [19]. The hybrid composite specimen was a rectangular plate of 240mm(L)×25mm(W) with thickness 1.55mm and 2.50mm. Additionally, the 4.50mm thick $[\text{Ti}/\text{Kevlar}/\text{Ti}/(0/90)_2/\text{Ti}]_s$ nine-layered laminates were covered by Kevlar layers via vacuum assisted resin transfer molding.

An MTS-810 servohydraulic computer-controlled dynamic material testing machine was used to conduct the tensile tests after the free drop impact.

Numerical Analysis

In simulation we adopted the same samples as fabricated in our lab such as $[\text{Ti}/(0/90)_s/\text{Ti}]$, $[\text{Ti}/(0/90)_2/\text{Ti}]_s$ and $[\text{Ti}/\text{Kevlar}/\text{Ti}/(0/90)_2/\text{Ti}]_s$ three hybrid composite laminates. The finite element analysis and ANSYS/LS-DYNA 3D software were used to simulate the impact process starting from the penetration of top layer step by step to the full penetration of bottom layer with the zero velocity left. Due to symmetry only one quarter of sample was considered. The 3D Solid 164 elements with eight nodes and nine degrees of freedom for each node were used to construct the model. The bullet was a hemispherical nosed projectile of 10 mm diameters. The assumed boundary conditions were that no displacement along the symmetry edges, totally constrained for other two free edges and a plastic cushion placed at the bottom of laminate to avoid rebound. The failure criterion was based on the value of principal strain. Herein, the failure values of principal strains were 0.0088 for 90° Carbon fibers, 0.189 for Ti alloy and 0.08 for Kevlar fibers, respectively.

The received results would provide valuable references for the next step of impact tests.

Experimental work

The APC-2 prepregs were sandwiched with the Ti alloy sheets to produce Ti/APC-2 hybrid $[\text{Ti}/(0/90)_s/\text{Ti}]$ three-layered and $[\text{Ti}/(0/90)_2/\text{Ti}]_s$ five-layered laminated composites. The hot press and modified diaphragm curing process were adopted to fabricate laminates [19]. The hybrid composite specimen was a rectangular plate of 240mm(L)×25mm(W) with thickness 1.55mm and 2.50mm. Additionally, the 4.50mm thick $[\text{Ti}/\text{Kevlar}/\text{Ti}/(0/90)_2/\text{Ti}]_s$ nine-layered laminates were covered by Kevlar layers via vacuum assisted resin transfer molding.

Results

The numerical simulation results of velocity, impact energy, height and failure mechanisms of $[\text{Ti}/(0/90)_s/\text{Ti}]_s$ three-layered laminates were listed in Table 1. Also, the data of three-layered laminates due to 5 kg free drop tests were listed in Table 1 for contrast. The numerical results and the data 10 kg free drop tests were listed in Table 2 together. The numerical results and

test data of $[Ti/(0/90)_2/\overline{Ti}]_s$ five layered laminates due to 5 kg free drop were tabulated in Table 3 and due to 10 kg free drop were listed in Table 4, respectively. The results and data of $[Ti/Kev/Ti(0/90)_2/\overline{Ti}]_s$ nine layered laminates due to 10 kg free drop were tabulated in Table 5.

The simulation picture of 5 kg free drop impact onto the five layered laminates at different height were shown in Fig.1. The pictures of damage impact for five layered laminates due to 5 kg free drop of height of 1.5m were shown in Fig. 2. The relationships of load and initial compliance vs. the height of free drop were plotted in Fig. 3.

Table 1. The numerical results and impact damage mechanisms of three layered samples due to 5kg free drop tests.

Velocity (m/s)	Energy (J)	Height of simulation (m)	Height of impact (m)	Damage model of simulation	Damage mechanisms of impact test
2.34	13.69	0.28	0.49	APC-2: f Ti:1 st , f	Ti:1 st , d; 2 nd , f; depression: 5.79mm
2.43	14.76	0.30	0.53	APC-2: f Ti:1 st and 2 nd , f	-
2.47	15.25	0.31	0.55	APC-2: f Ti: 1 st and 2 nd , f	Ti:1 st , d; 2 nd , f; depression: 6.40mm
2.51	15.75	0.32	0.57	ballistic limit	-
2.62	17.16	0.35	0.62	p	Ti:1 st , d; 2 nd , f; depression: 6.71mm
2.80	19.6	0.40	0.71	p	Ti:1 st , f; 2 nd , f; depression: 7.02mm
2.97	22.05	0.45	0.79	p	Ti:1 st , f; 2 nd , f; depression: 7.72mm
3.13	24.49	0.50	0.88	p	Ti:1 st , f; 2 nd , f; depression: 8.04mm
3.28	26.90	0.55	0.97	p	Two samples near penetration, one penetrated; depression: 9.30mm
3.34	27.89	0.57	1.01	p	p

Notes: 1st denotes the first layer; 2nd denotes the second layer; d: depressed; f: fractured; p: penetration; APC-2: APC-2 laminates; Ti: Ti sheet; -: not available

Table 2. The numerical results and impact damage mechanisms of three layered samples due to 10kg free drop tests.

Velocity (m/s)	Energy (J)	Height of simulation (m)	Height of impact (m)	Damage model of simulation	Damage mechanisms of impact test
1.4	9.80	0.10	0.18	APC-2: f Ti: 1 st and 2 nd , f	Ti: d; depression: 5.01 mm
1.72	14.79	0.15	0.27	APC-2: f Ti: 1 st and 2 nd , f	Ti: Ti:1 st , d; 2 nd , f; depression:6.13 mm
1.77	15.66	0.16	0.28	ballistic limit	-
1.98	19.60	0.20	0.35	p	Ti: Ti:1 st and 2 nd , f; depression:7.41 mm
2.21	24.42	0.25	0.44	p	p
2.34	27.38	0.28	0.49	p	p

Notes: 1st denotes the first layer; 2nd denotes the second layer; d: depressed; f: fractured; p: penetration; APC-2: APC-2 laminates; Ti: Ti sheet; -: not available

Table 3. The numerical results and impact damage mechanisms of five layered samples due to 5kg free drop tests.

Velocity (m/s)	Energy (J)	Height of simulation (m)	Height of impact (m)	Damage model of simulation	Damage mechanisms of impact test
2.43	14.76	0.30	0.53	APC-2: f Ti: 1 st , f	Ti: d; depression: 5.11 mm
2.62	17.16	0.35	0.62	APC-2: f Ti: 1 st , f	-
2.80	19.60	0.40	0.71	APC-2: f Ti: 1 st , f	Ti: 3 rd , f; depression: 5.78 mm
2.97	22.05	0.45	0.79	APC-2: f Ti: 1 st and 2 nd f	-
3.13	24.49	0.50	0.88	APC-2: f Ti: 1 st , 2 nd and 3 rd , f	Ti: 3 rd , f; depression: 6.22 mm
3.19	25.44	0.52	0.92	ballistic limit	Ti: 3 rd , f; depression: 6.52 mm
3.43	29.41	0.6	1.06	p	Ti: 3 rd , f; depression: 6.96 mm
3.57	31.86	0.65	1.15	p	Ti: 1 st and 3 rd , f; depression: 7.20 mm
3.71	34.41	0.70	1.24	p	Ti: 1 st and 3 rd , f; depression: 7.63 mm
3.84	36.86	0.75	1.33	p	Ti: 1 st and 3 rd , f; depression: 7.85 mm
3.96	39.20	0.80	1.41	p	Ti: 1 st and 3 rd , f; depression: 8.00 mm
4.08	41.62	0.85	1.50	p	Ti: 1 st and 3 rd , f; depression: 9.10 mm

Notes: 1st denotes the first layer; 2nd denotes the second layer; 3rd denotes the third layer; d: depressed; f: fractured; p: penetration; APC-2: APC-2 laminates; Ti: Ti sheet; -: not available

Table 4. The numerical results and impact damage mechanisms of five layered samples due to 10kg free drop tests.

Velocity (m/s)	Energy (J)	Height of simulation (m)	Height of impact (m)	Damage model of simulation	Damage mechanisms of impact test
1.40	9.80	0.10	0.18	APC-2: f Ti: 1 st , f	Ti: d; depression: 4.72 mm
1.98	19.60	0.20	0.35	APC-2: f Ti: 1 st and 2 nd , f	Ti: 1 st , d; 3 rd , f; depression: 6.09 mm
2.21	24.42	0.25	0.44	APC-2: f Ti: 1 st , f,	-
2.30	26.45	0.27	0.48	ballistic limit	-
2.34	27.38	0.28	0.49	p	-
2.43	29.52	0.30	0.53	p	Ti: 1 st , d; 3 rd , f; depression: 7.44 mm
2.80	39.20	0.40	0.71	p	Ti: 1 st , d; 2 nd and 3 rd , f; depression: 8.70 mm
3.13	48.98	0.50	0.88	p	Ti: 1 st , 2 nd and 3 rd , f; depression: 9.56 mm
3.28	53.79	0.55	0.97	p	Near penetration; depression: 10.68 mm
3.43	58.82	0.60	1.06	p	p

Notes: 1st denotes the first layer; 2nd denotes the second layer; 3rd denotes the third layer; d: depressed; f: fractured; p: penetration; APC-2: APC-2 laminates; Ti: Ti sheet; -: not available

Table 5. The numerical results and impact damage mechanisms of nine layered samples due to 10kg free drop tests.

Velocity (m/s)	Energy (J)	Height of simulation (m)	Height of impact (m)	Damage model of simulation	Damage mechanisms of impact test
2.43	29.52	0.30	0.53	APC-2: f; Kevlar: f; Ti: 1 st , f	-
2.62	34.32	0.35	0.62	APC-2: f; Kevlar: f; Ti: 1 st 5 th , f	-
2.80	39.2	0.40	0.71	APC-2: f; Kevlar: f; Ti 1 st , 2 nd and 5 th , f	Ti, APC-2, and Kevlar, d; depression: 8.13 mm
2.87	41.18	0.42	0.74	APC-2: f; Kevlar: f; Ti: 1 st , 2 nd and 5 th , f	-
2.97	44.1	0.45	0.79	APC-2: f; Kevlar: f; Ti: 1 st , 2 nd , 4 th and 5 th , f	-
3.13	48.98	0.50	0.88	APC-2: f; Kevlar: f; Ti: f	Ti: 1 st -4 th , APC-2, and Kevlar, d; Ti: 5 th , f; depression: 8.13 mm
3.22	51.84	0.53	0.93	ballistic limit	-
3.28	53.79	0.55	0.97	p	-
3.43	58.82	0.60	1.06	p	Ti: 2 nd -4 th , APC-2, and Kevlar, d; Ti: 1 st and 5 th , f; depression: 9.68 mm
3.71	68.82	0.70	1.24	p	Ti: 2 nd -4 th , APC-2, and Kevlar, d; Ti: 1 st and 5 th , f; depression: 10.61 mm
4.08	82.23	0.85	1.50	p	Ti: 2 nd -4 th , APC-2, and Kevlar, d; Ti: 1 st and 5 th , f; depression: 11.61 mm

Notes: 1st denotes the first layer; 2nd denotes the second layer; 4th denotes the fourth layer; 5th denotes the fifth layer; d: depressed; f: fractured; p: penetration; APC-2: APC-2 laminates; Ti: Ti sheet; -: not available

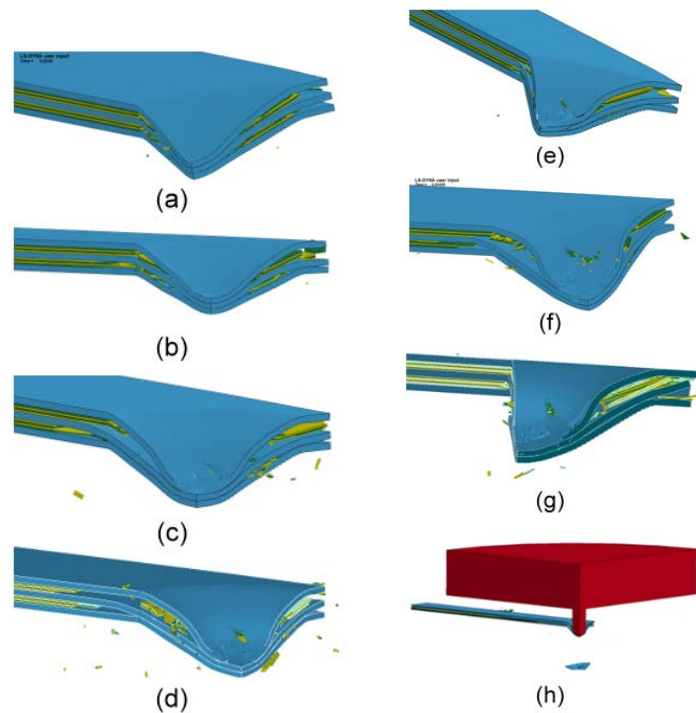


Fig. 1. Damage simulation pictures of five layered laminates due to 5kg free drop impact at heights (a) 0.2m (b) 0.25m (c) 0.3m (d) 0.35m (e) 0.4m (f) 0.45m (g) 0.5m (h) scheme of

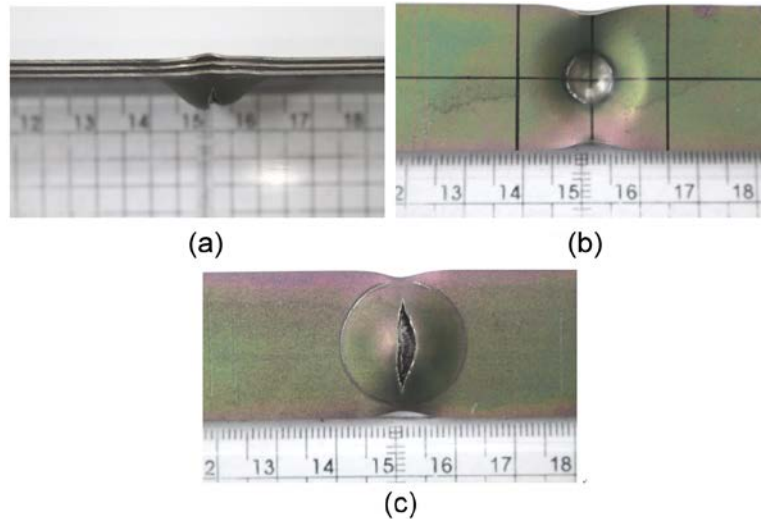


Fig. 2. The photos of impact damage on five layered laminates due to 5kg free drop at 1.5m high (a) side view (b) top view (c) bottom view.

Discussion

The numerical simulation by LS DYNA-3D software and finite element method we adopted first was to obtain the results as valuable information and references for the next step impact tests. Otherwise, the waste of many samples can not be avoided. Inversely, the test data provided an important contrast in comparison with the numerical results. We also found that the used software was acceptably feasible without the impact tests because the results were more serious damage than that of test data.

Due to the limited space of our lab the free drop impact tests were adopted alternatively. As can be seen in Tables 1-5 the real heights of free drop were much higher than those of predicted heights by simulation. It was mainly attributed to the friction of the testing system. To keep the same impact energy acting on the samples the heights adjusted and elevated were necessarily needed.

All the three types of samples after free drop impact tests were subjected to tensile tests as illustrated in Fig. 3. The general trend was that the applied loads decreased with the increasing height, however, the initial compliances of damage samples increased with the increasing height inversely.

Conclusion

Three types of Ti/APC-2 hybrid composite laminates were fabricated. The numerical simulation by using ANSYS LS DYNA-3D software and finite element method were completed to provide references for free drop impact tests. The equipments of free drop tests were set up. After all the impact tests the damage samples were due to tensile tests to obtain their residual capabilities of loads and compliances. The work can be concluded that the numerical results were more serious damage than those of test data. Thus, the adopted software was well acceptable. In the consideration of friction the height of free drop should be elevated to meet the requirement of equal impact energy in both testing and simulation.

References

- [1] Asundi, A., & Choi, A. Y. N. (1997). Fiber metal laminates: An advanced material for future aircraft. *Journal of Materials Processing Technology*, 63(1-3), 384-394.
- [2] Castrodeza, E. M., Ipinia, J. E. P., & Bastian, F. L. (2002). Experimental techniques for fracture instability toughness determination of unidirectional fibre metal laminates. *Fatigue & Fracture of Engineering Materials & Structures*, 25(11), 999-1008.
- [3] Dan-Jumbo, E., Leewood, A. R., & Sun, C. T. (1989). Impact Damage Characteristics of Bismaleimides and Thermoplastic Composite Laminates. 356-356-317.
- [4] Ditchek, B. M., Breen, K. R., Sun, T. S., & Venables, J. D. (1980). *Morphology and composition of titanium adherends prepared for adhesive bonding*. Paper presented at the 25th National SAMPE Symposium.
- [5] Dorey, G., Bishop, S. M., & Curtis, P. T. (1985). On the impact performance of carbon fibre laminates with epoxy and PEEK matrices. *Composites Science and Technology*, 23(3), 221-237.
- [6] Gunnink, J. W. (1988). Damage Tolerance and Supportability Aspects of Arall Laminare Aircraft Structures. *Composite Structures*, 10(1), 83-104.
- [7] Jen, M.-H. R., Tseng, Y. C., & Li, P. Y. (2007). Fatigue response of hybrid magnesium/carbon-fiber/PEEK nanocomposite laminates at elevated temperature. *Journal of the Japanese Society for Experimental Mechanics*, 7(special issue), s56-s60.
- [8] Jen, M.-H. R. C., C.-K.; Sung, Y.-C.; Hsu F.-C. (2010). *Fifth International Conference on Fatigue of Composites*. Paper presented at the Fifth International Conference on Fatigue of Composites, Nanjing.
- [9] Jen, M. H. R., Chang, C. K., & Sung, Y. C. (2015). Fabrication and mechanical properties of Ti/APC-2 hybrid nanocomposite laminates at elevated temperatures. *Journal of Composite Materials*, 50(15), 2035-2045.
- [10] Joshi, S. P., & Sun, C. T. (1985). Impact Induced Fracture in a Laminated Composite. *Journal of Composite Materials*, 19(1), 51-66.
- [11] Khalili, S. M. R., Mittal, R. K., & Kalibar, S. G. (2005). A study of the mechanical properties of steel/aluminium/GRP laminates. *Materials Science and Engineering a-Structural Materials Properties Microstructure and Processing*, 412(1-2), 137-140.
- [12] Lin, C. T., Kao, P. W., & Yang, F. S. (1991). Fatigue Behavior of Carbon Fiber-Reinforced Aluminum Laminates. *Composites*, 22(2), 135-141.
- [13] Ramani, K., Weidner, W. J., & Kumar, G. (1998). Silicon sputtering as a surface treatment to titanium alloy for bonding with PEKEKK. *International Journal of Adhesion and Adhesives*, 18(6), 401-412.
- [14] Sun, C. T., & Rechak, S. (1988). Effect of Adhesive Layers on Impact Damage in Composite Laminates. 97-97-27.
- [15] Vlot, A. (1993). Impact Properties of Fiber Metal Laminates. *Composites Engineering*, 3(10), 911-927.
- [16] Vlot, A. (1996). Impact loading on Fibre Metal Laminates. *International Journal of Impact Engineering*, 18(3), 291-307.
- [17] Vlot, A., & Gunnink, J. W. (2001). *Fibre metal laminates : an introduction.*: Springer.
- [18] Wu, C. L., Zhang, M. Q., Rong, M. Z., & Friedrich, K. (2005). Silica nanoparticles filled polypropylene: effects of particle surface treatment, matrix ductility and particle species on mechanical performance of the composites. *Composites Science and Technology*, 65(3-4), 635-645.
- [19] Zhou, Y. X., Wang, Y., & Mallick, P. K. (2004). An experimental study on the tensile behavior of Kevlar fiber reinforced aluminum laminates at high strain rates. *Materials Science and Engineering a-Structural Materials Properties Microstructure and Processing*, 381(1-2), 355-362.

A data-driven diagnostic tool for wind turbines under operational variability

†*Simona Bogoevska¹, Eleni Chatzi², Elena Dumova-Jovanoska¹, and Rüdiger Höffer³

¹Faculty of Civil Engineering, Ss. Cyril and Methodius University in Skopje, Republic of Macedonia.

²Department of Civil, Environmental and Geomatic Engineering, ETH Zürich, Switzerland.

³Department of Civil and Environmental Engineering, Ruhr-University Bochum, Germany.

*Presenting author: simona.bogoevska@gf.ukim.edu.mk

†Corresponding author: simona.bogoevska@gf.ukim.edu.mk

Abstract

The need for real-time condition assessment of complex systems relies on implementation of holistic Structural Health Monitoring (SHM) strategies that are capable of tracking structural behavior in a complete operational spectrum of the structure, distinguishing between true system changes and nonthreatening variations.

The proposed data-driven framework utilizes an autonomous bi-component tool able to link monitored structural response with random evolution of Environmental and Operational Parameters (EOP) affecting the monitored system. The approach combines the implementation of a Smoothness Priors Time Varying Autoregressive Moving Average (SP-TARMA) method for modeling the temporal variability in structural response, and a Polynomial Chaos Expansion (PCE) probabilistic model for modeling the propagation of response uncertainty. The computational tool is applied on long-term data, collected from an active sensing system installed for four years on a real operating WT structure located in Dortmund, Germany.

The twenty one-month tracking of the proposed PCE-SPTARMA diagnostic index, further assessed by means of statistic-based analysis, demonstrates that the proposed symbiotic treatment yields a robust model, capable of separating benign EOP fluctuations from potential pattern alterations due to actual structural damage. The obtained data-driven model verifies the future prospective of the strategy for development of an automated SHM diagnostic tool.

Keywords: Data-driven diagnostics, Operating wind turbine, Structural variability, Uncertainty propagation

Introduction

Latest technological advancements have fostered extensive application of various sensing techniques and acquisition systems on real engineering structures, thus shifting the focus towards hybrid analysis approaches (data/model) or purely data-based schemes (machine-learning black box approaches). Indeed, owing to various existing sources of uncertainty, complex behavior and variability characterizing the system and environment of actual in-service structures, data-aided assessment of operational engineering structures often remains a more accurate and computationally inexpensive alternative to approaches relying on physical law-based models.

In this context, continuous monitoring strategies facilitate the utilization of more objective and flexible tools pertinent to structural diagnostics and prognosis. Particularly for wind

turbines, as systems characterized with time-varying dynamics and alternating operating nature, the adoption of automated identification tools, capable of unprejudiced diagnosis of the structural condition, becomes most valuable.

Semi-data-driven approaches, which rely on fusion of updated Finite Element Model (FEM) of the structure and recorded vibration responses, are reported as promising long-term strategies for monitoring fatigue accumulation, as well as acceleration and strain predictions at unmeasured locations of WT structures [1]-[3]. It is worth mentioning however that Operational Modal Analysis (OMA)-based methods are limited to implementation with time invariant systems [4]-[5], i.e. parked or idling condition of the structure, mode-by-mode or case-by-case investigation [6].

EOP-born variations in structural responses, known to compromise structural performance signatures and mimic real damage states of the structure, have placed data-driven diagnostics as a highly potential approach in tackling the challenge. New emerged strategies rely on eliminating influences of environmental factors from estimated performance indicators with algorithms adopted from the area of statistics, like fitting regression models, or projection methods when influencing variables are not attainable [7]-[9]. As systems are often monitored in an unknown or healthy baseline condition, robust novelty detection strategies and manifestation of detected outliers, related to changing environmental and operational conditions, versus structural damages, are recently gaining popularity in monitored full-scale engineering structures as well [10]-[12]. Comprehensive overviews of further commonly applied statistics based concepts in SHM can be found in [11]-[13].

Instead of filtering out EOPs, an alternative in the data-driven domain lies in integrating both structural response data and influencing agents within probabilistic models [14]-[17]. The workings of a PCE-SPTARMA data-driven tool, previously tested by the authoring team on two operating WT structures [17], are herein further explored by expanding the validation periods of monitored data. Whereas the twenty one-month long implementation on a real operating WT structure confirms the robustness of the strategy, fusion of the proposed strategy with a novelty detection algorithm and probability distribution divergence measure demonstrates the high potential for further automated structural health assessment.

Conceptual model

Conceptualized as a holistic approach, the proposed strategy addresses both behavioral signatures associated to collected WT response data, i.e., (short-term) non-stationarity and long-term temporal variability. This is accomplished through tracking of measured structural responses by an algorithm capable of capturing short and long-term variability of the observed system, thus providing a link between output-only vibration response data and measured EOPs, Fig. 1.

Fluctuations that are typical for the inherent (short-term) system dynamics are modeled by means of a parametric SP-TARMA method. Identified structural performance indicators, corresponding to short-term modeled responses, are then integrated into a PCE tool. The PCE probabilistic modeling approach enables long-term monitoring of structural response variability, further associated to the randomness of measured EOPs.

With this “binocular” visualization of the problem a selected PCE-SPTARMA output feature can serve as a robust diagnostic indicator for separating benign pattern alterations from actual structural damage.

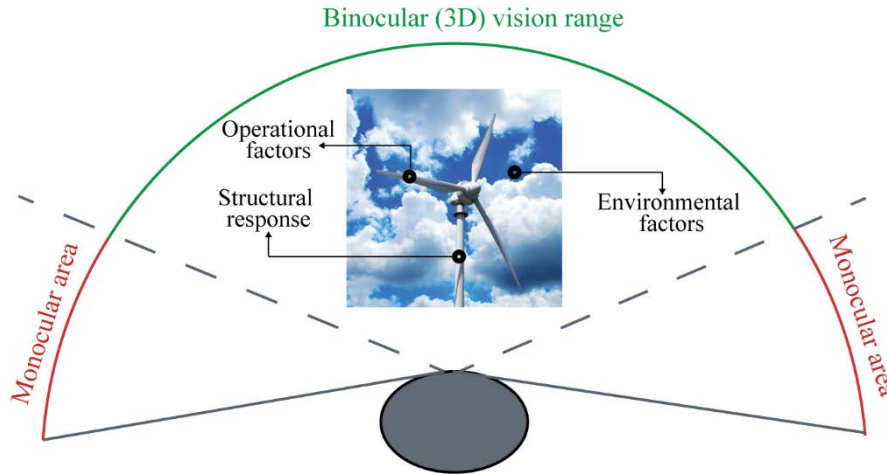


Figure 1. Conceptual model of the SHM strategy as a binocular eye vision metaphor

Theoretical framework

The proposed PCE-SPTARMA tool is a multicomponent algorithm comprising several computational methods commonly applied in a wide range of areas of research and application. The separate methods are herein summarized and presented with a concise theoretical overview. The reader is guided to further appropriate references for more detailed information on the theoretical background.

Modeling non-stationarity

The nonstationary dynamics typical for an operating WT structure can be successfully tracked via the compact parametric formulation provided by the SP-TARMA models [18]. A full SP-TARMA model is completely described by an assemblage of three equations, one representing the modeled signal (system response), and two stochastic difference equations governing the time evolution of the unknown AR and MA parameters of the model. Thus, an adequate modelling of a measured nonstationary signal is ensured by proper selection of three user-defined parameters, i.e. the AR/MA order n , the ratio of the residual variances ν , and the order of the stochastic difference equations κ [19]. Statistical approaches such as minimization of the AIC (Akaike information criterion) or the BIC (Bayesian information criterion) improve the optimal selection of these values without overfitting the modeled signal. Finally, for a selected model $M(n, \nu, \kappa)$ the SP-TARMA model parameters are obtained via the Kalman Filter scheme combined with an Extended Least Squares-like algorithm [18].

Modeling uncertainty

The PCE tool is an uncertainty quantification method, which enables the relationship between outputs (structural response performance indicators) and inputs (environmental and operational loads) to the system. A PCE model can be described by a mathematical expansion of a random system output variable on multivariate polynomial chaos basis functions [17]. Spectral representations, such as the PCE method, rely on several regularity requirements, namely finite variance of the outputs, orthonormality of the polynomial basis, and statistical independence of the input variables [20]. Hence, the polynomial chaos basis functions orthonormal with respect to the probability space of the system's random inputs have to be properly selected to ensure the necessary orthogonality relationship. Furthermore, the statistical independence of input data needs to be properly verified and possibly addressed via computational approaches capable of extracting independent (latent) variables from observed

data, such as the Independent Component Analysis (ICA) tool [21]. Then for a selected family of polynomial functions and maximum polynomial order P , the solution of the deterministic unknown parameters of a truncated PCE model are estimated via the least squares approach based on minimization of the sum of the squared residuals between true (observed) and modeled (predicted) system outputs [20].

Application case study

The described SHM strategy is implemented and tested for a 0.5MW WT erected in 1997, located in the vicinity of Dortmund, Germany, Fig. 2. A continuous measurement of acceleration response is recorded by triaxial accelerometers (PCB-3713D1FD3G MEMS sensors) mounted at five different height positions on the inner side of the WT shaft. Along with the vibration data, SCADA data are recorded with the same sampling frequency of 100 Hz. Within this paper results are presented for records corresponding to almost two complete years of continuously monitored data (January 2012 to September 2013). The last three months of year 2013, as well as scattered weeks in the previous period, are disregarded from the assessment as a result of missing temperature data from various sensor malfunctions.

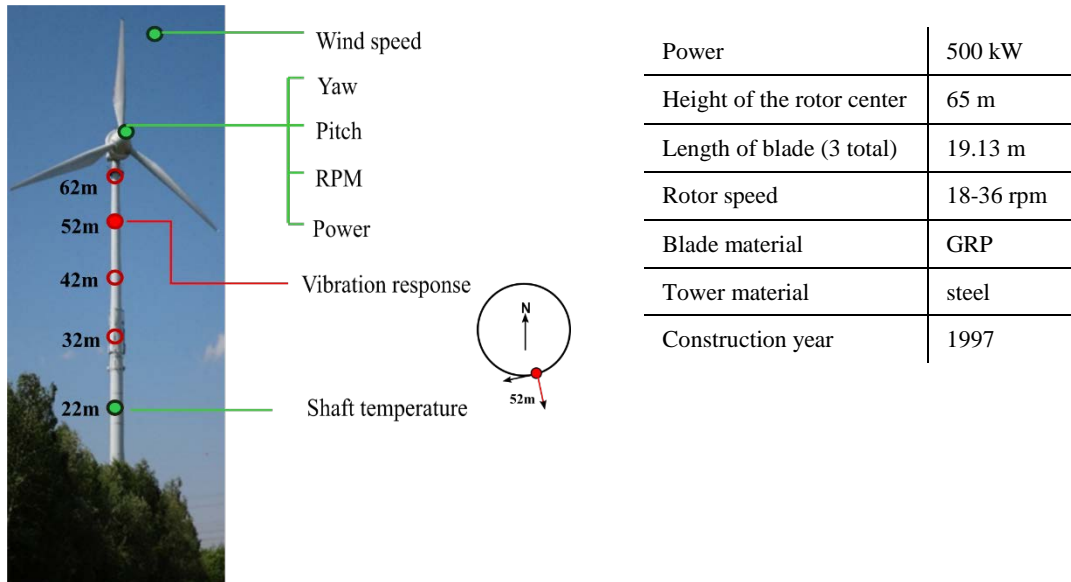


Figure 2. Schematic overview of measured data (left), WT structure characteristics (right)

As a first step, acceleration records from a selected sensor location (marked at Fig. 2) were low-pass filtered and down-sampled to 12.5 Hz, with a cut-off frequency at 6 Hz. Subsequently, 10-min long preprocessed data sets were implemented within the short-term framework. The tuning of an appropriate SP-TARMA model to actual 10-min long signals is a crucial point of the short-term modeling phase. Towards this end, plots of the AIC and BIC for model order selection are significant indicative tools that facilitate the fitting process of the user-defined parameters of the SP-TARMA model (i.e. the model order n , the smoothness constraint order κ and the residual variance ratio ν). A detailed inspection of a selection of response data sets in conjunction with their estimated statistical criterion plots revealed an optimal fitting with the parameter values equal to $n=18$, $\kappa=1$, $\nu=0.0001$. For a graphic comparison, Fig. 3 presents a fitted and an over fitted 10-min long data set signal with $\nu=0.0001$ and $\nu=0.001$, respectively. Further discussion and graphical outputs on the SP-TARMA tuning process for the actual WT structure can be found in [17].

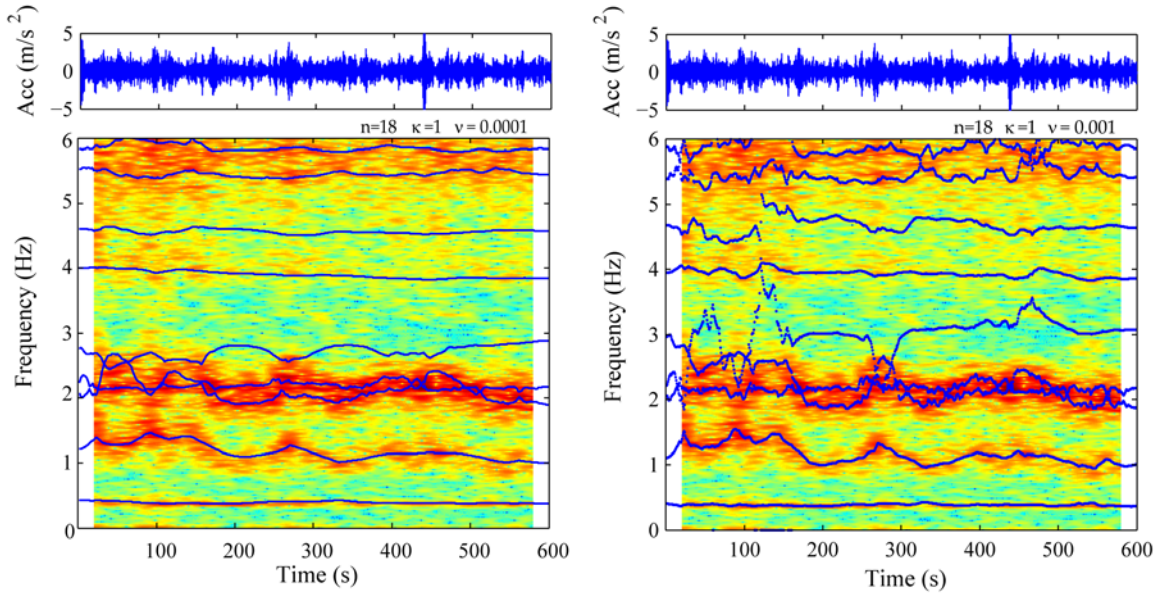


Figure 3. SP-TARMA model tuning, $v=0.0001$ (Left), $v=0.001$ (Right)

Measured data corresponding to operational and environmental parameters were organized as 10-min averages and further processed to be utilized as input variables into the long-term framework. More precisely, five SCADA parameters (wind velocity, RPM of the rotor, power production, yaw angle, and shaft temperature) were transformed to independent variables via the ICA algorithm. In order to preserve all the existing information the number of ICA latent variables was kept same as the maximum number of available EOP. For the purpose of satisfying the second PCE prerequisite, the ICA estimates are further transformed into uniformly distributed variables via use of the non-parametrically estimated cumulative distribution functions. Hence, in accordance with the uniform PDFs of the input data, the Legendre polynomials are selected as the PC functional basis. The standard deviation (std) of the SP-TARMA residuals for the 10 minute intervals, analyzed as part of the short-term framework, is selected as the PCE output parameter.

The selection of the second PCE user-defined parameter, the maximum polynomial order, is achieved via supervision of a PCE modeled output parameter Y^{PC} for a selected validation data range that clearly contains new ranges of input data. As presented in Fig. 4, the highest sensitivity to new records (marked red) of measured temperature and RPM values is linked to the maximum order $P=5$. This results in an evident discrepancy between the original output variable (Y) and the PCE modeled one (Y^{PC}). In addition, it was concluded that further increasing the maximum order does not significantly improve the accuracy of the expansion. More details regarding the PCE model generation, as well as the ICA transformation of the SCADA parameters are further elaborated in [17].

The previously described framework is herein utilized for the twenty one – month period of monitored data of the operating WT. In order to attain acceptable accuracy and alertness, as well as low computational cost, the assessment is performed for one 10-min data set per hour, resulting in total of 14064 analyzed data sets for the stated period. The standard deviation (std) of the PCE residuals is selected as a reliable Diagnostic Index (DI) able to directly demonstrate responsiveness to varying EOP, further verified with outlier analysis of validation input data sets. The proposed SHM strategy as a comprehensive three step tool is summarized in Fig. 5.

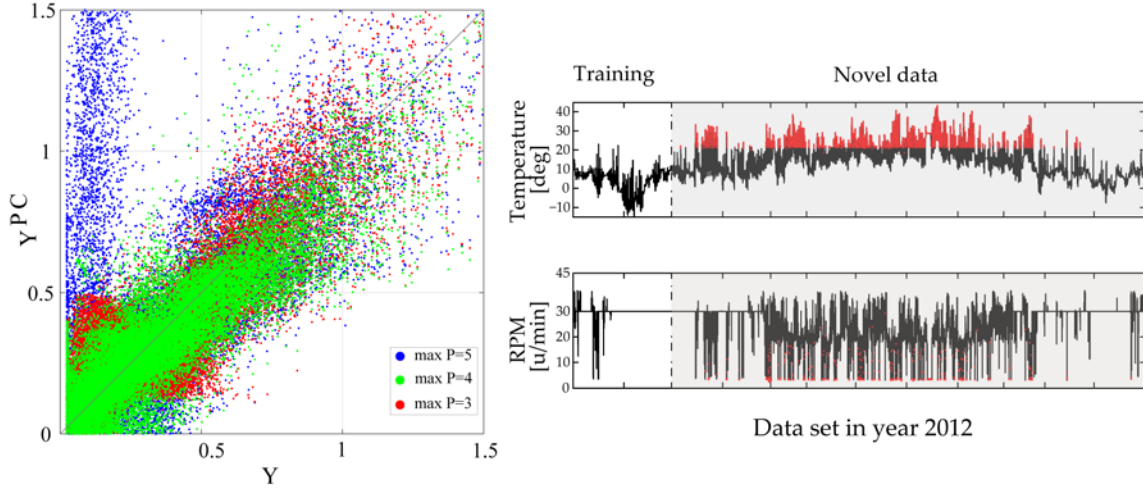


Figure 4. Tuning of PCE maximum order: modeled output (Left), input data (Right)

A preliminary testing of the sensitivity of the obtained DI is performed by means of outlier analysis on the input data time histories with the well-known Mahalanobis Distance (MD) discordancy measure. More precisely, for a p -dimensional multivariate sample x_1, \dots, x_n , the MD is defined as, [22]:

$$MD_i = \sqrt{(x_i - t_{tr})^T C_{tr}^{-1} (x_i - t_{tr})} \quad \text{for } i = 1, \dots, n \quad (1)$$

where t_{tr} is the arithmetic mean and C_{tr} is the sample covariance matrix, estimated for a certain training period of an input data set. The x_i samples from a testing set which have MD beyond a predefined value are interpreted as novelties. Hence, the definition of thresholds is vital part of the process. An adaptive method that takes into account the actual empirical chi-square distribution function of the estimated MD (instead of a fixed quantile) is herein applied [22].

The sensitivity of the index values to unfamiliar EOP fluctuations is tested for two, four and twelve-month training periods. In Fig. 6, for a two-month training period, the validation sets of the estimated DI and statistical outlier analysis (univariate MD plot) of the input data time histories illustrate that index values exceeding the ± 3 std thresholds (99.7% confidence intervals calculated for the fitted Gaussian distribution of the PCE estimation set errors) can be linked to novel data ranges of the measured influencing agents, more precisely temperature and RPM values between months March and November year 2012, as well as April and September in year 2013.

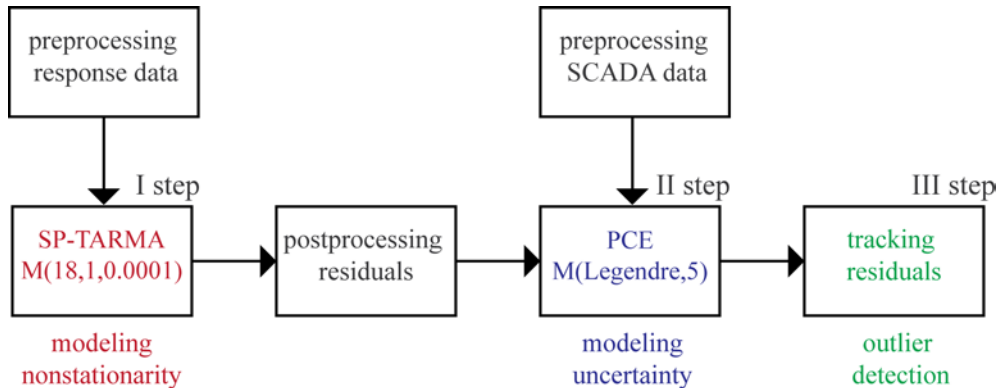


Figure 5. Schematic overview of the proposed SHM framework

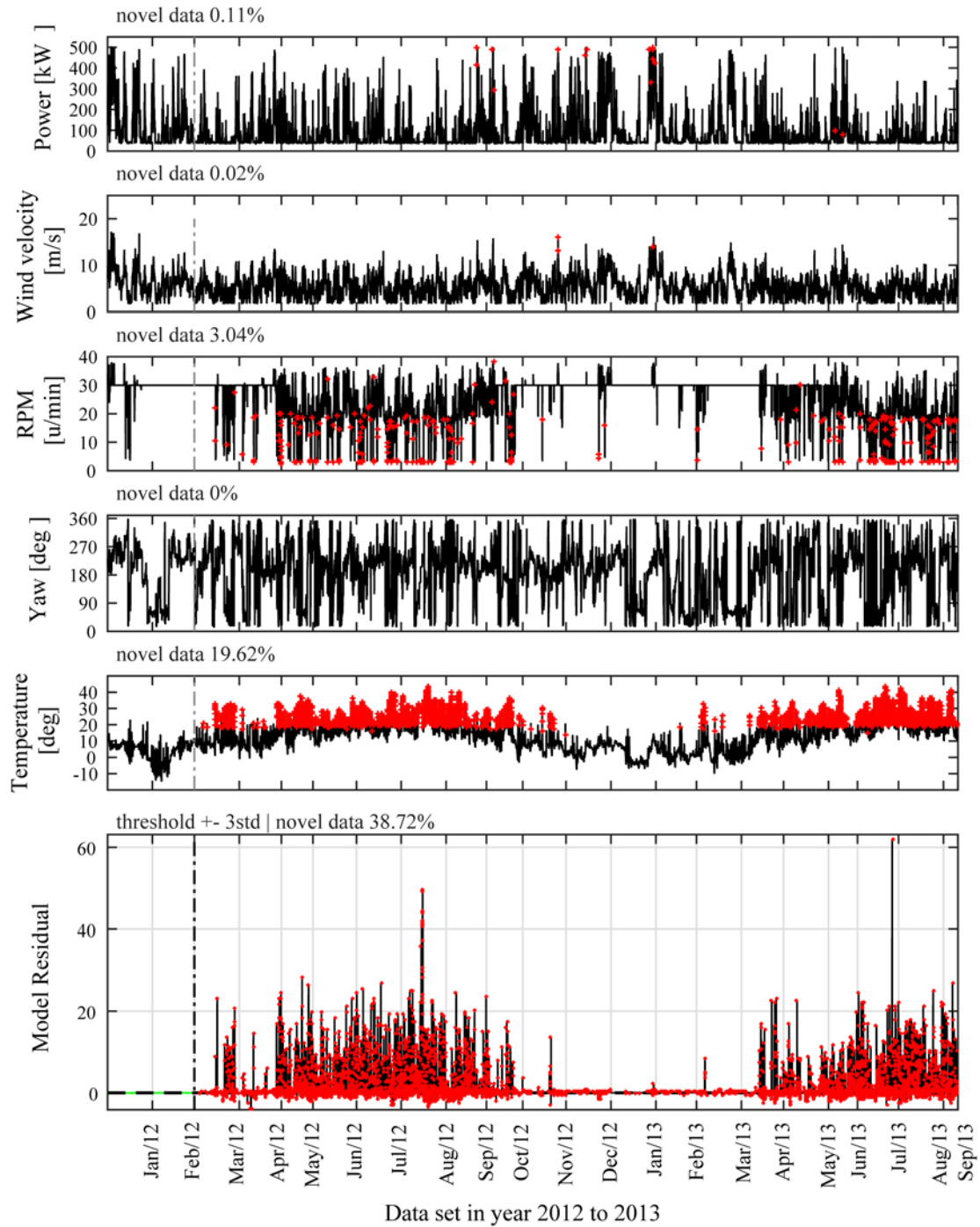


Figure 6. Two-month training set. Identified novel data (red points) within time history of 10-min mean values of measured SCADA and X- chart of the Model residual

With further increase of the training period and redefining the normal condition to include more points on fluctuating EOP (Figures 7-8), the MD outlier percentages decrease (SCADA variables with zero percent are not included in the plots) and correspondingly the DI becomes significantly reduced. In the case of the twelve-month period of training, the MD outlier percentages drop below 0.2% and the DI distribution pattern of the testing set is evidently improved, with substantially less points above the threshold values.

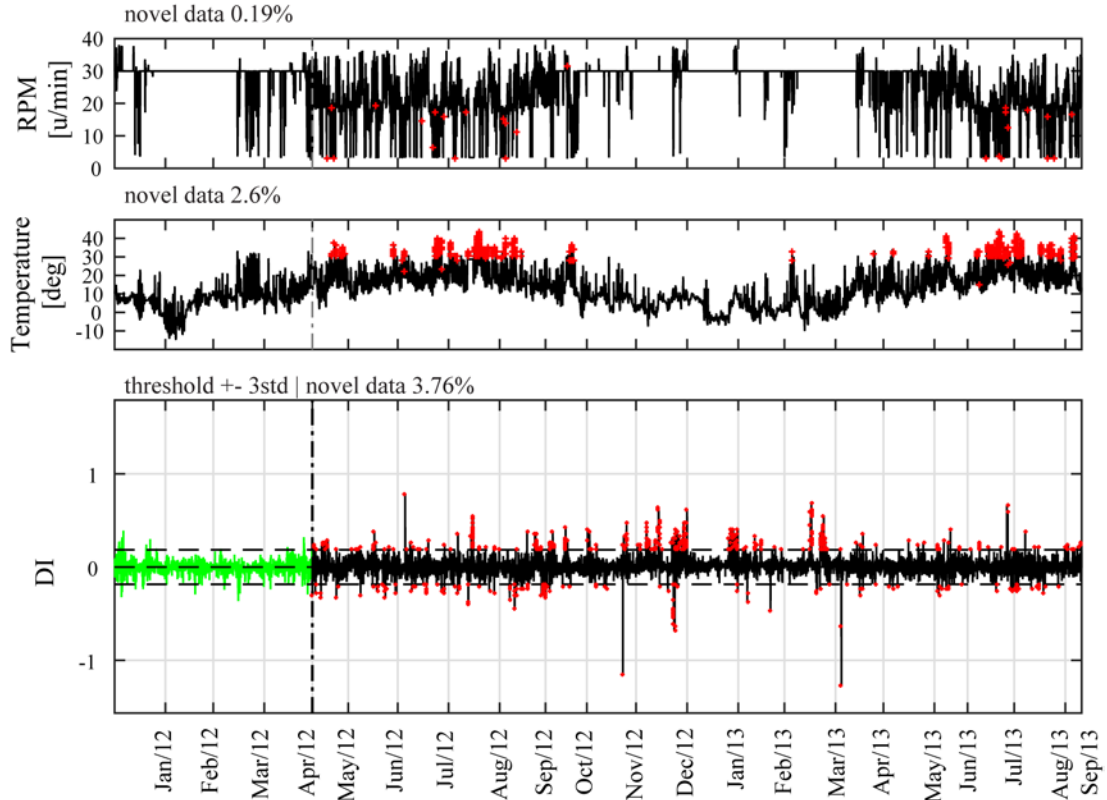


Figure 7. Four-month training set. Identified novel data (red points) within time history of 10-min mean values of measured SCADA and X- chart of the DI

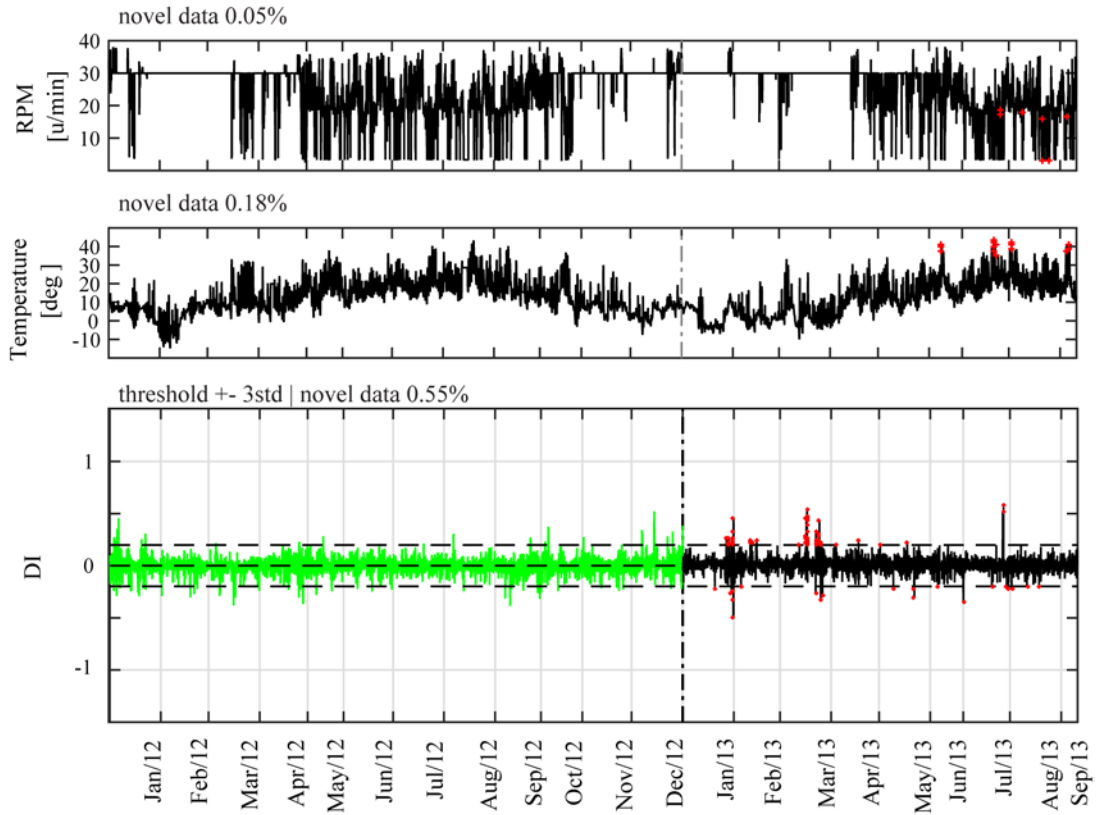


Figure 8. Twelve-month training set. Identified novel data (red points) within time history of 10-min mean values of measured SCADA and X- chart of the DI

Future work

In order to identify connections of specific patterns of structural behavior to relevant operating regimes of the WT system, future research will focus on the long-term tracking of the estimated PCE-SPTARMA diagnostic index. As preliminary presented in Fig. 9, simple statistical measures like the Kullback–Leibler divergence, applied daily on the obtained DI, demonstrate sensitivity to new data ranges of measured EOPs and agree well with the MD-based analysis of SCADA variables (Figs. 6-8).

However, in order to develop a holistic and computationally efficient tool capable of separating benign EOP fluctuations from indicator distortions due to actual structural damage or system malfunction, proper threshold tuning and pattern analysis is crucial. Towards this end, simulated damages will be introduced to the baseline training data of the monitored healthy structure. Finally, autonomous routines, based on robust outlier analysis or similar statistical measures, and an application- ready monitoring mapping for an appropriate timely reaction (model retraining or structural intervention), will be sought as a last step.

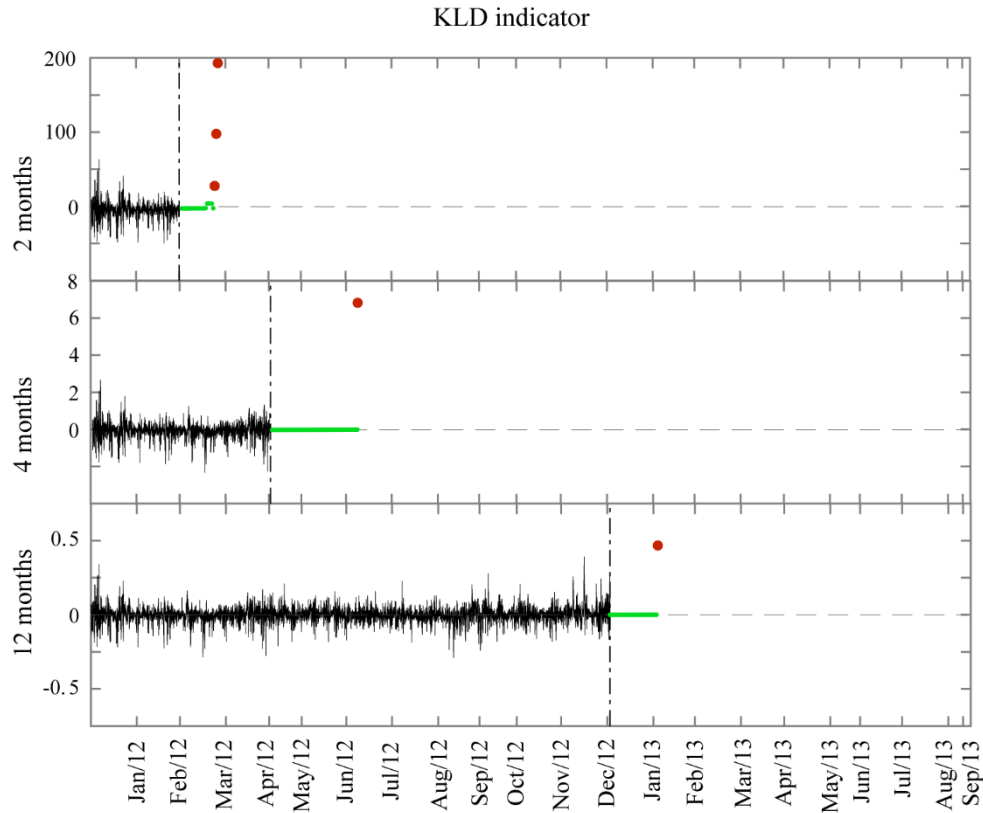


Figure 9. KL-Divergence indicator applied on the DI with cumulative assessment of one DI value per day, and tested for 2, 4, 12 months of training

Conclusions

The proposed strategy delivers a PCE-SPTARMA robust diagnostic index able to capture the non-stationary response and the long-term response variability of an actual operating WT structure for a monitoring period of twenty one months. The potential for further enhancements of the tool, towards real-time computing platform able to guide operators in the management of WT structures, is verified by outlier analysis of recorded SCADA data and preliminary utilization of statistical divergence measures on the obtained index.

References

- [1] Iliopoulos, A., Devriendt, C., Guillaume, P. and Van Hemelrijck, D., Continuous fatigue assessment of an offshore wind turbine using a limited number of vibration sensors, *Proceedings of the 7th European Workshop on Structural Health Monitoring*, Nantes, France, 2014.
- [2] Maes, K., Iliopoulos, A., Weijtjens, W., Devriendt, C. and Lombaert, G. (2016) Dynamic strain estimation for fatigue assessment of an offshore monopile wind turbine using filtering and modal expansion algorithms, *Mechanical Systems and Signal Processing* **76–77**, 592–611.
- [3] Van der Male, P. and Lourens, E., Operational vibration-based response estimation for offshore wind lattice structures, *Proceedings of the 33rd International Modal Analysis Conference*, Niezrecki, C., Ed., Orlando, Florida USA, 2015, Volume 7, 83–96.
- [4] Tcherniak, D., Chauhan, S. and Hansen, M. H., Applicability limits of operational modal analysis to operational wind turbines, *Proceedings of 28th International Modal Analysis Conference*, Jacksonville, Florida USA, 2010.
- [5] Ozbek, M., Meng, F. and Rixen, D. J. (2013) Challenges in testing and monitoring the in-operation vibration characteristics of wind turbines, *Mechanical Systems and Signal Processing* **41**, 649–666.
- [6] Limongelli, M.P., Chatzi, E., Döhler, M., Lombaert, G. and Reynders, E., Towards extraction of vibration-based damage indicators, *Proceedings of 8th European Workshop On Structural Health Monitoring*, Bilbao, Spain, 2016.
- [7] Weijtjens, W., Verbelen, T., De Sitter, G. and Devriendt, C. (2015) Foundation structural health monitoring of an offshore wind turbine- a full-scale case study, *Structural Health Monitoring* **15**, 389–402.
- [8] Oliveira, G., *Vibration-based structural health monitoring of wind turbines*, PhD Thesis, University of Porto, Portugal, 2016.
- [9] Hu, W. H., Thöns, S., Rohrmann, R. G., Said, S. and Rücker, W. (2015) Vibration-based structural health monitoring of a wind turbine system Part II: Environmental/operational effects on dynamic properties, *Engineering Structures* **89**, 273–290.
- [10] Dervilis, N., Worden, K. and Cross, E. J. (2015) On robust regression analysis as a means of exploring environmental and operational conditions for SHM data, *Journal of Sound and Vibration* **347**, 279–296.
- [11] Worden, K., Baldacchino, T., Rowson J. and Cross E., Some recent developments in SHM based on nonstationary time series analysis, *Proceedings of the IEEE* **104**, 1589–1603, 2016.
- [12] Deraemaeker, A. and Worden, K. (2018) A comparison of linear approaches to filter out environmental effects in structural health monitoring, *Mechanical Systems and Signal Processing* **105**, 1–15.
- [13] Martinez-Luengo, M., Kolios, A. and Wang, L. (2016) Structural health monitoring of offshore wind turbines: A review through the statistical pattern recognition paradigm, *Renewable and Sustainable Energy Reviews* **64**, 91–105.
- [14] Spiridonakos, M. and Chatzi, E., Polynomial chaos expansion models for SHM under environmental variability, *Proceedings of 9th International Conference on Structural Dynamics*, Porto, Portugal, 2014.
- [15] Spiridonakos, M., Chatzi, E. and Sudret, B. (2016) Polynomial chaos expansion models for the monitoring of structures under operational variability, *ASCE- ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* **2**.
- [16] Avendaño-Valencia, L.D. and Chatzi, E., (2017) Sensitivity driven robust vibration-based damage diagnosis under uncertainty through hierarchical Bayes time-series representations, *Procedia Engineering* **199**, 1852–1857.
- [17] Bogoevska, S., Spiridonakos, M., Chatzi, E., Dumova-Jovanoska, E. and Höffer, R. (2017) A Data-Driven Diagnostic Framework for Wind Turbine Structures: A Holistic Approach, *Sensors* **17**, 720.
- [18] Spiridonakos, M. D., Poulimenos, A. G. and Fassois, S. D. (2009) Output-only identification and dynamic analysis of time-varying mechanical structures under random excitation: A comparative assessment of parametric methods, *Journal of Sound and Vibration* **329**, 768–785.
- [19] Akaike, H. and Kitagawa, G., Eds. (1999) *The practice of time series analysis*, Chapter 11, Springer-Verlag, New York.
- [20] Le Maître, O. P. and Knio, O. M. (2010) *Spectral Methods for Uncertainty Quantification*, Springer, Netherlands.
- [21] Hyvärinen, A. and Oja, E. (2000) Independent component analysis: algorithms and applications, *Neural Networks* **13**, 411–430.
- [22] Filzmoser, P., Garrett, R. G. and Reimann, C., *Multivariate outlier detection in exploration geochemistry*, Technical report TS 03-5, Department of Statistics, Vienna University of Technology, Austria, 2003.

Numerical modeling of free-surface wave effects on flexural vibration of floating structures

†Shahrokh Sepehrirahnama, Eng Teo Ong, Heow Pueh Lee, and Kian-Meng Lim

Department of Mechanical Engineering, National University of Singapore, Singapore.

†Corresponding and presenting author: mpeshse@nus.edu.sg

Abstract

To investigate flexural vibration of structures in a fluid, a numerical algorithm was developed to relate the added mass and damping effects of the fluid to each mode of vibration. These are separate from the traditional added mass associated with rigid body motion, such as the translational motion along Cartesian axes. In this formulation, small-amplitude free surface waves were accounted for by using a non-singular implementation of the free-surface Green's function for a potential flow solver based on Boundary Element Method. The formulation was applied to the free and forced vibration of structures, namely a hemispherical shell and a simplified ship model, to obtain their dynamic response at various excitation frequencies. The results show the influence of added damping at lower frequencies as well as the simplicity of relating the fluid added mass to mode shapes of the structures.

Keywords: Free Surface Wave, Non-singular Boundary Element Method, Fluid-Structure Interaction

Introduction

For a structure interacting with surrounding fluid, any sudden change in the motion, for example onset of motion or change in the acceleration, results in additional resistance from the fluid in the form of a pressure load. This fluid loading can be represented by an equivalent system of mass and damper, which are called added mass and added damping, respectively [1]-[9]. Added-mass depends on the geometry of the fluid-structure interface, density of the fluid and the type of motion [1]-[3]. Added damping occurs due to the viscosity and condition of the free-surface of the fluid if it exists [4]-[9].

Added-mass and damping have been studied for different geometries of structures, either being fully or partially submerged in a fluid domain, under translational rigid-body motion [4]-[10][14][19][20]. A few studies were reported about relation between the flexural vibration of structures and added mass [21]-[24]. In most of the theoretical studies of applications such as offshore mobile structures (ships or submarines), the fluid or sea water is typically modeled as an incompressible fluid with negligible viscosity; thus, added-damping only arises from the free-surface condition [1][3][5][9][10][14].

For such a case of potential flow, the typical numerical method for calculating the fluid pressure is Boundary Element Method with proper Green's function [3][12][13][15]-[20][24]. The total pressure at the free-surface of the fluid is set to zero. If the pressure head due to the gravity is included in the total pressure, the free-surface elevation, and hence the velocity, is related to the unsteady pressure which rises from changes in the motion. For small amplitude oscillatory flow, this relationship is modeled by the Airy wave equation [25]. To use the

Boundary Element Method, modification of the Green's function is required to satisfy the free-surface wave condition. Several studies reported the surface-wave Green's function as an analytical expression which includes semi-infinite integral of the modified Bessel functions [12][13][15]-[18]. Although several analytical derivations for a limited number of simple geometries existed, numerical implementation of this Green's function is challenging due to the singularity of the Bessel function as well as the infinite bound of the integral. The surface-wave Green's function results in complex-valued pressure and velocity [11][16][17].

In this paper, the aim is to investigate the flexural vibration of shell structures interacting with fluid modelled as potential flow. A non-singular formulation is proposed for implementing the surface-wave Green's function. Then, by using the modal superposition, the fluid loading is calculated for each selected mode shape which are derived for the dry-state of structure. The added-mass and damping are represented the real and imaginary parts of the fluid loading, respectively. By including these fluid effects in the vibration equation, the flexural response of the wet-state of the structure can be calculated. The results show that the proposed numerical formulation provides an efficient way of vibration design of ship structures in sea water.

Theory and formulation

Vibration and modal superposition

Vibration of structures is governed by the following equation,

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{D}\mathbf{x} = \mathbf{g}(\mathbf{x}) + \mathbf{h}, \quad (1)$$

where \mathbf{x} is structure displacement vector, \mathbf{M} , \mathbf{C} and \mathbf{D} are the structural mass, damping and stiffness matrices, respectively; time differentiation is denoted by the dot ($\dot{}$) operator. The fluid loading on fluid-structure interface, denoted by \mathbf{g} , is a function of the displacement of the fluid-structure interface. Other external forces are denoted by \mathbf{h} . For time harmonic response, $e^{-i\omega t}$, Eq. (1) is written as follow,

$$(-\omega^2\mathbf{M} - i\omega\mathbf{C} + \mathbf{D})\mathbf{X} = \mathbf{G}(\mathbf{x}) + \mathbf{H}. \quad (2)$$

where ω is the circular frequency; \mathbf{X} , \mathbf{G} and \mathbf{H} are the corresponding complex amplitudes. To determine fluid loading \mathbf{G} as a function of displacement, the modal superposition technique is applied. First, structure displacement is written as

$$\mathbf{X} = \sum_{j=0}^N \boldsymbol{\psi}_j Q_j = [\boldsymbol{\psi}]\{\mathbf{Q}\}, \quad (3)$$

where $\boldsymbol{\psi}_j$ is the j th mode shape of the dry structure, which is called j th dry mode-shape hereinafter, and $\boldsymbol{\psi}_j$ is associated with the j th dry frequency $f_j^{(dry)}$. The participation factor of the j th dry mode in the response is denoted by Q_j . In the literature, $[\boldsymbol{\psi}]$ and $\{\mathbf{Q}\}$ are also known as modal matrix and natural coordinates, respectively. Since the dry mode-shapes are linearly independent and orthogonal, a partial fluid loading \mathbf{F}_i is calculated for each of them by using the Boundary Element Method, as described in the next subsection. Again by invoking the concept of modal superposition, the total fluid loading is calculated as follows,

$$\mathbf{G}(\mathbf{x}) = \sum_{j=0}^N \mathbf{F}_j Q_j = [\mathbf{F}]\{\mathbf{Q}\}. \quad (4)$$

The fluid loading matrix $[\mathbf{F}]$ can be replaced by the equivalent system of mass and stiffness as follows,

$$[\boldsymbol{\psi}]^T [\mathbf{F}] = \omega^2 \mathbf{M}_a + i\omega \mathbf{C}_a \quad (5)$$

where \mathbf{M}_a and \mathbf{C}_a are the modal added mass and damping, respectively, and $[\boldsymbol{\psi}]^T$ is the transpose of the modal matrix. The sizes of the added mass and damping matrices are the same as the number of selected dry mode-shapes for modal analysis. The advantage of calculating modal added mass and damping is that the magnitude of the fluid loading and its non-uniform distribution over the interface, for the case of flexural vibration, is reported by a single value which appears on the main diagonal of the matrices. The off-diagonal terms indicate the interaction between different dry mode-shapes in terms of the fluid loading.

By pre- and post-multiplying Eq. (2) with $[\boldsymbol{\psi}]$ and using Eq. (3), (4) and (5), one can write

$$(-\omega^2(\hat{\mathbf{M}}+\mathbf{M}_a) - i\omega(\hat{\mathbf{C}}+\mathbf{C}_a) + \hat{\mathbf{D}})\{\mathbf{Q}\} = \hat{\mathbf{H}} \quad (6)$$

where $\hat{\mathbf{M}}$, $\hat{\mathbf{C}}$, $\hat{\mathbf{D}}$ and $\hat{\mathbf{H}}$ are the modal structural mass, damping, stiffness and loading, respectively and $\hat{\mathbf{M}} = [\boldsymbol{\psi}]^T \mathbf{M} [\boldsymbol{\psi}]$. For free vibration analysis, damping terms and external load are ignored and Eq. (6) becomes

$$(-\omega^2(\hat{\mathbf{M}}+\mathbf{M}_a) + \hat{\mathbf{D}})\{\mathbf{Q}\} = \mathbf{0}. \quad (7)$$

Eq. (7) is an eigenvalue problem from which the frequencies and mode shapes of the immersed structure, which are called wet frequencies and mode-shapes, can be computed. It is noted that the wet mode-shapes are actually calculated by means of the modal superposition and the $\{\mathbf{Q}\}$ obtained from Eq. (7).

Flow simulation with free-surface wave

For a linear inviscid and incompressible flow, Navier-Stokes equation is reduced to a potential flow equation, as follows,

$$\Delta\varphi = 0 \quad (8)$$

$$p = -\rho_f \dot{\varphi} + \rho_f g z \quad (9)$$

$$\nabla\varphi = \mathbf{v} = \dot{\mathbf{x}} \quad (10)$$

where φ is the velocity potential, p is the total pressure, ρ_f is the density of the fluid, g is the gravity acceleration and z is the position along the vertical. Eq. (9) is also known as the linearized Bernoulli's equation and indicates the contribution of the unsteady motion (first term) and the gravity potential (second term) in the total fluid pressure. On the free surface of the fluid (where $p = 0$) for an oscillatory flow, one can rewrite Eq. (9) by using Eq. (10), as follows,

$$-i\omega\varphi + \frac{g}{-i\omega} \frac{\partial\varphi}{\partial n} = 0|_{@free-surface}, \quad (11)$$

or

$$\frac{\partial\varphi}{\partial n} = \frac{\omega^2}{g} \varphi|_{@free-surface}, \quad (12)$$

where n is the unit normal to the free surface which is in the z -direction in this derivation. Eq. (12) is known as the Airy wave equation that governs small amplitude wave on the free surface of a fluid due to gravity effects on an oscillatory flow. It shows that the velocity potential and its normal derivative are related on the free surface. The wavelength of the surface wave can be obtained as follows,

$$\lambda_f = \frac{2\pi}{\alpha} = \frac{2\pi}{\omega^2}, \quad (13)$$

where α is the wavenumber. Eq. (8) can be represented as a Boundary Integral Equation, as follow,

$$c_p \varphi(\mathbf{x}_p) + \int_S \frac{\partial G(|\mathbf{x}_p - \mathbf{x}_q|)}{\partial n(x_q)} \varphi(\mathbf{x}_q) dS(\mathbf{x}_q) = \int_S G(|\mathbf{x}_p - \mathbf{x}_q|) \frac{\partial \varphi}{\partial n}(\mathbf{x}_q) dS(\mathbf{x}_q) \quad (14)$$

where $G(|\mathbf{x}_p - \mathbf{x}_q|)$ is the Green's function, \mathbf{x}_p and \mathbf{x}_q are the position vectors of point p and q , respectively, c_p is the solid-angle constant, which is 0.5 if point p is on the boundary Γ and 1 if located in the fluid domain. The area of the surface element at point q is denoted by $dS(\mathbf{x}_q)$. Eq. (13) can be solved by using boundary elements on the fluid-structure interface as long as conditions on other boundaries are satisfied by an appropriate Green's function. A modified Green's function was derived to impose the wave condition without discretizing the infinite free-surface [15]-[17]. The analytical expression for this Green's function, which is called the surface-wave Green's function and denoted by G_w , is

$$G_w = \frac{1}{4\pi r} + \frac{1}{4\pi \bar{r}} + \frac{i\alpha}{2} e^{\alpha(z_p + z_q)} H_0^{(1)}(\alpha R) + I_\infty, \quad (15)$$

where $R = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2}$, $r = \sqrt{R^2 + (z_p - z_q)^2}$, $\bar{r} = \sqrt{R^2 + (z_p + z_q)^2}$,

$$I_\infty = -\frac{1}{\pi^2} \int_0^\infty \frac{\alpha K_0(\eta R)}{\alpha^2 + \eta^2} (\alpha \cos \eta(z_p + z_q) - \eta \sin \eta(z_p + z_q)) d\eta. \quad (16)$$

Here, $H_0^{(1)}$ and K_0 are the zeroth order Hankel function of the first kind and modified Bessel function of the second kind, respectively [17]. Since K_0 is a fast decaying function, the semi-infinite integral in Eq. (16) is computed by using Gauss quadrature method with 61 Gauss points which gives a relative error of less than 0.01%. Also, by using this numerical technique, the singularity at $\eta = 0$ is also avoided. Since both $H_0^{(1)}$ and K_0 are singular at zero, the choices of \mathbf{x}_p and \mathbf{x}_q that gives $R = 0$ should be treated separately. For this purpose, a non-singular series expansion is proposed based on Newman's derivation to calculate G_w for the case of $R = 0$ [11]. It is noted that the special case of $R = 0$ occurs when the two points are located along a line parallel to the z axis.

From Newman's series derivation for small distances between the two points p and q [11], an alternative formulation was derived to take care of the $R = 0$ singularity of the Bessel functions, as follows,

$$G_w = \frac{1}{4\pi r} + \frac{1}{4\pi \bar{r}} + \frac{i\alpha}{2} e^{\alpha(z_p + z_q)} J_0(\alpha R) + \mathcal{T}_\infty \quad \text{for } R < 10^{-3}, \quad (17)$$

$$\mathcal{T}_\infty = -\frac{\alpha}{2\pi} J_0(\alpha R) e^{\alpha(z_p + z_q)} \text{Ei}(-\alpha(z_p + z_q)) + \mathcal{S}_\infty \quad (18)$$

$$\mathcal{S}_\infty = \sum_{n=1}^{\infty} \frac{(-1)^n}{\Gamma(n)^2} \left(\frac{R}{2(z_p + z_q)} \right)^{2n} \sum_{m=0}^{2n-1} \Gamma(2n - m - 1) (-1)^m (z_p + z_q)^m, \quad (19)$$

where J_0 is the order zero regular Bessel function, $\Gamma(n)$ is the Factorial function, Ei is the standard Exponential Integral function, written as follows,

$$\text{Ei}(z) = -PV \left(\int_{-z}^{\infty} \frac{e^{-t}}{t} dt \right) \quad (20)$$

with the principal value being denoted by PV . Eq. (17) to (20) provides a non-singular Green's function for $R < 10^{-3}$ and are easy to implement since the Bessel, Factorial and Exponential Integral functions have standard implementation based on reference [26], and are accessible from any standard math library such as GSL for C/C++ programs. It is noted that

there is an infinite sum in Eq. (19); however, using more than 5 terms for $R < 10^{-3}$ changes the final value of the sum by less than 0.1%. Hence, the infinite sum is truncated to only five terms for our computer implementation of the non-singular surface-wave Green's function.

Since this Green's function is a complex-valued function, the calculated pressure amplitudes on the fluid-structure interface are also complex values. The real and imaginary parts of the fluid pressure correspond to the added mass and damping, respectively. By multiplying the pressure with surface area, the fluid loading is obtained at the centroid of each element. The force is distributed to the nodes equally due to the linear shape function. Partial fluid loading for the dry mode shapes are assembled in the same order as the mode shapes to obtain the fluid loading matrix $[F]$.

Proposed numerical formulation

To summarize, the steps in the proposed numerical formulation are listed here, as follows.

- Step 1: Extract N dry mode-shapes by using Finite Element analysis of the structure
- Step 2: For each dry mode-shape, calculate the fluid pressure loading by using Boundary Element Method (Eq. (9), (14), (15) and (17))
- Step 3: Extract the equivalent added mass and damping by using Eq. (5)
- Step 4: Solve the free (or forced) vibration by using Eq. (7) (or (6))
- Step 5: Calculate the structure response by using modal superposition, Eq. (3)

The size of the matrices for in Equations (7) and (6) is the same as the number of selected mode shapes which is much smaller than the size of the discretized model. By using the modified Green's functions to include the free-surface wave, the size of the fluid problem which is solved by Boundary Element Method is also kept to a minimal size. In term of computation time, the fluid solver is the most expensive part of this formulation and it is performed for all selected dry mode-shapes.

Simulation Results

Two case studies, a hemispherical shell and a simplified ship model with internal partitions, were considered; numerical results of structural vibration of these two cases based the proposed numerical algorithm are presented. The normalized frequency, which is reported in this section, is defined as

$$f_j^* = f_j a \sqrt{E/\rho} \quad (2)$$

where f_j is the natural frequency of the j th mode, a is a characteristic length of the given geometry, and E and ρ are the mass density and Young modulus of the solid material, respectively. Variables with the unit of length are also normalized by the characteristic length a .

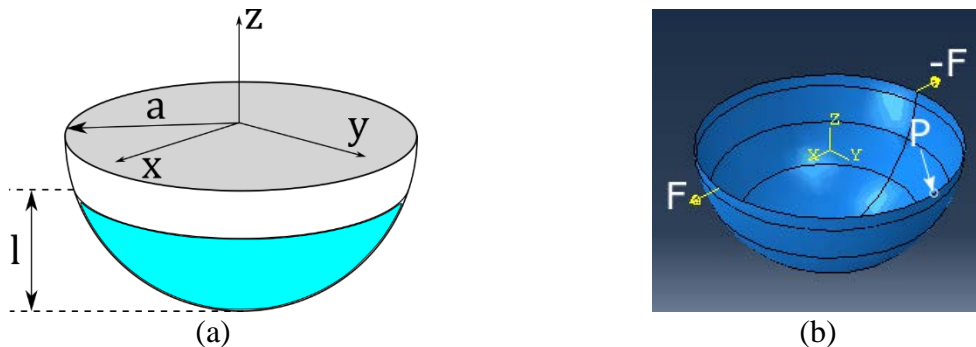


Figure 1: (a) illustration of a hemispherical shell of radius a , immersed partially by l and

(b) the location of external forces acting on the rim for forced vibration study

Fig. 1 shows the hemispherical shell (case 1) which is immersed by $l/a = 0.9$, where the immersion depth is denoted by l . The external forces applied to the structure for forced vibration study are shown in Fig. 1b as F . The forces are acting on two opposite points on the rim in the x direction.

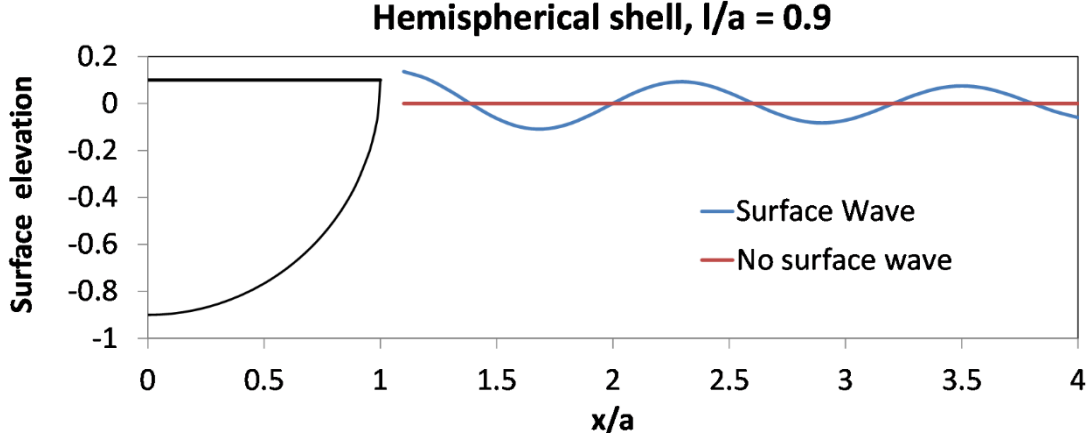


Figure 2: Normalized amplitude of surface elevation along a line segment on the free surface of the fluid domain for two cases, with and without free surface wave. Half of the shell is illustrated at the left side to demonstrate the relative location of the line segment.

For each dry mode-shape of the hemispherical shell, the fluid flow was solved for two scenarios: with and without accounting for the free-surface wave. To verify the implementation of the special Green's function G_w , the normalized amplitude of the surface elevation was plotted for a line segment on the free surface, as shown in Fig. 2, for the first dry mode shape with two nodal lines associated with $f_1^* = 2.20 \times 10^{-4}$. The free-surface wave length was obtained from Eq. (13) to be $1.2a$. The same wavelength can be observed for the surface undulations shown in Fig. 2. This implies that the surface condition is satisfied automatically by using the special Green's function. When the surface wave is neglected, the surface elevation becomes zero, which is presented in Fig. 2 by a red solid line.

After calculating the fluid pressure acting on the hemispherical shell, the added mass and damping were calculated and subsequently incorporated into the vibration equation to obtain the wet frequencies and their associated participation factors from Eq. (7). The wet mode-shapes were then calculated from Eq. (3), by multiplying the participation factors with the modal matrix of the dry structure. Fig. 3 shows the first four dry and wet mode shapes of the shell in ascending sequence of their frequencies. It is noted that the axisymmetric shape of the shell gives the repeated frequencies, for which the corresponding mode shapes have the same number of nodal lines. The same feature was observed for the wet state of the shell, as shown in Fig. 3b and 3d for the first repeated frequencies, and Fig. 3f and 3h for the next pair. Dry mode-shapes were normalized by the unitary normalization technique. The displacement profiles of the wet mode-shapes were derived from the dry mode-shapes and the participation factors. It is noted that the wet mode-shapes are similar to the dry ones, implying that the chosen dry mode-shapes give an appropriate set of basis function to construct the arbitrary response of the shell to any excitation.

As explained in the formulation section, the additional fluid resistance is presented by the modal added mass and damping. For free vibration analysis, only the added mass is accounted for to obtain the wet natural frequencies. The modal added mass matrix is reported in Table 1.

For each mode, the values are normalized by the modal structural mass of that mode. The reported values are rounded up to two decimal places hereinafter.

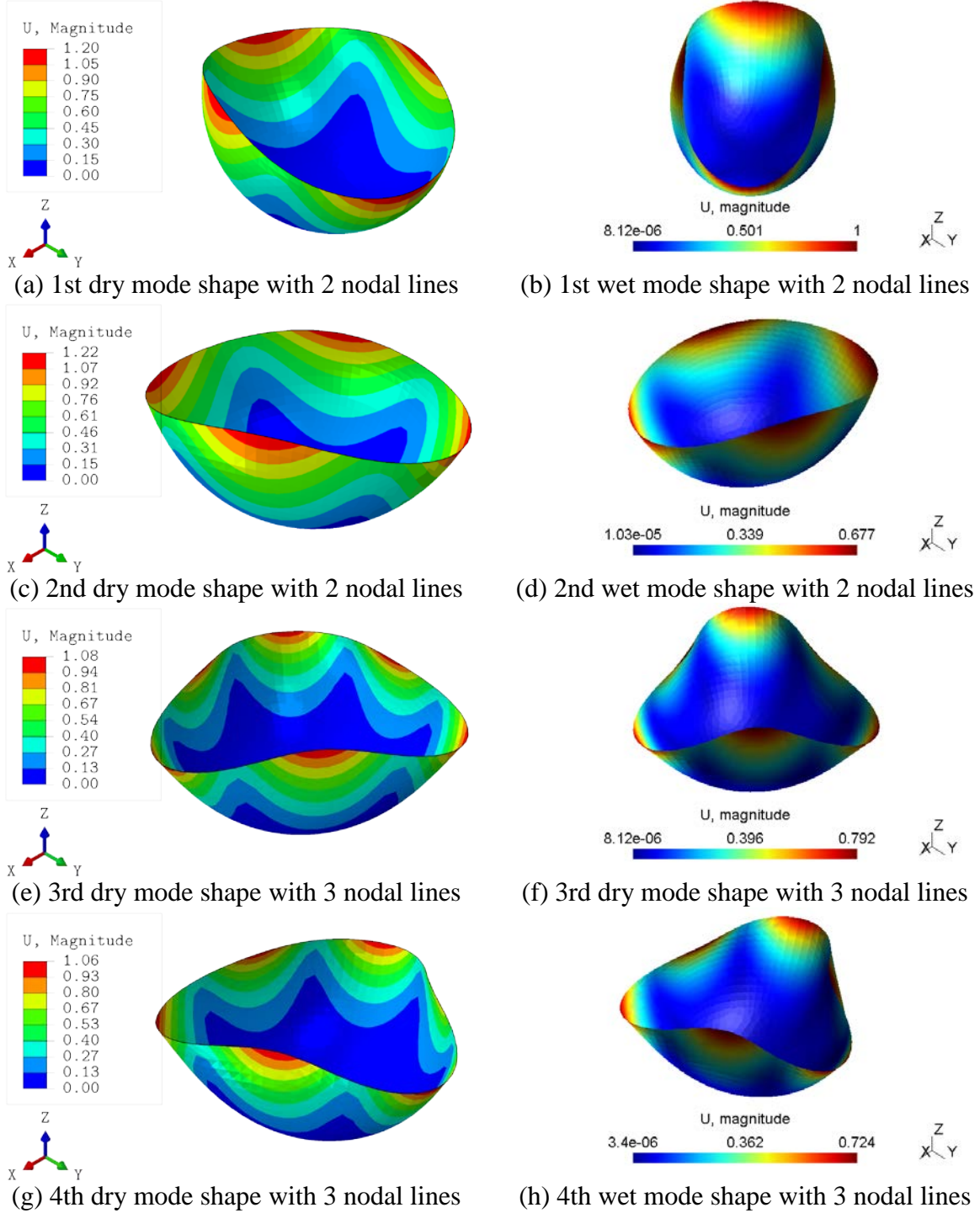


Figure 3: Panels (a), (c), (e) and (g) are the first four dry mode shapes of a hemispherical shell and panels (b), (d), (f) and (h) are their wet counterparts, respectively, for the immersion depth of $\hat{l} = 0.9$

It is noted that all the diagonal elements are larger than one, indicating that the modal fluid added-mass is larger than structural modal mass. The added mass matrix is not symmetric, due to the formulation of Boundary Element Method used for estimating the fluid loading. Despite the structural added-mass matrix being symmetric, the total mass matrix is non-symmetric which results in non-orthogonal eigenvectors. It was also observed that the off-diagonal elements are smaller than the diagonal ones by at least two orders of magnitude. By solving Eq. (7), wet frequencies and participation factors were computed by using an in-house Eigenvalue solver. The participation factors for the four mode shapes are reported in Table 2 to demonstrate the contribution of each dry mode shape in constructing the wet modes of vibration.

Table 1: Normalized modal added mass matrix for the first four mode shapes of the hemispherical shell

$[m_a]_{jk}/[m]_{jj}$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
$j = 1$	7.95	-0.04	0.00	-0.01
$j = 2$	0.03	7.95	-0.02	0.00
$j = 3$	0.00	-0.02	6.26	-0.01
$j = 4$	-0.02	0.00	0.01	6.25

Table 2: Eigen vector of participation factors with their associated wet frequency

	$\{Q\}_1$	$\{Q\}_2$	$\{Q\}_3$	$\{Q\}_4$
$\left\{ \begin{array}{l} \text{dry mode 1} \\ \text{dry mode 2} \\ \text{dry mode 3} \\ \text{dry mode 4} \end{array} \right\}$	$\left\{ \begin{array}{l} -0.73 \\ 0.68 \\ 0.00 \\ 0.00 \end{array} \right\}$	$\left\{ \begin{array}{l} -0.69 \\ -0.72 \\ 0.00 \\ 0.00 \end{array} \right\}$	$\left\{ \begin{array}{l} 0.00 \\ 0.00 \\ 0.61 \\ 0.79 \end{array} \right\}$	$\left\{ \begin{array}{l} 0.00 \\ 0.00 \\ 0.78 \\ -0.62 \end{array} \right\}$

The vectors of participation factor in Table 2 are linearly independent. The first two dry mode shapes which are associated to the first repeated frequency have dominant contributions in the first two wet mode shapes. This relationship can be observed in Fig. 3a, 3c and 3b, 3d, respectively. Similarly, wet modes 3 and 4 are constructed mainly from the third and fourth dry modes.

Table 3: Ratio between the wet and dry frequencies of the hemispherical shell

	$f_j^{*(wet)}/f_j^{*(dry)}$	$\epsilon_j = 100 \times (f_j^{*(dry)} - f_j^{*(wet)})/f_j^{*(dry)}$
$j = 1$	0.33	66.65%
$j = 2$	0.33	66.50%
$j = 3$	0.37	62.83%
$j = 4$	0.37	62.92%

The ratio of dry and wet frequencies and the relative downshift ϵ_j in natural frequency due to fluid added-mass effect are reported in Table 3. For all the modes, the wet frequencies are lower than their dry counterparts, demonstrating the additional resistance from fluid motion induced by the deflection of the interface. The considerable reduction in the frequencies implies the significant impact of fluid loading on the vibration response of such shell structures.

Table 4: Normalized modal added damping matrix for the first four mode shapes of the hemispherical shell

$[C_a]_{jk}/([m]_{jj} \times f_j^{(wet)})$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
--	---------	---------	---------	---------

$j = 1$	171.93	1.90	-0.03	0.01
$j = 2$	-2.19	171.12	-0.12	-0.02
$j = 3$	-0.02	0.10	4.01	0.08
$j = 4$	0.11	0.03	-0.07	4.04

To study the forced vibration of the shell for the given external loads shown in Fig. 1b, the modal added damping, which represents the dissipative effects of the free-surface wave, were included. Table 4 shows the modal added damping which are normalized by the critical damping $[C_{cr}]_{jj} = [m]_{jj} \times f_j^{(wet)}$ with $f_j^{(wet)}$ being the wet frequency of the j th mode. It can be seen that for all the four modes, the diagonal elements of the matrix which represent the added-damping factors are all greater than one. This implies that the shell is overdamped due to the dissipative effect of the free-surface waves. The off-diagonal elements are smaller than the diagonal damping factors by at least two orders of magnitude. It is noted that the modal damping tends to be larger at lower frequencies, and the added damping is the same for mode shapes associated with repeated frequencies.

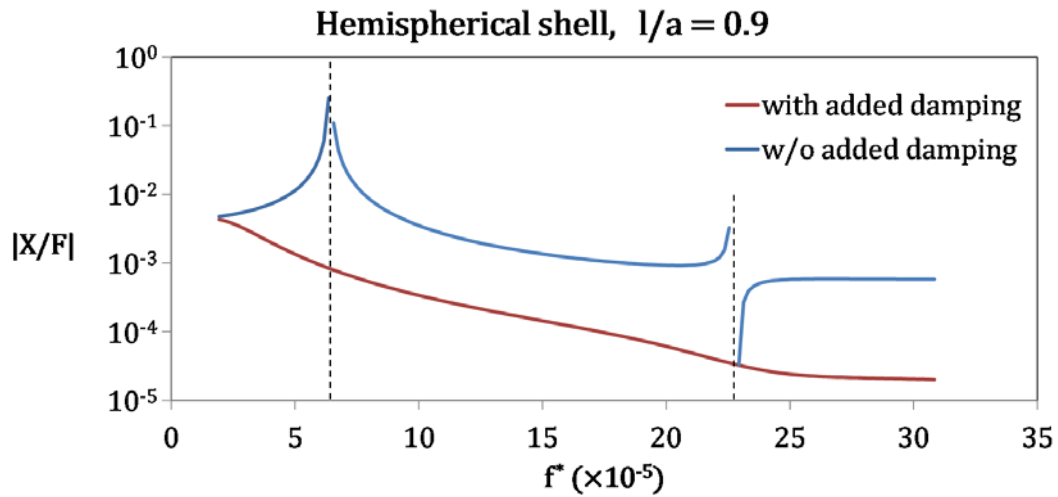
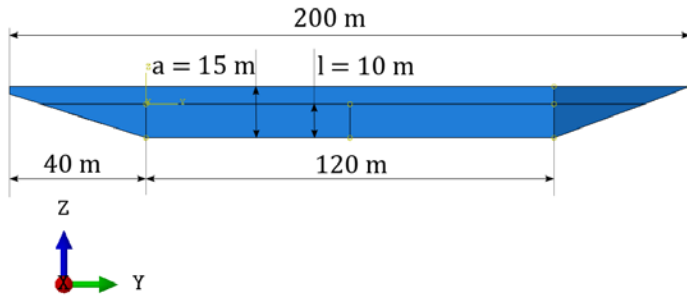


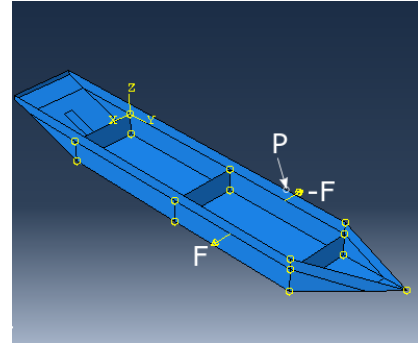
Figure 4: Forced vibration response of the hemispherical shell subjected to force F at point P , as shown in Fig. 1b, with and without including the added damping from the free-surface wave effect.

The displacement response at an observation point P due to the given load over a frequency range that contains the two natural frequencies is plotted in Fig. 4. The results are shown for two cases, with and without free-surface wave effect, to investigate the added-damping. When the surface wave was neglected, strong resonance can be observed when the excitation frequency approaches the calculated natural wet frequencies. By including the free-surface wave and hence its added-damping effect, the vibration response shows that the system is in the overdamped state as the free-surface wave carried energy away from the shell. This implies that resonance will not occur in this frequency range as long as the shell is partially submerged in a fluid.

For the next case study, a simplified ship model was developed for analysis, as shown in Fig. 5. The model dimensions are 200m (length) \times 30m (width) \times 15m (height). Three plate partitions with 10 m height were placed inside hull at 40, 100 and 160 m. The ship draft is considered to be 10 m, which is illustrated by a horizontal line on the ship hull. Two equal and opposite forces shown in Fig. 5b are exerted on the sides of ship hull for forced vibration analysis.



(a)



(b)

Figure 5: (a) schematic illustration of a simplified ship model, immersed partially by l and (b) the location of external forces acting on the side walls for forced vibration study

The first four dry frequencies and mode-shapes of the ship model were obtained using finite element analysis. Similar to the previous case, the fluid loading was calculated for each dry mode-shape by using the Boundary Element Method, with and without including the free-surface wave effects. After deriving the added mass matrix, the free vibration of the wet ship model was solved to obtain the wet mode-shapes and natural frequencies. It is noted the internal plates in the model were not considered in fluid flow simulation, since only the fluid-solid interface is required. However, the stiffeners affect the dry mode shapes and consequently the fluid pressure experienced by the ship structure.

Fig. 6 shows the first four dry and wet mode-shapes of the ship model. The first four wet modes were closely related to the corresponding first four dry modes, indicated by the dominant contributing factor. Direct correspondence between the dry and wet mode-shapes shows that the choice of modes for modal superposition is appropriate and no further iteration is required in this numerical algorithm. The normalized frequencies are also reported for each mode. As expected, the wet frequencies are lower than the dry ones by one order of magnitude. This implies that the impact of the fluid added-mass is considerable for this type of structure.

Fig. 7 shows the response of the ship at an observation point P due to the given excitation force over a range of frequencies. In the first scenario, the free-surface wave is neglected to detect the resonance occurrence by performing a frequency sweep, as shown by the blue solid line and markers. The second scenario with the damping from the free-surface wave was then conducted with the forced vibration analysis. The displacement results show slightly lower displacement amplitudes compared to the undamped case, indicating that the structure is underdamped. It can be inferred that the free-surface wave only dissipates a small fraction of the vibration energy from the ship within this range of excitation frequencies. This is in contrast with the overdamped case of the hemisphere discussed previously. Thus, it is concluded that including the free-surface wave may result in either underdamped, critically damped, or overdamped vibration, depending on the displacement profile of the mode-shapes. From Table 5, it can be verified that the forced vibration response of the ship is indeed underdamped since all the added damping factors are smaller than one.

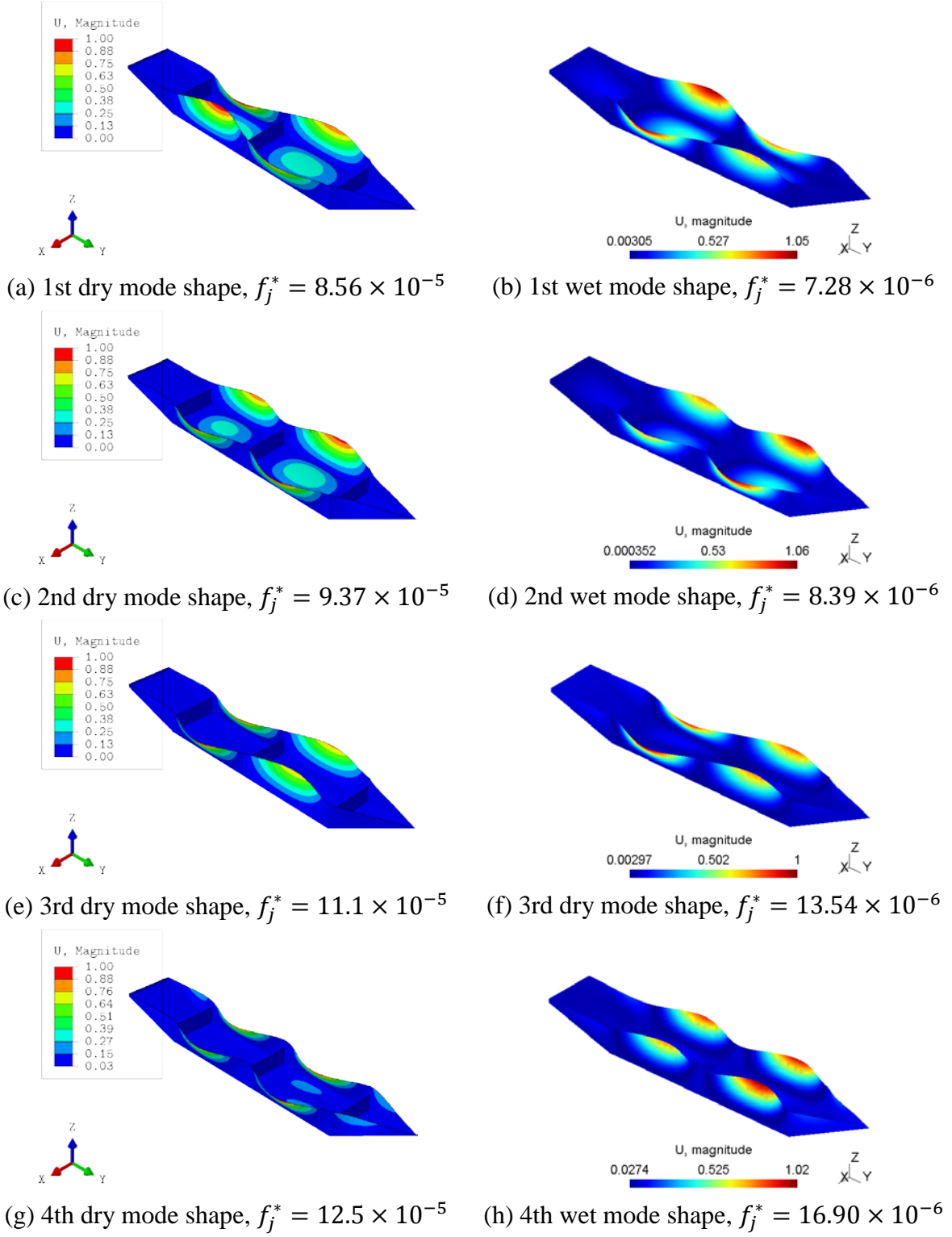


Figure 6: Panels (a), (c), (e) and (g) are the first four dry mode shapes of a simplified ship model with internal plate partitions and panels (b), (d), (f) and (h) are their wet counterparts, respectively, for the immersion depth of $\hat{l} = 2/3$

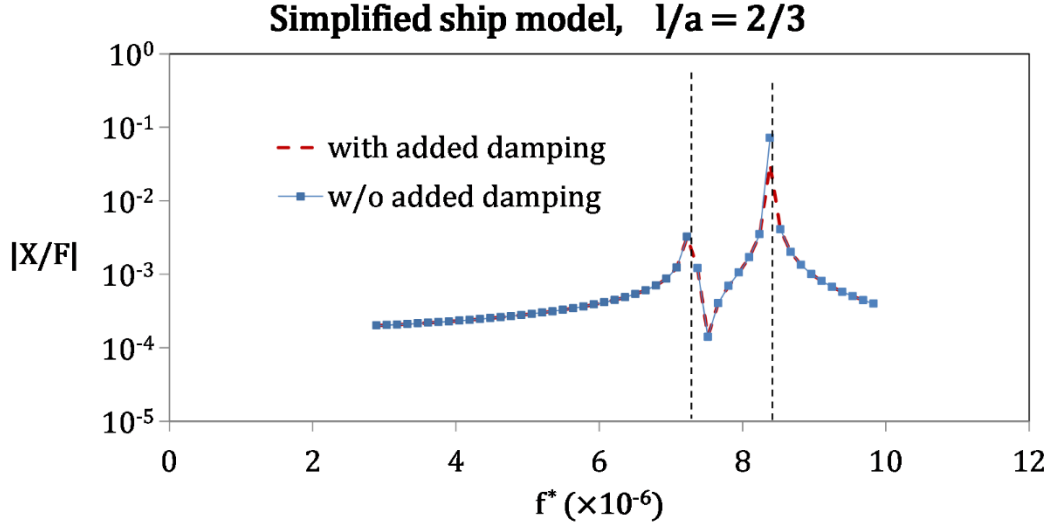


Figure 7: Forced vibration response of the simplified ship model subjected to force F at point P , as shown in Fig. 5b, with and without including the added damping from the free-surface wave effect.

Table 5: Normalized modal added damping matrix for the first four mode shapes of the simplified ship model

$[C_a]_{jk}/([m]_{jj} \times f_j^{(wet)})$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
$j = 1$	5.3×10^{-3}	-0.03	0.00	0.01
$j = 2$	-0.02	0.20	0.02	-0.06
$j = 3$	0.00	0.01	1.5×10^{-3}	0.00
$j = 4$	0.00	-0.05	0.00	0.02

Discussion

The added damping due to the free-surface wave represents a part of vibrational energy that is carried away from the structure. The added-damping may be neglected for free vibration analysis since the purpose is to determine only the natural frequencies. The natural frequencies are functions of mass and stiffness only. However, for steady-state forced vibration analysis, it is necessary to include the added-damping effect.

Conclusions

The proposed numerical algorithm for structural vibration interacting with a fluid combines the use of finite element method, boundary element method and modal superposition. Finite element solver is used to derive the mode-shapes of structure in the absence of fluid (dry modes). Modal superposition is applied to reduce the problem size and calculate the partial fluid loading. The boundary element method is used to calculate the fluid loading, for two scenarios of with and without free-surface waves. A numerical implementation of the modified Green's function was proposed to impose the free-surface wave condition automatically. This ensures that only the fluid-structure interface needs to be discretized for the Boundary Element simulation, leading to a much smaller problem-size.

The proposed numerical scheme was used to study the vibration response of a partially submerged hemispherical shell and simplified ship structure. The results showed the impact of modal added mass on lowering the natural frequencies of vibration. It was also shown that added-damping is large for the hemisphere at low frequencies. The simplified model of a ship

structure was used to study the effects of fluid added-mass and damping for more practical applications. It was demonstrated that the fluid added-mass is significantly larger than the structural modal mass, especially for the lower modes. The proposed formulation provides an efficient algorithm for solving forced vibration problems of fluid-structure interaction since the problem size is reduced to the number of selected mode shapes.

Acknowledgement: This study is supported by the Singapore Maritime Institute (Project ID: SMI-2015-MA-08).

References

- [1] Newman, J. N. (1979), The theory of ship motions, *Advances in applied mechanics* **18**, 221–283.
- [2] Brennen C. E. (1982) A Review of Added Mass and Fluid Inertial Forces. Naval Civil Engineering Laboratory, Sierra Madre CA
- [3] Korotkin A. I. (2008) Added masses of ship structures. Springer Science & Business Media **88**.
- [4] Yeung, R. W. (1981) Added mass and damping of a vertical cylinder in finite-depth waters, *Applied Ocean Research* **3** (3) 119–133.
- [5] Evans, D. V., McIver, P. (1984) Added mass and damping of a sphere section in heave, *Applied Ocean Research* **6** (1) 45–53.
- [6] Rahman, M., Bhatta, D. D. (1993) Evaluation of added mass and damping coefficient of an oscillating circular cylinder, *Applied Mathematical Modelling* **17** (2) 70–79.
- [7] Conca, C., Osses, A., Planchard, J. (1997) Added mass and damping in fluid-structure interaction, *Computer methods in applied mechanics and engineering* **146** (3-4) 387–405.
- [8] Lopes, D. B. S., Sarmento, A. (2002) Hydrodynamic coefficients of a submerged pulsating sphere in finite depth, *Ocean engineering* **29** (11) 1391–1398.
- [9] Sutulo, S., Rodrigues, J. M., Soares, C. G. (2009) Computation of inertial and damping characteristics of ship sections in shallow water, *Ocean Engineering* **36** (14) 1098–1111.
- [10] Sutulo, S., Rodrigues, J. M., Soares, C. G. (2010) Hydrodynamic characteristics of ship sections in shallow water with complex bottom geometry, *Ocean Engineering* **37** (10) 947–958.
- [11] Newman, J. N. (1984) Approximations for the Bessel and Struve functions, *Mathematics of Computation* **43** (168) 551–556.
- [12] Liapis, S. (1992) Numerical methods for water-wave radiation problems, *International journal for numerical methods in fluids* **15** (1) 83–97.
- [13] B. Ponizy, B., Noblesse, F., Ba, M., Guilbaud, M. (1994) Numerical evaluation of free-surface green functions, *Journal of Ship Research* **38** (3) 193–202.
- [14] Chen, G.-R., Fang, M.-C. (2001) Hydrodynamic interactions between two ships advancing in waves, *Ocean Engineering* **28** (8) 1053–1078.
- [15] Kim, B., Shin, Y. S. (2003) A NURBS panel method for three-dimensional radiation and diffraction problems, *Journal of ship research* **47** (2) 177–186.
- [16] Noblesse, F., Huang, F., Yang, C. (2013) The Neumann–Michell theory of ship waves, *Journal of Engineering Mathematics* **79** (1) 51–71.
- [17] Peter, M. A., Meylan, M. H. (2004) The Eigenfunction expansion of the infinite depth free surface Green function in three dimensions, *Wave Motion* **40** (1) 1–11.
- [18] Wang, C. D., Meylan, M. H. (2004) A higher-order-coupled boundary element and finite element method for the wave forcing of a floating elastic plate, *Journal of Fluids and Structures* **19** (4) 557–572.
- [19] Ergin, A., Temarel, P. (2002) Free vibration of a partially liquid-filled and submerged, horizontal cylindrical shell, *Journal of Sound and vibration* **254** (5) 951–965.
- [20] Ergin, A., Price, W. G. (1992) Dynamic characteristics of a submerged, flexible cylinder vibrating in finite water depths, *Journal of Ship Research* **36** (1992) 154–167.
- [21] Chen, S. S., Wambsganss, M. W., Jendrzejczyk, J. A. (1976) Added mass and damping of a vibrating rod in confined viscous fluids, *Journal of Applied Mechanics* **43** (2) 325–329.
- [22] Haddara, M. R., Cao, S. (1996) A study of the dynamic response of submerged rectangular flat plates, *Marine Structures* **9** (10) 913–933.
- [23] Chang, T.-P., Liu, M.-F., (2000) On the natural frequency of a rectangular isotropic plate in contact with fluid, *Journal of Sound and Vibration* **236** 547–553.
- [24] Sepehrirahnama, S., Wijaya, F. B., Ong, E. T., Lee, H. P., Lim, K.-M. (2018) Modal added mass for flexural vibration of floating structures, *Applied Ocean Research* (submitted)
- [25] Craik, A. D. D. (2004) The origins of water wave theory, *Annual Review of Fluid Mechanics* **36** 1–28.
- [26] Abramowitz M, Stegun I. A. (1965) Handbook of mathematical functions with formulas, graphs, and mathematical table, New York: Dover

Computational Fluid Dynamics Modelling of Slurry Transport by Pipeline

***Mingzhi Li¹, †Yanping He^{1,2,3}, Yadong Liu^{1,2,3}, Chao Huang^{1,2,3}**

¹ School of Naval Architecture, Ocean & Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

² State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

³ Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration (CISSE), Shanghai 200240, China

*Presenting author: limz_2008@sjtu.edu.cn

†Corresponding author: hyp110@sjtu.edu.cn

The flow velocity and solid concentration distributions of solid–liquid (slurry) flows transported by pipeline are investigated using a steady three-dimensional (3D) hydrodynamic model based on the kinetic theory of granular flow. Slurries of varying solid particle concentration, grain diameter, and flow conditions are studied, and the effects of particle–particle and particle–wall collisions and near-wall lift force on the concentration distribution are modelled. The simulation agrees well with various experimental results in the literature. The simulation shows that the solid concentration distribution is asymmetric in the vertical plane, and its degree of asymmetry increases as the solid concentration decreases, the mixture velocity decreases, the particle size increases, or the pipe diameter increases. The solid concentration decreases rapidly near the pipe wall due to collisions with the wall. Fine particles smaller than the thickness of the viscous sublayer are most concentrated near the pipe bottom (maximum concentration at the relative location $y_{\max}/D \approx 0.02D$, where D is the pipe diameter) in the viscous sublayer, while the greatest concentration of coarser particles is away from the pipe bottom ($y_{\max}/D \approx 0.1D$), outside of the viscous sublayer. The solid velocity distribution is also asymmetric: maximum-velocity points deviate from the pipe centre, and increasing the solid concentration gradually shifts the point of maximum velocity downward. These results lay a solid foundation for further study of the resistance mechanism and pipe wear, and can be used as a reference for analysing the mesoscopic processes of slurry transport by pipeline.

Keywords: Eulerian multiphase model; Concentration distribution; Particle kinetic; Slurry transport; Pipeline; Dredging; Deep-sea mining

1. Introduction

In recent decades, pipelines have been promoted for the transport of various solids (as slurry mixtures) owing to their insulation from the environment and ability to run uninterrupted, thus reducing the required investment and operating costs. Pipelines can achieve high efficiency, low energy use, environmental protection, and ease of implementation and control. They have been widely applied in many fields such as coal, metallurgy, and mining. Pipelines used for dredging alone have contributed hundreds of billions of US dollars to the global economy in recent years. Pipelines have also been considered as potentially useful in emerging technologies such as ocean mining.

Most research about slurry pipeline transport has focused on predicting friction loss and critical velocity (i.e., no stable particle bed, the lowest pressure loss point), but recently more attention has been paid to the operational costs arising from pipeline wear and maintenance. Industrially transported slurry is generally heterogeneous within the pipeline because the lower half has a greater solid concentration than the upper half (due to gravity), and will thus suffer worse wear. The safety and lifetime of a pipeline can be improved by rotating according to the degree of abrasion around the circumference. This requires accurate prediction of pipeline wear around the circumference, and a proper model for predicting the slurry concentration and velocity distributions is the precondition.

Existing models of slurry pipeline transport can predict with varying degrees of accuracy parameters such as pressure drop, particle settling velocity, and solid concentration distribution under different working conditions, given data regarding the pipe diameter, particle size, slurry concentration, etc. However, most are empirical formulae, based on dimensionless parameters such as excess pressure, Froude number, and solid concentration derived from experimental data, or are semi-empirical formulae based on the theories of gravity, energy, etc. It is impossible theoretically to characterize, for example, the turbulence intensity or particle momentum exchange in a pipeline, but these microscopic characteristics often greatly affect the pipeline characteristics in practice with varying in-situ gradations and operating conditions. Understanding the variability of these parameters in different positions within a pipeline is critical to the proper modelling of factors such as pipe wear, energy loss, and slurry flow regime in practice.

Existing models cannot accurately predict the particle concentration distribution near the bottom of the pipe, especially when the maximum concentration of coarse particles is located away from the bottom (Kaushal and Tomita, 2007). However, the solid concentration distribution near the wall determines the local solid pressure, wall shear stress, and friction resistance, and thus has a critical effect on pipeline wear. Therefore, accurate prediction of the solid concentration distribution, especially near the pipeline wall, is the key not only to predicting wear but also to calculating friction resistance. The solid concentration and velocity distribution (and their variation in a pipe under different conditions) are important to understanding the mechanism of pressure drop and predicting the degree of wear in a pipeline. They can also help improve economic efficiency.

Developments of computer technology and calculation methods have allowed computational fluid dynamics (CFD) to be widely applied in engineering. While CFD has yet to develop full models of solid–liquid two-phase flow (most models refer to gas–liquid two-phase flow theory), simulations of velocity distribution and solid concentration distribution are relatively mature. Current CFD technology includes 3D horizontal pipeline CFD models based on granular kinetic theory, which have been established for comparison with published experimental results and to study the effects on velocity distribution and concentration distribution of the particle concentration, particle size, slurry velocity, and pipe diameter. Before discussing the calculation method and results, a brief view of current work in this field is presented.

2. Previous Work

While the study of slurry pipeline transport has shown continual progress, research has tended to focus on pressure drop and critical velocity. For example, the established Durand (1952)

formula, based on experimental data, is favoured by much of the European dredging industry. The Wilson and Addie (1997) formula is widely used by American dredgers. Wasp *et al.*'s (1977) two-phase flow model considers particle concentration distribution during transport. Turian and Yuan (1977) developed a formula that can fit different dimensionless parameters to experiment data. Doron and Barnea's (1993) formula for a three-layer model is based on mechanical balance, and Lahiri and Ghanta's (2008) formula fits existing experimental data by means of genetic algorithm. The Delft head loss and limit deposit velocity framework (DHLLDV) reported by Miedema and Ramsdell (2015a) probes the mechanism of pressure drop, and developed an applicable and convenient system by using parameters that are easily obtained. Theoretical analysis has improved from the earliest empirical formulae based on purely dimensionless analysis to lift force theory, energy theory, and two-phase flow theory; however, microscopic parameters such as turbulent dissipation force, particle collision force, and particle momentum exchange are not yet properly modelled.

Many scholars have made significant contributions to the study of concentration distributions, including Karabelas (1977), Roco and Shook (1983), Kaushal and Tomita (2002, 2007), Kaushal *et al.* (2005), and Gillies *et al.* (2004). Their various experimental studies have considered variables such as pipe diameter, particle size, and flow conditions. Miedema (2017) and others have published methods to calculate vertical solid-concentration profiles in pipelines given previous experimental data for flow parameters such as eddy diffusivity and particle settling velocity. Each formula can represent the concentration distribution with varying degrees of accuracy, but they rely less on flow parameters than empirical coefficients, thus limiting their applicability and accuracy to the quality of the experimental data and the experience of the user. Experiments by Kaushal and Tomita (2007) for specific particle sizes (diameters of 0.125 and 0.44 mm) found maximum concentrations of coarse particles in the zone away from the wall at about $0.2 D$ rather than at the bottom of the pipe. This result supported the speculation of Wilson and Sellgren (2003) about the effect of near-wall lift on the particle concentration near the bottom of the pipe, but no mathematical model has yet predicted and interpreted this finding (Kaushal *et al.*, 2012).

Numerical simulations include Ling *et al.*'s (2003) simulation of low-density slurry flows in a fully developed turbulent model using the algebraic slip mixture (ASM) in ANSYS Fluent; the results agree well with experimental data. Kaushal *et al.* (2012) carried out numerical simulations of mono-dispersed fine particles at high concentration using the Eulerian model and mixture model in ANSYS Fluent (0.125 mm diameter glass beads in a 54.9 mm pipe); comparison with experimental results showed the Eulerian model to give more accurate predictions for both the pressure drop and concentration profile than the mixture model. However, the model results differed slightly from Kaushal and Tomita's (2007) experimental data, especially near the bottom of the pipeline. Ekambara *et al.* (2009) predicted horizontal solid-liquid pipeline flows under a wide range of conditions using the two-fluid model in ANSYS-CFX, and simulated results close to observed data. Messa *et al.* (2014, 2015) developed a two-fluid model and used it in PHOENICS software to simulate fully suspended liquid-solid slurry flows in horizontal pipes. The model considered turbulent dissipation, momentum exchange, and the influence of wall shear stress on grains, and it provided a method of wall function calculation with improved computing speed and accuracy. No previous models can accurately calculate the particle concentration distribution near the wall, especially for

coarse particles, or calculate the effect of the wall lift force.

In addition, the existing formula for horizontal pipes, although it can predict slurry characteristics under various operating conditions such as friction resistance and solid concentration, cannot be applied to complex geometric spaces such as loop lines and gate valves owing to its empirical nature not taking into account factors such as slurry turbulence, particles collisions, and energy exchange. This limits its applicability in practice, because the pipeline systems for dredging, mining, and coal inevitably include pumps, angular pipes, pipe branches, and other complex spaces. The model will thus fail to understand the properties of the entire system. Therefore, to develop a universal model is an important research goal.

To overcome the above limitations, an integrated model is developed using ANSYS Fluent, based on granular kinetic theory. This model can accurately describe the dynamic characteristics of slurry transport by pipeline.

3. Mathematical Modelling

The Eulerian multiphase model is used here. It mathematically treats the different phases as interpenetrating continua. Granular kinetic theory is used to describe interactions between the particles. A single pressure is shared by all phases. The conservation equations of mass, momentum, and energy are solved individually for each of the phases. Coupling of all phases is then achieved by pressure and interphase exchange coefficients. The model considers energy dissipation and energy exchange caused by particles. Interfacial forces such as the drag force caused by speed differences between phases, the virtual mass force by particle acceleration, the lift force by phase velocity gradient, and other forces are also considered. The Eulerian multiphase model is suitable for simulating slurry transport in pipelines over a wide range of operating conditions.

3.1 Conservation of Mass

Multiphase flow is modelled as a primary phase and n secondary phases. The primary phase is designed as water, and each secondary phase presents particles of different size ranges, which may or may not be of equal volume fraction. These volume fractions, including that for water, are assumed to be continuous in space and time, and their sum is equal to one.

$$\sum_{q=1}^{n+1} \alpha_q = 1 \quad (1)$$

For multiphase flows, each phase volume fraction is less than its maximum allowed value. Therefore, each phase in the model is considered a compressible fluid satisfying the Eulerian continuity equation.

$$\frac{1}{\rho_{rq}} \left[\frac{\partial}{\partial t} (\alpha_q \rho_q) + \nabla \cdot (\alpha_q \rho_q \bar{v}_q) \right] = \sum_{p=1}^n (\dot{m}_{pq} - \dot{m}_{qp}) \quad (2)$$

where ρ_{rq} is the q^{th} phase reference density, or the volume averaged density of the q^{th} phase in the solution domain, α_q is the volume fraction of phase q , \bar{v}_q is the velocity of phase q , and \dot{m}_{pq} characterizes the mass transfer from the p^{th} to q^{th} phase. All these mechanisms can be specified separately.

3.2 Conservation of Momentum

Each phase q in the Eulerian multiphase model must conserve momentum via the following equation:

$$\frac{\partial}{\partial t}(\alpha_q \rho_q \bar{v}_q) + \nabla \cdot (\alpha_q \rho_q \bar{v}_q \bar{v}_q) = -\alpha_q \nabla p - \nabla p_q + \nabla \cdot \bar{\tau}_q + \alpha_q \rho_q \bar{g} + \sum_{p=1}^n (\bar{R}_{pq} + \dot{m}_{pq} \bar{v}_{pq} - \dot{m}_{qp} \bar{v}_{qp}) + \bar{F}_q \quad (3)$$

where p is the pressure, which is equal for each phase at any given point; p_q is the q^{th} phase solid pressure, which is equal 0 for any liquid phase; \bar{g} is acceleration due to gravity; \bar{R}_{pq} is the interphase force; \bar{v}_{pq} is the interphase velocity; n is the total number of phases; and \bar{F}_q is the sum of the external forces (such as lift force, virtual mass force, wall lubrication force and turbulent dispersion force); $\bar{\tau}_q$ is the q^{th} phase stress–strain tensor.

3.3 Conservation of Energy

The first law of thermodynamics is applied in ANSYS Fluent to solve the conservation of energy:

$$\frac{\partial}{\partial t}(\alpha_q \rho_q h_q) + \nabla \cdot (\alpha_q \rho_q \bar{u}_q h_q) = \alpha_q \frac{dp_q}{dt} + \bar{\tau}_q : \nabla \bar{u}_q - \nabla \bar{q}_q + S_q + \sum_{p=1}^n (Q_{pq} + \dot{m}_{pq} h_{pq} - \dot{m}_{qp} h_{qp}) \quad (4)$$

where h_q is the specific enthalpy of the q^{th} phase, \bar{q}_q is the heat flux, S_q is a source term that includes sources of enthalpy, Q_{qp} is the intensity of heat exchange between the q^{th} and p^{th} phases, and h_{qp} is the interphase enthalpy. The heat exchange between phases must comply with the local balance conditions $Q_{qp} = -Q_{pq}$ and $Q_{pp} = Q_{qq} = 0$.

3.4 Solids Pressure

For granular flows, the solids pressure is determined by the intensity of the particle collisions and velocity fluctuations. This work calculates the solids pressure using the model of Lun *et al.* (1984):

$$p_q = \alpha_q \rho_q \Theta_q + \sum_{p=1}^n 2 \frac{d_{pq}^3}{d_q^3} (1 + e_{qp}) g_{0,pq} \alpha_p \alpha_q \rho_q \Theta_q \quad (5)$$

The particle pressure consists of a kinetic term corresponding to the momentum transport caused by particle velocity fluctuations and a second term related to particle collisions.

Granular temperature, Θ_q , is solved by the transport equation derived from kinetic theory.

3.5 Transport Equation

The Eulerian multiphase flow model uses a multi-fluid granular model to describe the flow behaviour of a fluid–solid mixture. Solid-phase stresses are key elements derived by making an analogy between the random particle motion arising from particles' inelastic collisions with each other and the walls. The stresses are defined as a function of granular temperature representing the particle velocity fluctuations proportional to the mean square of the random motion of particles. The granular temperature is solved by the transport equation in the model, as follows:

$$\frac{3}{2} \left[\frac{\partial}{\partial t}(\alpha_q \rho_q \Theta_q) + \nabla \cdot (\alpha_q \rho_q \bar{v}_q \Theta_q) \right] = (-p_q \bar{I} + \bar{\tau}_q) : \nabla \bar{v}_q + \nabla \cdot (k_{\Theta q} \nabla \Theta_q) - \gamma_{\Theta q} + \phi_{pq} \quad (6)$$

The left-hand side of the equation represents the net change in fluctuating energy. The first term, $(-p_q \bar{I} + \bar{\tau}_q) : \nabla \bar{v}_q$, on the right-hand side represents the generation of energy by the solid stress tensor. The second term, $\nabla \cdot (k_{\Theta q} \nabla \Theta_q)$ is the diffusion of energy in the solid phase. The third term, $\gamma_{\Theta q}$, is the collisional dissipation of energy, and ϕ_{pq} is the exchange of fluctuating energy between the liquid and the solid phase.

The term $k_{\Theta q}$ is the diffusion coefficient given by Gidaspow *et al.* (1992) as an optional model in ANSYS Fluent:

$$k_{\Theta q} = \frac{150\rho_q d_q \sqrt{\Theta_q \pi}}{384(1+e_q)g_{0,q}} \left[1 + \frac{6}{5} \alpha_q g_{0,q} (1+e_q) \right]^2 + 2\rho_q \alpha_q^2 d_q (1+e_q) g_{0,q} \sqrt{\frac{\Theta_q}{\pi}} \quad (7)$$

where $\gamma_{\Theta q}$ is the rate of energy dissipation within the q^{th} solid phase due to collisions between particles. Lun *et al.* (1984) give $\gamma_{\Theta q}$ as an optional model:

$$\gamma_{\Theta q} = \frac{12(1+e_q^2)g_{0,q}}{d_q \sqrt{\pi}} \rho_q \alpha_q^2 \Theta_q^{1.5} \quad (8)$$

Although equation (6) can be solved for the granular temperature, the procedure is complex and convergence is difficult. ANSYS Fluent by default uses a simpler and computationally more efficient model called “algebraic formulation” that neglects convection and diffusion in the transport equation.

3.6 Turbulence Equations

The per-phase turbulence model is used here. It includes a set of k - ε transport equations for each phase:

$$\begin{aligned} \frac{\partial(\alpha_q \rho_q k_q)}{\partial t} + \nabla \cdot (\alpha_q \rho_q \bar{U}_q k_q) = & \nabla \cdot \left(\alpha_q \left(\mu_q + \frac{u_{t,q}}{\sigma_k} \right) \nabla k_q \right) + (\alpha_q G_{k,q} - \alpha_q \rho_q \varepsilon_q) \\ & + \sum_{p=1}^n K_{pq} (C_{pq} k_p - C_{qp} k_q) - \sum_{p=1}^n K_{pq} (\bar{U}_p - \bar{U}_q) \frac{u_{t,p}}{\alpha_p \sigma_p} \nabla \alpha_p \\ & + \sum_{p=1}^n K_{pq} (\bar{U}_p - \bar{U}_q) \frac{u_{t,q}}{\alpha_q \sigma_q} \nabla \alpha_q + \Pi_{kq} \end{aligned} \quad (9)$$

and

$$\begin{aligned} \frac{\partial(\alpha_q \rho_q \varepsilon_q)}{\partial t} + \nabla \cdot (\alpha_q \rho_q \bar{U}_q \varepsilon_q) = & \nabla \cdot \left(\alpha_q \left(\mu_q + \frac{u_{t,q}}{\sigma_\varepsilon} \right) \nabla \varepsilon_q \right) + \frac{\varepsilon_q}{k_q} (C_{1\varepsilon} G_{k,q} - C_{2\varepsilon} \rho_q \varepsilon_q) \\ & + C_{3\varepsilon} \frac{\varepsilon_q}{k_q} \left[\sum_{p=1}^n K_{pq} (C_{pq} k_p - C_{qp} k_q) - \sum_{p=1}^n K_{pq} (\bar{U}_p - \bar{U}_q) \frac{u_{t,p}}{\alpha_p \sigma_p} \nabla \alpha_p \right] \\ & + C_{3\varepsilon} \frac{\varepsilon_q}{k_q} \sum_{p=1}^n K_{pq} (\bar{U}_p - \bar{U}_q) \frac{u_{t,q}}{\alpha_q \sigma_q} \nabla \alpha_q + \Pi_{\varepsilon q} \end{aligned} \quad (10)$$

where $C_{1\varepsilon}$ is the C1-epsilon number, $C_{2\varepsilon}$ is the C2-epsilon number, $C_{3\varepsilon}$ is the C3-epsilon number, σ_k is the TKE Prandtl number, and σ_ε is the TKR Prandtl number. Their default values are 1.44, 1.92, 1.3, 1.0, and 1.3, respectively.

4. Simulation Method

4.1 Physical Model

To ensure the model’s adaptability and universal applicability, a large range of pipe diameters, particle sizes, and solid concentrations are considered here. To allow comparison with experimental data, the horizontal pipes are modelled with inner diameters, D , of 51.5, 54.9, 103, 206, and 495 mm; in each case the length, $L \approx 60D$. To ensure the computations are of good quality and able to converge, 30 boundary layers are established along the surface with a growth factor of 1.2 (i.e., each row of the boundary layer mesh is 20% thicker than the previous one), and the outermost layer has a height of about 0.08 mm. The first layer height from the wall of these models, expressed as a dimensionless parameter y^+ , is $y^+ < 30$, and reaches $y^+ < 15$. These 3D models include around 10 million meshes, as shown in figure 1.

The following physical properties are employed: for the liquid phase, density

$\rho_l = 998.2 \text{ kg/m}^3$ and dynamic viscosity $\mu_l = 1.003 \times 10^{-3} \text{ Pa}\cdot\text{s}$, and for the solid phase, particle density $\rho_s = 2470\text{--}2650 \text{ kg/m}^3$, particle diameter $d_p = 0.09, 0.125, 0.165, 0.27, \text{ and } 0.44 \text{ mm}$, the limiting volume concentration is $0.6\text{--}0.7$, and the default value of the internal friction angle is 30° .

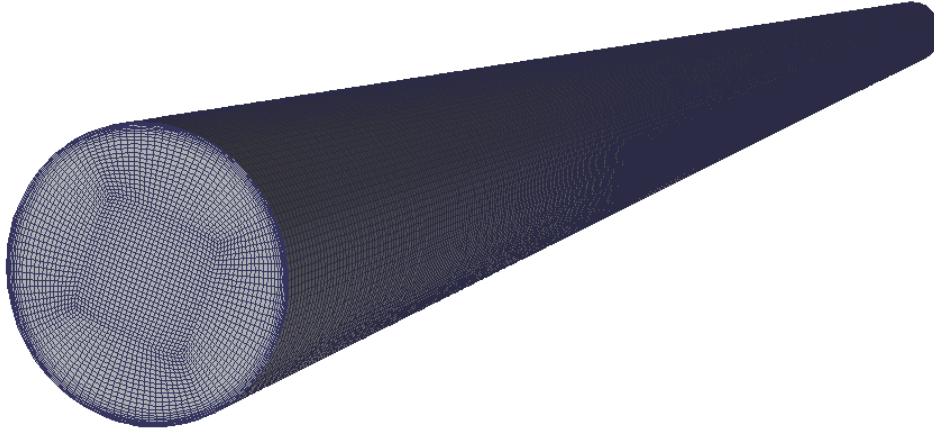


Figure 1: Grid structure for pipeline model.

4.2 Boundary Conditions

At the inlet, a velocity-inlet condition is selected. The velocities and concentrations of both phases are given specified values, with the particle phase having a slightly lower velocity than the liquid phases and the coarse particles being about 5% slower than fine particles. Turbulent intensity, turbulent viscosity ratio, and temperature are all set to their default values of 5%, 10, and 20°C , respectively.

At the outlet, the pressure-outlet is selected, and the pressure is atmospheric. At the wall, the velocity of the liquid phase is set to zero (i.e., no-slip condition). The wall roughness is set to 0.02 mm , and the specularity coefficient (taken as 0.451) is selected for the shear condition of the solid phase. The particle condition follows the Johnson–Jackson model, and the restitution coefficient is set to 0.2 .

4.3 Solving Process and Convergence Scheme

Commercial CFD software ANSYS Fluent 17.0 is used to solve the above continuum equations and boundary conditions. Convergence of the root-mean square residual is set to 10^{-5} . The solving method follows the phase-coupled SIMPLE function to ensure convergent, steady, and accurate results. The second-order upwind method is adopted to solve momentum equations, with the pressure relaxation factor set to 0.2 , the momentum relaxation factor set to 0.3 , and the volume fraction set to 0.4 ; other factors retain their default values.

5. Results and Discussion

To analyse a broad range of simulation results and experimental data in a limited space, some typical experimental conditions (Roco and Shook, 1983; Kaushal and Tomita, 2005; and Gillies *et al.*, 2004) are simulated with ANSYS Fluent. They cover a wide range of particle diameters ($0.09\text{--}0.44 \text{ mm}$), particle volume concentrations ($9\%\text{--}50\%$), slurry flow rates ($2\text{--}5 \text{ m/s}$), and pipe diameters ($51.5\text{--}495 \text{ mm}$). The simulation results are validated through comparison with the corresponding measured data for the steady-state particle concentration distribution and the particle and liquid velocity distributions in the pipes.

Three works by Kaushal and Tomita (2002, 2007) and Kaushal et al. (2005) give similar concentration profiles, but there are differences among a few cases for coarse (0.44 mm) particles at low flow velocities (Kaushal and Tomita, 2007). The values tabulated in Kaushal et al. (2005) are clearer and more convenient than the data given as figures in the other papers, and are thus selected here for comparison with the simulation results.

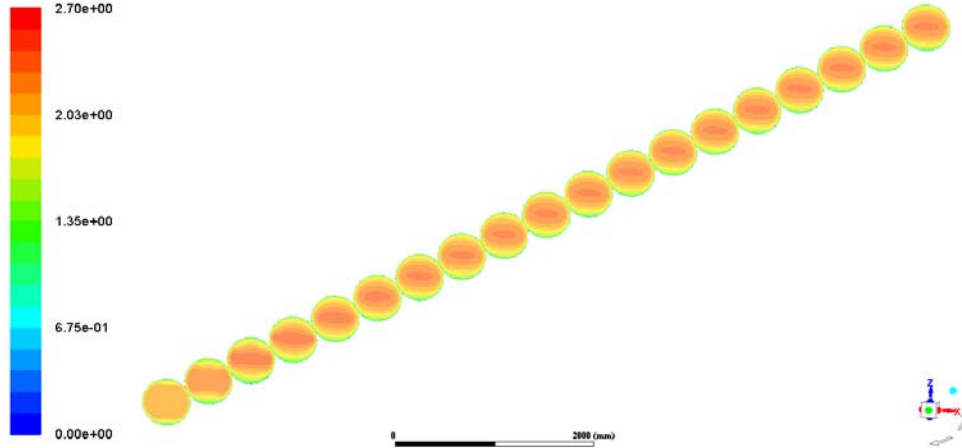


Figure 2: Contour plots for particle velocity taken at regularly spaced axial positions along the pipeline for the following conditions: $D = 495$ mm, $d_p = 0.165$ mm, $C_{vf} = 27.3\%$, and $v = 3.16$ m/s

Figure 2 shows contour plots of liquid velocity distribution along the pipe cross-section at axial positions separated by $0.05L$, where L is the length of the pipe simulation model. The simulation conditions are pipe diameter 495 mm, particle diameter $d_p = 0.165$ mm, solid volume concentration 27.3%, and slurry rate $v = 3.16$ m/s. The distributions differ significantly among the first six sections, but all the subsequent distributions appear nearly identical in each case. This shows that the model pipeline has sufficient length (60 times its diameter) and that the numerical slurry simulations provide fully developed results. The following simulation results are obtained near the outlet of the pipe model.

5.1 Solid Concentration Distribution

figure 3 shows a set of solid volume concentration contours on the left, and the curve on its centreline compared with experimental data on the right for 0.44 mm glass ball slurries flowing at a constant mixture velocity (4 m/s) in a 54.9 mm pipeline. The slurries differ in their solid volume concentrations: 20%, 30%, 40%, and 50%. The experimental data were initially reported by Kaushal et al. (2005). The simulation reasonably coincides with the experimental results, indicating that the model is suitable for simulating the particle concentration distribution for a wide range of solid concentrations.

As figure 3 shows, the solid volume concentration is asymmetric in the perpendicular direction. The asymmetry is reduced as the solid volume concentration increases at a given velocity, pipe diameter, and particle size because of increased particle–particle collisions. In figure 3A–C, the distances of the points of greatest particle concentration from the bottom of the pipe (about $0.1D$) are simulated here for the first time: the results fit the experimental data well, indicating that the model can simulate the effect of wall lift force on the particle concentration distribution.

The point of greatest solid concentration in figure 3D is nearer to the pipe bottom than in the other images, because increasing the particle concentration leads to more particle–particle

and particle–wall collisions, which have a uniform effect far beyond the effect of the near-wall lift force. Increased interactions between particles increases the particle virtual mass force, and thus increases the ability of particles to remain suspended, leading to a more symmetric distribution in the pipe.

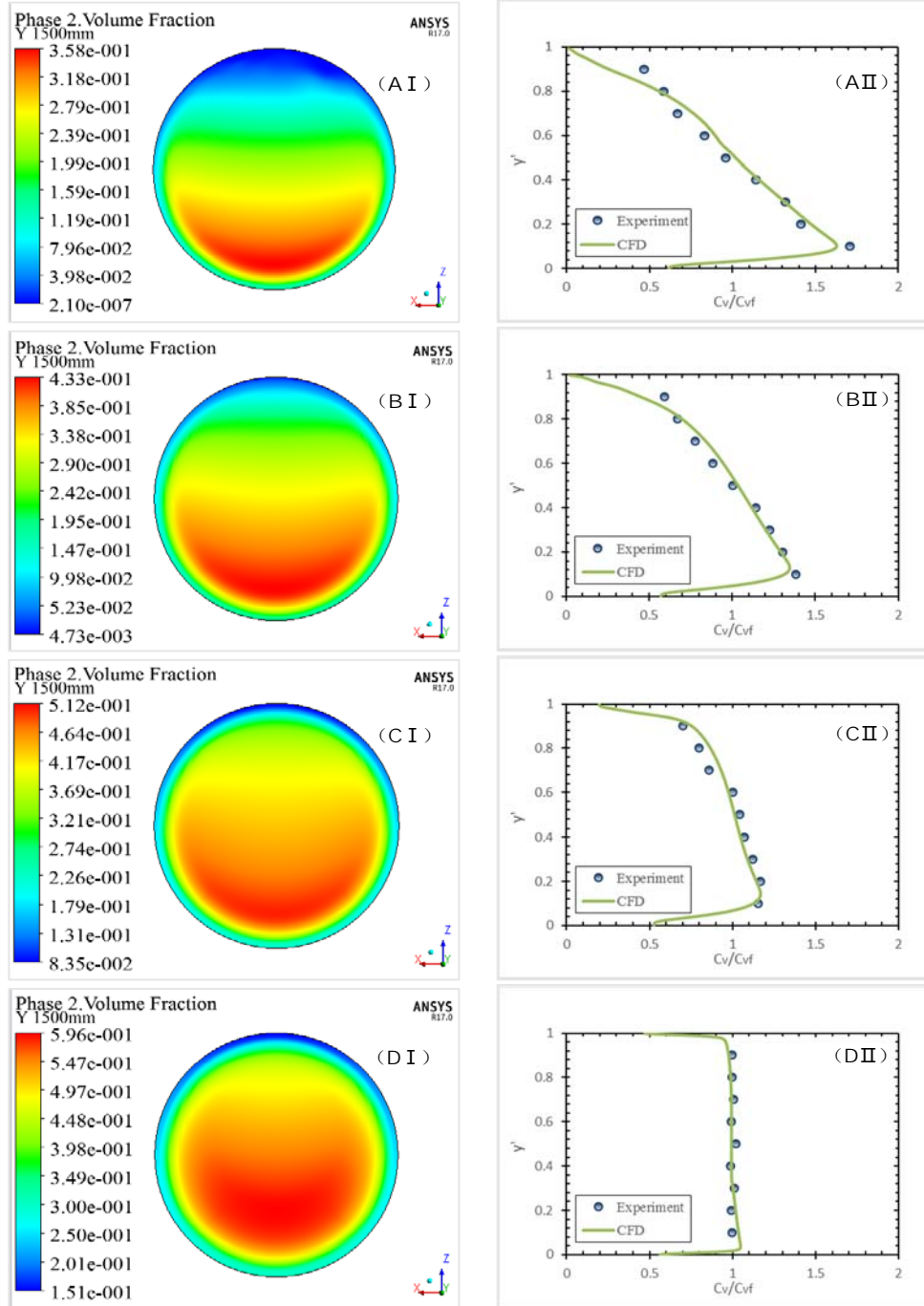


Figure 3: Particle volume concentration distribution for $D = 54.9$ mm, $d_p = 0.44$ mm, $v = 4$ m/s, and (A) $C_{vf} = 20\%$, (B) $C_{vf} = 30\%$, (C) $C_{vf} = 40\%$, and (D) $C_{vf} = 50\%$.

The ordinate axis, $y' = y/D$, is the dimensionless position along the pipe's vertical axis, where y is the distance from the pipe bottom, and the abscissa, C_v/C_{vf} , represents the relative solid volume concentration ratio on the vertical centreline of the pipe, where C_v is the local solid volume concentration and C_{vf} is the efflux solid volume concentration.

Figure 4 gives simulated solid concentration distributions for 0.125 mm glass balls in a

54.9 mm pipe with 30% solid concentration at different slurry velocities (2, 3, 4, and 5 m/s), and compares them with the experimental data of Kaushal et al. (2005). The results are nearly the same as the experimental data, demonstrating that the model can accurately simulate the concentration distribution of slurry pipeline transportation at a wide range of velocities.

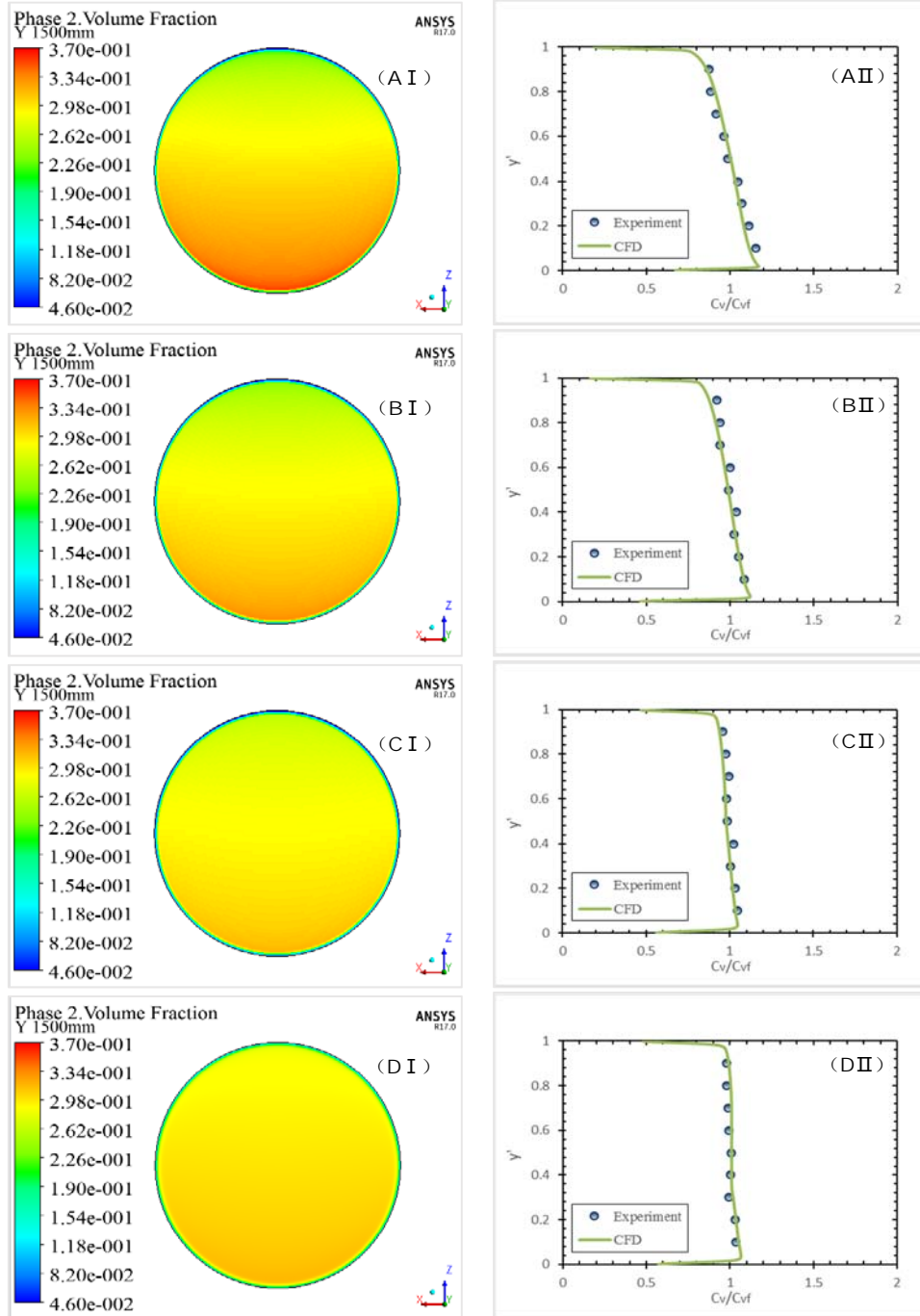


Figure 4: Particle volume concentration distributions for $D = 54.9$ mm, $d_p = 0.125$ mm, $C_v = 30\%$, and (A) $v = 2$ m/s, (B) $v = 3$ m/s, (C) $v = 4$ m/s, and (D) $v = 5$ m/s.

The asymmetry of the slurry concentration curves is significantly reduced as the slurry velocity increases for the given conditions of constant concentration, pipe diameter, and particle size. This result arises because the particles are more easily suspended at increased velocity, which increases turbulence intensity and turbulent dissipation forces.

The maximum solid concentration is situated closer to the pipe bottom than in figure 3A–

C, and appears in the viscous sublayer. It remains close to the wall, and the particle concentrations then decrease rapidly because the particle size is less than the thickness of the viscous sublayer; thus, particles are not influenced by the near-wall lift force, and instead suffer from collisions with the wall.

The simulation results in figure 5 for four different particle sizes (0.09, 0.125, 0.27, and 0.44 mm) all agree well with experimental values by Kaushal et al. (2005) and Gillies *et al.* (2004).

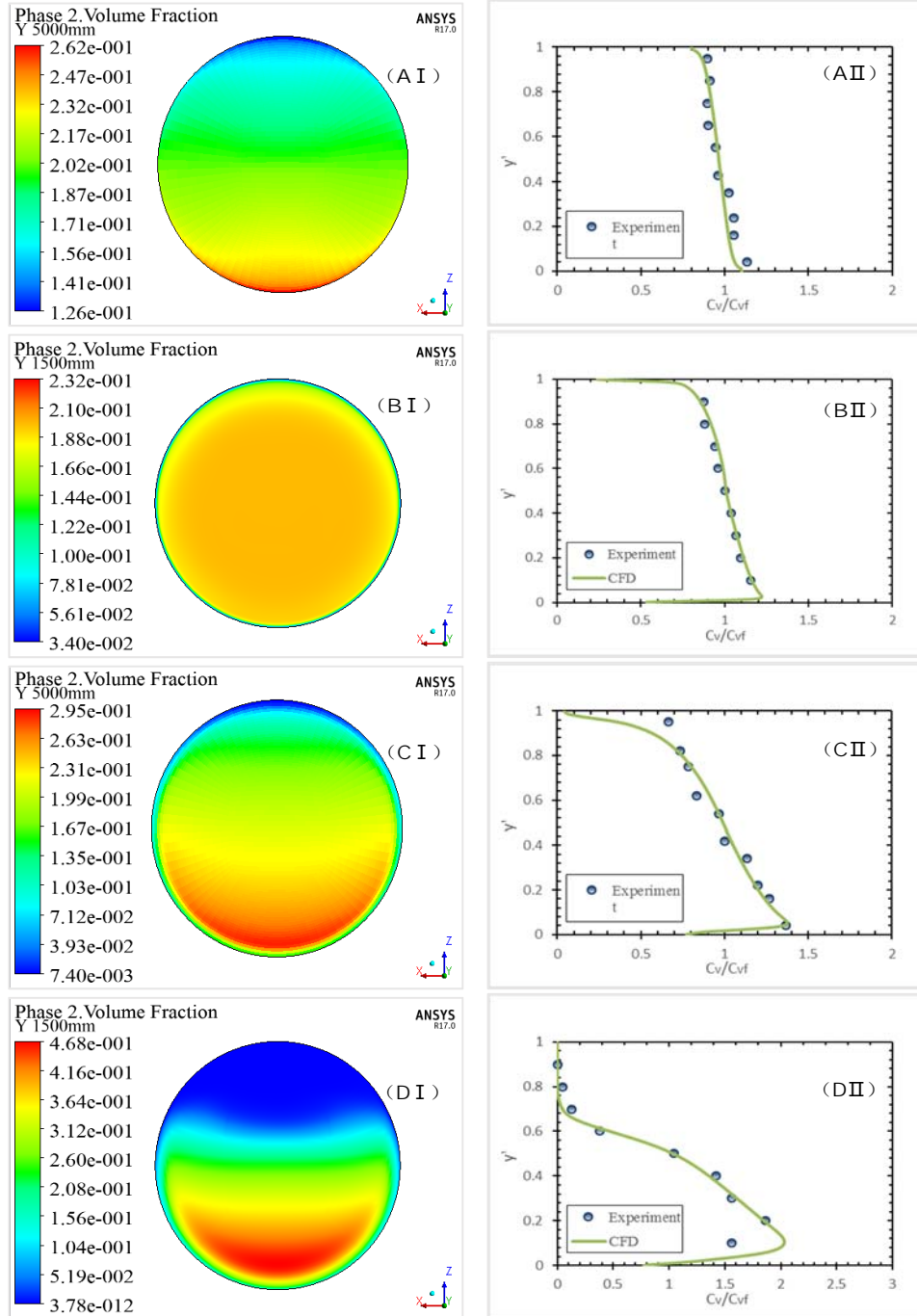


Figure 5: Particle volume concentration distributions for (A) $d_p = 0.09$ mm, $D = 103$ mm, $C_{vf} = 19\%$, and $v = 3$ m/s; (B) $d_p = 0.125$ mm, $D = 54.9$ mm, $C_{vf} = 20\%$, and $v = 2$ m/s; (C) $d_p = 0.27$ mm, $D = 103$ mm, $C_{vf} = 30\%$, and $v = 5.4$ m/s; and (D) $d_p = 0.44$ mm, $D = 54.9$ mm, $C_{vf} = 20\%$, and $v = 2$ m/s.

For similar flow conditions, the figure shows that as the particles become larger, their concentration distributions become more asymmetric in the pipe cross-section. The method of classifying slurry flow regimes given by Wasp et al. (1977) describes figure 5A (0.09 mm) as homogeneous, figure 5B and C (0.125 and 0.27 mm, respectively) as heterogeneous, and figure 5D (0.44 mm) as a sliding bed. Three different flow regimes and their changing trends with respect to particle size are thus simulated successfully.

Figure 5C and D appear similar to figure 3A–C, in terms of the location of the point of maximum solid volume concentration relative to the pipe bottom. Figure 5A and B are similar to figure 3D, with the reversal situated in the viscous sublayer.

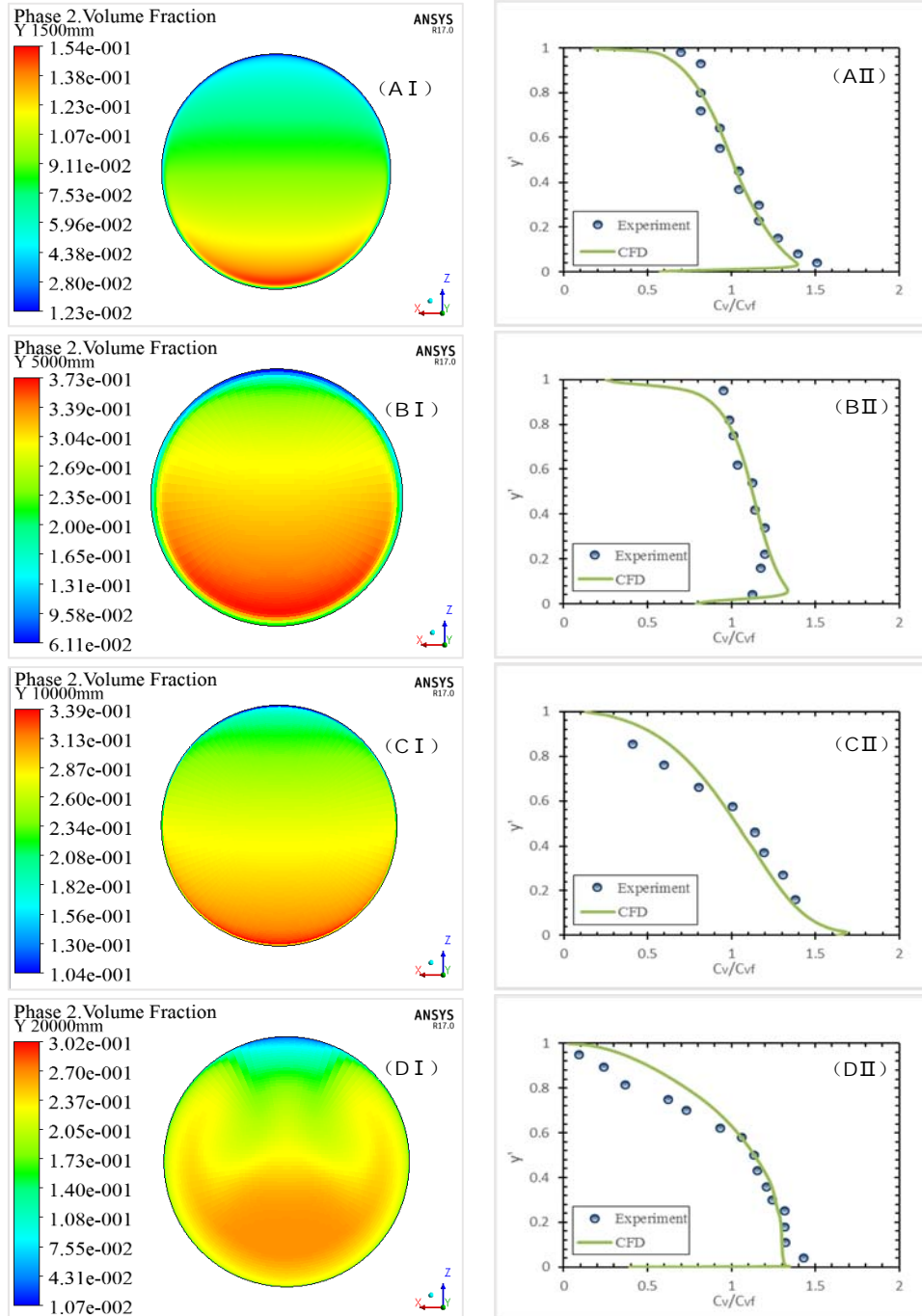


Figure 6: Particle volume concentration distribution for (A) $D = 51.5$ mm, $d_p = 0.165$ mm, $C_{vf} = 9.18\%$, and $v = 3.78$ m/s; (B) $D = 103$ mm, $d_p = 0.27$ mm, $C_{vf} = 40\%$, and $v = 5.4$ m/s;

(C) $D = 263$ mm, $d_p = 0.165$ mm, $C_{vf} = 27\%$, and $v = 2.9$ m/s; and (D) $D = 495$ mm, $d_p = 0.165$ mm, $C_{vf} = 27.3\%$, and $v = 3.16$ m/s.

The effects of pipe diameter on particle concentration distribution are modelled and compared with experimental data. The pipe models have diameters of 51.5, 103, 263, and 495 mm, corresponding to the experiments conducted by Roco & Shook (1983) and Gillies *et al.* (2004). The simulated results are consistent with the experimental results. Within a certain range, increasing the pipe diameter increases the asymmetry of the slurry concentration distribution on the vertical axis in the pipeline.

The clear difference between the simulation (which shows the greatest particle concentration situated at $0.1D$ above the pipe bottom) and the experiment (which shows it at $0.2D$ above the pipe bottom) in figures 4CII, 6DII, and 7BII may arise owing to the sparse concentration-measuring points near the pipe wall in the experiment.

5.2 Velocity Distribution

The solid velocity distribution is inextricably linked to its concentration distribution within the pipeline. A symmetric solid concentration distribution will have a symmetric velocity distribution. Figure 7 shows solid velocity distribution contours on the left and curves for the centreline compared with corresponding experimental data on the right. Figure 7A shows the simulated solid velocity distribution for 0.09 mm particles at 19% concentration in a 103 mm pipe with 3 m/s flow speed and the comparison with experimental data by Gillies *et al.* (2004). The distribution is relatively symmetric, and the point of maximum speed is slightly over the pipe centre.

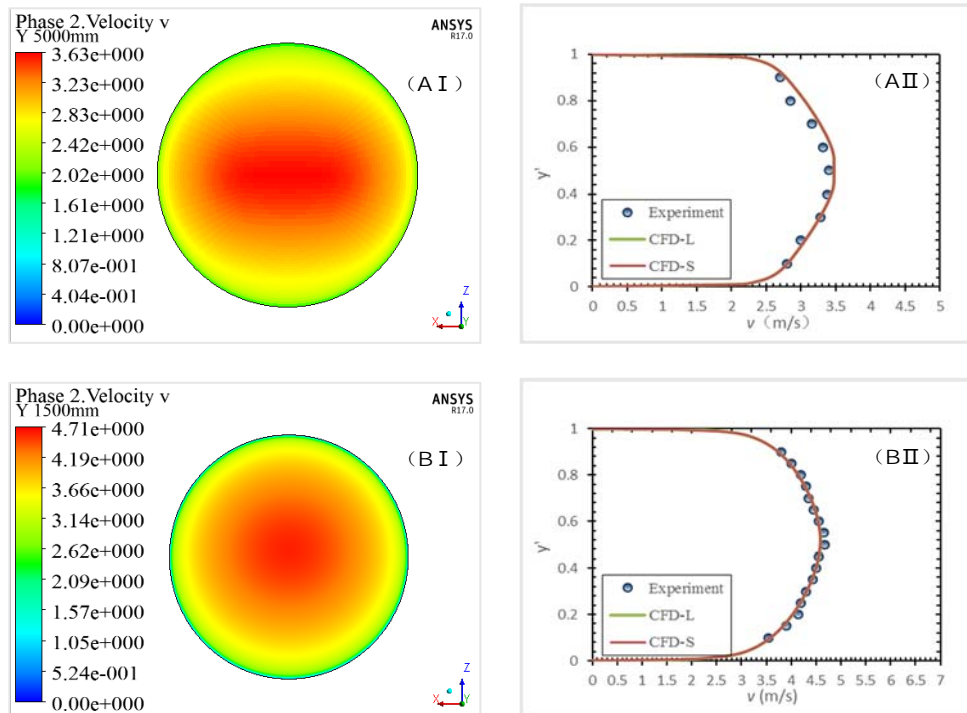


Figure 7: Solid velocity distribution for (A) $D = 103$ mm, $d_p = 0.09$ mm, $C_{vf} = 19\%$, and $v = 3$ m/s; and (B) $D = 51.5$ mm, $d_p = 0.165$ mm, $C_{vf} = 9.18\%$, and $v = 3.78$ m/s.

Figure 7 also shows that particle velocity and liquid velocity are similar; thus, particle slip velocity of the slurry is very small. Figure 7B is the simulated solid velocity distribution for 0.165 mm particles at 9.18% concentration in a 51.5 mm pipe with 3.78 m/s mixture speed and

a comparison with experimental data by Roco & Shook (1983). The simulation agrees well with the experimental results, and its distribution is symmetric in the horizontal direction.

Figure 8 shows the simulated solid velocity distribution for 0.44 mm particles at different concentrations (20%, 30%, 40%, and 50%) in a 54.9 mm pipe at 4 m/s velocity (the solid volume concentration distribution is shown in figure 3). As the solid concentration increases, the point of maximum velocity gradually moves downward.

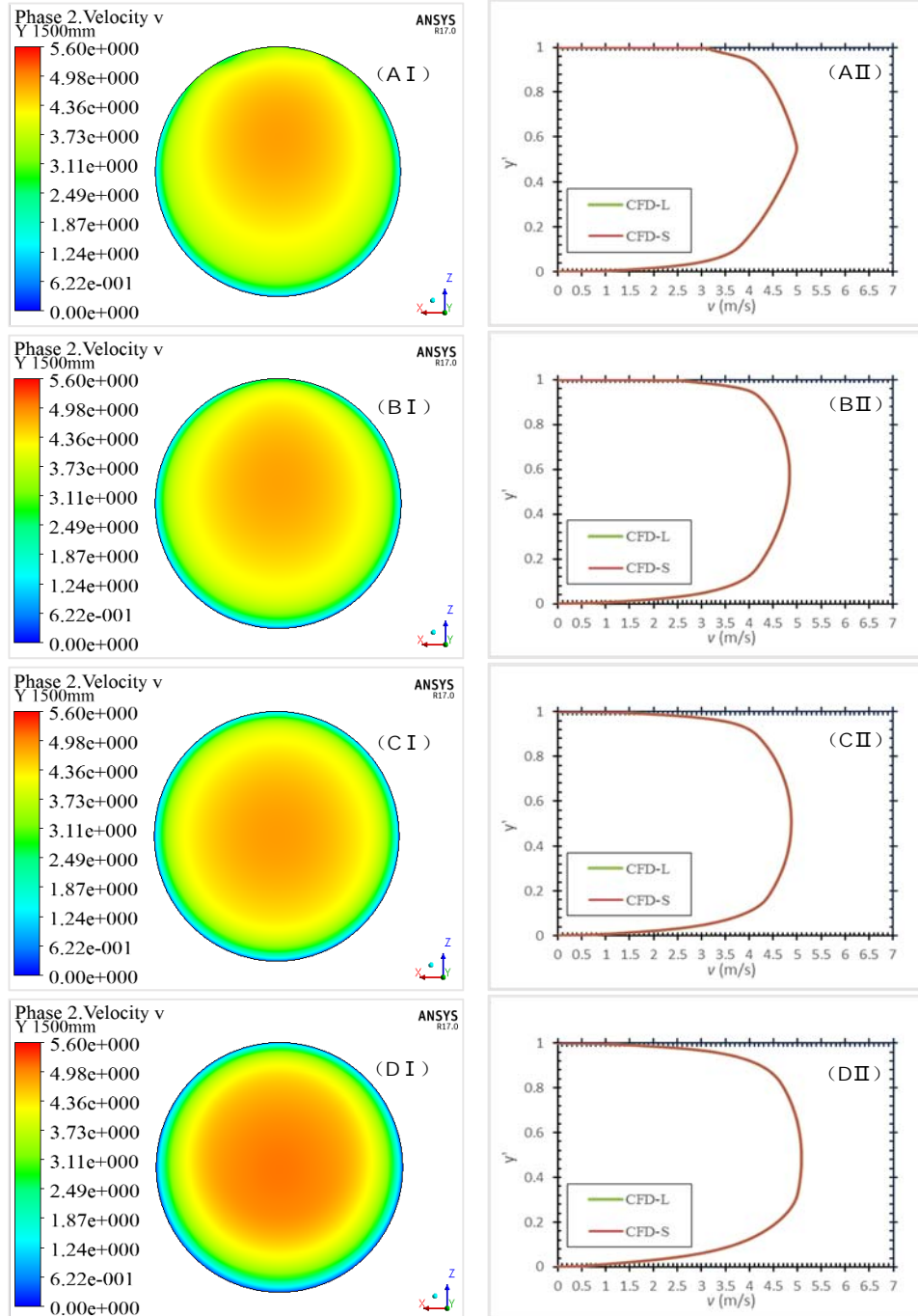


Figure 8: Solid velocity distribution for $D = 54.9$ mm, $d_p = 0.44$ mm, $v = 4$ m/s, and (A) $C_{vf} = 20\%$, (B) $C_{vf} = 30\%$, (C) $C_{vf} = 40\%$, and (D) $C_{vf} = 50\%$.

Figure 9 exhibits the simulated solid velocity distribution for 0.125 mm particles at 30% concentration in a 54.9 mm pipe with different velocities (2, 3, 4, and 5 m/s) (the solid volume

concentration distribution is shown in figure 4). The points of maximum velocity are close to the pipe centre.

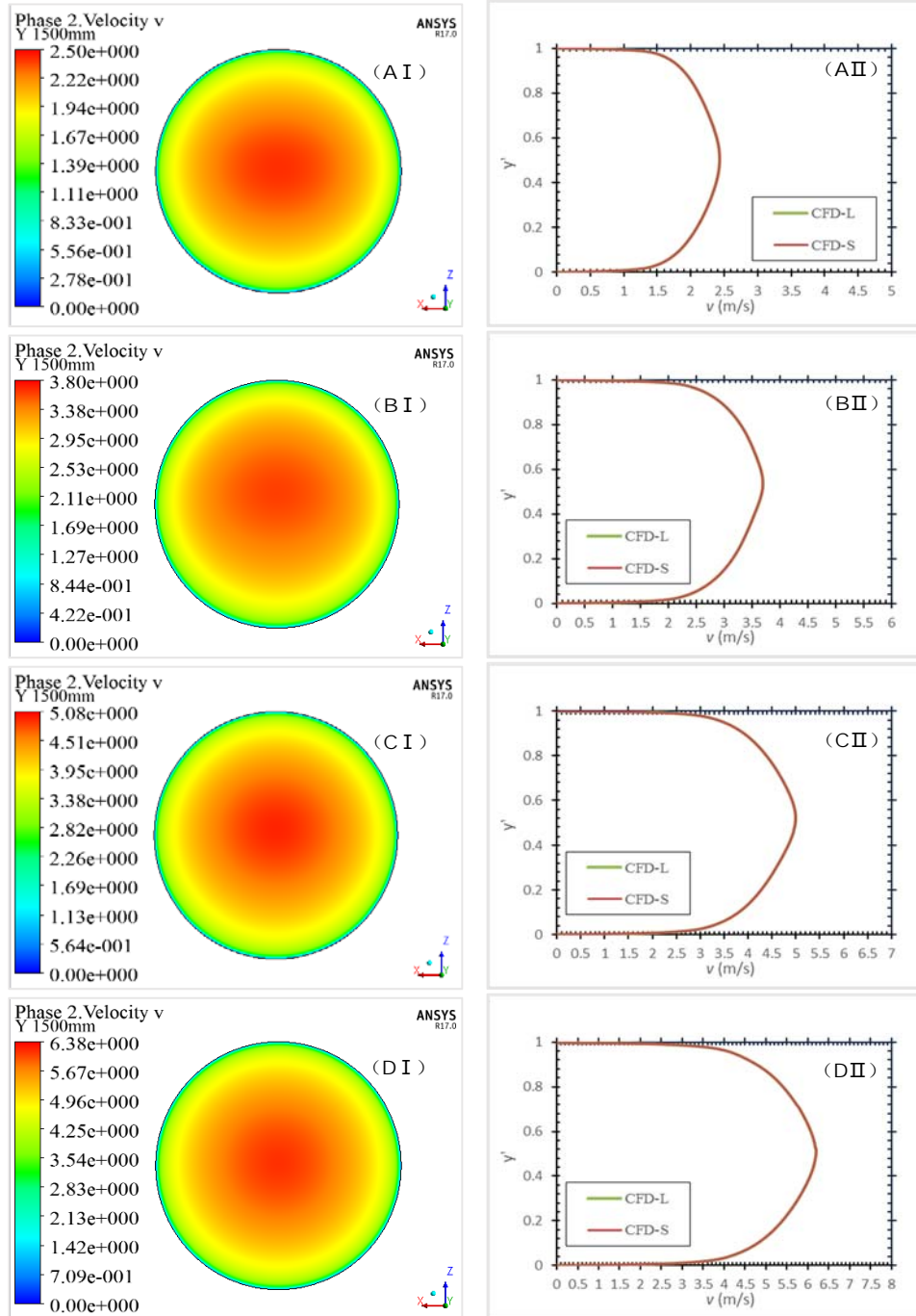


Figure 9: Solid velocity distribution for $D = 54.9$ mm, $C_{vf} = 30\%$, $d_p = 0.125$ mm, and (A) $v = 2$ m/s, (B) $v = 3$ m/s, (C) $v = 4$ m/s, and (D) $v = 5$ m/s.

Comparing figures 9B and 10C (particle diameter 0.44 vs. 0.125 mm) shows that larger particles at the same velocity and concentration condition will be more asymmetrically distributed with a larger deviation of the point of maximum velocity from the pipe centre. This demonstrates that larger particles experience a greater influence of gravity, and the concentration increases at the bottom of the pipe, leading to a lower concentration and higher velocity at the top of the pipe.

6. Conclusions

A steady three-dimensional hydrodynamic model of slurry transport by pipeline is developed here based on the kinetic theory of granular flow. Parameters including solid volume concentration and velocity distributions are simulated for a wide range of typical working conditions, and compared with experimental data by Roco and Shook (1983), Kaushal and Tomita (2005), and Gillies *et al.* (2004). Over a wide range of situations, such as different particle sizes, particle volume concentrations, mixture velocities, and pipe diameters, the model's predictions agree well with the experimental data. The difference between the simulation results (which show maximum particle concentration situated at $0.1D$ above the pipe bottom) and the results of Kaushal and Tomita (2007) (who found the maximum particle concentration significantly above the pipe bottom) is well depicted in concentration profiles such as figures 4CII, 6DII and 7BII. Overall, the Eulerian multiphase model based on the kinetic theory of granular flow appears capable of predicting the solid concentration and velocity of slurry flows in pipelines. It also handles well the regime of slurry flow under a range of conditions, the effect of the near-wall lift force on the coarse particle concentration distribution, and the effect of particle-wall collisions on the solid concentration distribution near the wall.

The simulation shows that the solid volume concentration and velocity distributions in a pipe depend on factors such as the mixture velocity, pipe diameter, particle size, slurry concentration, and solid density. On the vertical centreline of the pipe, particle concentration and velocity distribution are asymmetric, with the degree of asymmetry depending on the pipe diameter, particle size, mixture velocity, and solid volume concentration. The asymmetry of the particle concentration distribution increases with increasing particles size, pipe diameter, and mixture velocity, and with decreasing solid concentration. Increasing the solid concentration gradually shifts the point of maximum velocity downward.

Most of the boundary conditions of the model are default values for the ANSYS Fluent program, thus simplifying the input data. The calculations are fast and easily convergence, making the present scheme a widely applicable and useful model.

Acknowledgments

This research work is supported by the National Natural Science Foundation of China (Grant No. 51779143) and State Key Laboratory of Ocean Engineering of Shanghai Jiao Tong University (Grant No. GKZD010071).

References

- Durand R. (1952) *The Hydraulic Transportation of Coal and Other Materials in Pipes*, Collage of National Coal Board, London.
- Doron, P., and Barnea, D. (1993) A three-layer model for solid-liquid flow in horizontal pipes, *International Journal of Multiphase Flow* **19**, 1029-1043.
- Ekambara, K., Sanders, R. S., Nandakumar, K., and Masliyah, J. H. (2009) Hydrodynamic simulation of horizontal slurry pipeline flow using ANSYS-CFX, *Industrial and Engineering Chemistry Research* **48**, 8159-8171.
- Gidaspow D., Bezburuah R., Ding J. (1992) Hydrodynamics of Circulating Fluidized Beds: Kinetic Theory Approach, *In Fluidization VII, Proceedings of the 7th Engineering Foundation Conference on Fluidization* 75–82.
- Gillies, R. G., Shook, C. A., and Xu, J. (2004) Modelling heterogeneous slurry flows at high velocities, *Canadian Journal of Chemical Engineering* **82**, 1060-1065.
- Karabelas, A. J. (1977) Vertical distribution of dilute suspensions in turbulent pipe flow, *Aiche Journal* **23**, 426–

434.

- Kaushal, D. R., and Tomita, Y. (2002) Solids concentration profiles and pressure drop in pipeline flow of multisized particulate slurries, *International Journal of Multiphase Flow* **28**, 1697-1717.
- Kaushal, D. R., Sato, K., Toyota, T., Funatsu, K., and Tomita, Y. (2005) Effect of particle size distribution on pressure drop and concentration profile in pipeline flow of highly concentrated slurry. *International Journal of Multiphase Flow* **31**, 809-823.
- Kaushal, D. R., and Tomita, Y. (2007) Experimental investigation for near-wall lift of coarser particles in slurry pipeline using γ -ray densitometer, *Powder Technology* **172**, 177-187.
- Kaushal, D. R., Thinglas, T., Tomita, Y., Kuchii, S., and Tsukamoto, H. (2012) Cfd modeling for pipeline flow of fine particles at high concentration, *International Journal of Multiphase Flow* **43**, 85-100.
- Lahiri, S. K., and Ghanta, K. C. (2008) Prediction of pressure drop of slurry flow in pipeline by hybrid support vector regression and genetic algorithm model, *Chinese Journal of Chemical Engineering* **16**, 841-848.
- Ling, J., Skudarnov, P. V., Lin, C. X., and Ebadian, M. A. (2003) Numerical investigations of liquid–solid slurry flows in a fully developed turbulent flow region, *International Journal of Heat and Fluid Flow* **24**, 389-398.
- Lun, C. K. K., Savage, S. B., Jeffrey, D. J., and Chepurniy, N. (2006) Kinetic theories for granular flow: inelastic particles in couette flow and slightly inelastic particles in a general flow field, *Journal of Fluid Mechanics* **140**, 223-256.
- Messa, G. V., Malin, M., and Malavasi, S. (2014) Numerical prediction of fully-suspended slurry flow in horizontal pipes, *Powder Technology* **256**, 61-70.
- Messa, G. V., and Malavasi, S. (2015) Improvements in the numerical prediction of fully-suspended slurry flow in horizontal pipes, *Powder Technology* **270**, 358-367.
- Miedema, S. A. (2015) A head loss model for slurry transport in the heterogeneous regime, *Ocean Engineering* **106**, 360-370.
- Miedema, S. A. (2016a) A head loss model for homogeneous slurry transport for medium sized particles, *Journal of Hydrology and Hydromechanics* **63**, 1-12.
- Miedema, S. A. (2016b) The heterogeneous to homogeneous transition for slurry flow in pipes, *Ocean Engineering* **123**, 422-431.
- Miedema, S. A. (2017) *A new approach to determine the concentration distribution in slurry transport*, Dredging Summit and Expo. 2017, USA.
- Miedema, S. A., and Ramsdell, R. C. (2015a) *Preview: Slurry Transport: Fundamentals, A Historical Overview and The Delft Head Loss and Limit Deposit Velocity Framework*, Delft University of Technology.
- Miedema, S. A., and Ramsdell, R. C. (2015b) The limit deposit velocity model, a new approach, *Journal of Hydrology and Hydromechanics* **63**, 273-286.
- Ogawa S., Umemura A., Oshima N. (1980) On the equations of fully fluidized granular materials, *Zeitschrift Für Angewandte Mathematik Und Physik Zamp* **31**, 483-493.
- Roco, M. C., and Shook, C. A. (1983) Modeling of slurry flow: the effect of particle size, *Canadian Journal of Chemical Engineering* **61**, 494-503.
- Schaeffer, D. G. (1987) Instability in the evolution equations describing incompressible granular flow, *Journal of Differential Equations* **66**, 19-50.
- Turian, R. M., and Yuan, T. (1977) Flow of slurries in pipelines, *Aiche Journal* **23**, 232-243.
- Wasp, E. J., Kenny, J. P., and Gandhi, R. L. (1977) *Solid-liquid flow slurry pipeline transportation*, Ser. Bulk Mater. Handl. USA.
- Wilson, K. C., and Addie, G. R. (1997) Coarse-particle pipeline transport: effect of particle degradation on friction, *Powder Technology* **94**, 235-238.
- Wilson, K. C., Clift, R., and Sellgren, A. (2002) Operating points for pipelines carrying concentrated heterogeneous slurries, *Powder Technology* **123**, 19-24.
- Wilson, K. C., and Sellgren, A. (2003) Interaction of particles and near-wall lift in slurry pipelines, *Journal of Hydraulic Engineering* **129**, 73-76.
- Wilson, K. C., Addie, G. R., Sellgren, A., and Clift, R. (2006) *Slurry transport using centrifugal pumps*, Springer Science+Business Media Inc., USA.

Recovering historical urban texture by parametric computing modeling"

Historical Tehran as case study

*** Saeed Dolatkah**

Department of Architecture and Project, University of Roma" Sapienza", Italy.

*Presenting author: dolatkah.1808499@studenti.uniroma1.it

Abstract

Traditional urban texture shaped by nature, was forgotten by the industrial revolution and the advancement of technology in the field of architecture. Traditional urban design processes lack of the flexibility to deal with the complexity of the community. The complexity and uncertainty of urban environments by using traditional design methods have made it difficult to make definite plans for urban design in many cases. Since the structure of environment is having ability to convert to design parameters; therefore by using method of the parametricism, it will be possible to easily change the environment information. And the result of the change and adaptation of the design at each stage will be with new data. Because of this system dynamically updated, the design is updated with each change of input by the designer. The aim of this study is to identify features and benefits of using parametricism and computing for rebuilding a newly designed urban fabric after destroying and recovering the ancient special identity of a city. Unfortunately in the city of Tehran, the capital of Iran, the traditional face of the city is destroyed by builders and most of buildings which are our monuments convert to shopping mall and towers. In this article we first talk about the traditional urban texture of Tehran then we discuss about patterns obtained from urban texture of ancient Tehran. Finally by inputting these patterns as design parameters in grasshopper plugin, that is the most widespread plugin for parametric design, we rebuild a newly designed urban fabric of Tehran.

Keywords: Traditional Texture, Parametricism, Rebuilding, Recovering, Grasshopper

Introduction

Traditionally, urban plans are developed following methodologies aimed at the production of a single layout representing a rigid, definite solution. In addition, plans are centered on the definition of tight and interdependent urban parameters that tend to reduce design to a direct formalization of such parameters. However, legislation constrains neither design nor its representation to such an extent that they forbid design flexibility. In fact, it does not impose specific representational devices, nor does it imply any specific way of designing [1]. The context of a project, as a starting point of urban design, is changing constantly in non-linear way [2]. Therefore basic inputs of a design process, may change from time to time. However, the interrelations between elements of urban form can be fixed and predefined by designers. Current urban design practices ignore urban 'processes' and 'time' [3]. Planning processes have to become more flexible to reflect the opinions and actions of a wider range of stakeholders. [4] Therefore, objective of urban design need to move from static design of a specific layout toward complex and dynamic design of generic solutions [3]. The design of plans for cities can only be improved if designers are able to address measurements of some

of the relationships between the components of cities during the design process. These measurements are called urban indicators. By calculating such measurements, designers can grasp the meaning of the changes being proposed [5]. In this research, a urban design system is designed based on parametricism in order to achieve the above objectives. The purpose of the present research, besides introduce the features of this method of designing, is using the parametric urban design process capability to recover the old historical urban texture of Tehran. Since this method is possible and harmony with any cultural and linguistic context, parameters and patterns obtained from the traditional texture of the old city of Tehran, They are algorithmically linked with each other in "Reno + Grasshopper", one of the commonly used software in the field of parametric designing. The design results in this process show that, in addition to creating a process that makes any changes and updates the output of the design quickly, the resulting form is also largely aligned with the surrounding texture and the special needs and requirements of the design.

Parametric urban design process

An urban design system should be able to:

- be applicable in every design context for any district size and be able to apply different kinds of design programs;
- be interactive and responsive providing good visualization output both in terms of design layout and associated analytical data (indicators, attributes, indexes, etc);
- be able to implement and design the main features that compose a neighborhood.

Considering these aspects we implemented. A parametric urban design system using a NURBSCAD, environment using a parametric programming interface. The CAD environment used in this work was Rhinoceros and the programming interface was Grasshopper. The system aims at designing urban plans at the neighborhood level. Parametric design as a system of parameters, that are linked each other based on algorithm relationships. The relationship between the indices allows the general model, by changing the value of each parameter, final form is changed. The dependence of algorithmic relationships and mathematical formulas makes it possible to update the project in the design process. Urban design begins with a series of inputs. These elements are placed by the designer within the site and may be changed by the designer during the design process, and by changing the key points of the output, the design will also be displaced. Some inputs allow the designer to select the type of network [3].

There are two types of inputs – geometrical inputs and data inputs. Geometrical inputs, which is divided into 4 basic types: the site (defined by polygons); the composition elements, which are subdivided into main streets (defined by lines and curves) and focal points representing the location of the neighborhood center, local squares, public buildings and city objects in general; a vertical parameter to define the maximum number of floors; a set of grid types (rectangular, radial and recursive). Each of these inputs has a set of associated parameter inputs. A main street, for instance, has the attribute street width. [6].

To understand which design has more benefits (evaluate and analyze the design at each stage of the design), a geometric model can be connected to computing the metrics of density measurement or measurement of user levels. Whenever there is a change in the indexes, the calculations will be updated, so at any moment we will be able to get information about the changes in the design. By using this process, practitioners will be able to better understand their decision outcomes. In this process, any concept that can be expressed mathematically

can be related to a geometric model. [5]. The main steps of the proposed process will be as follows.

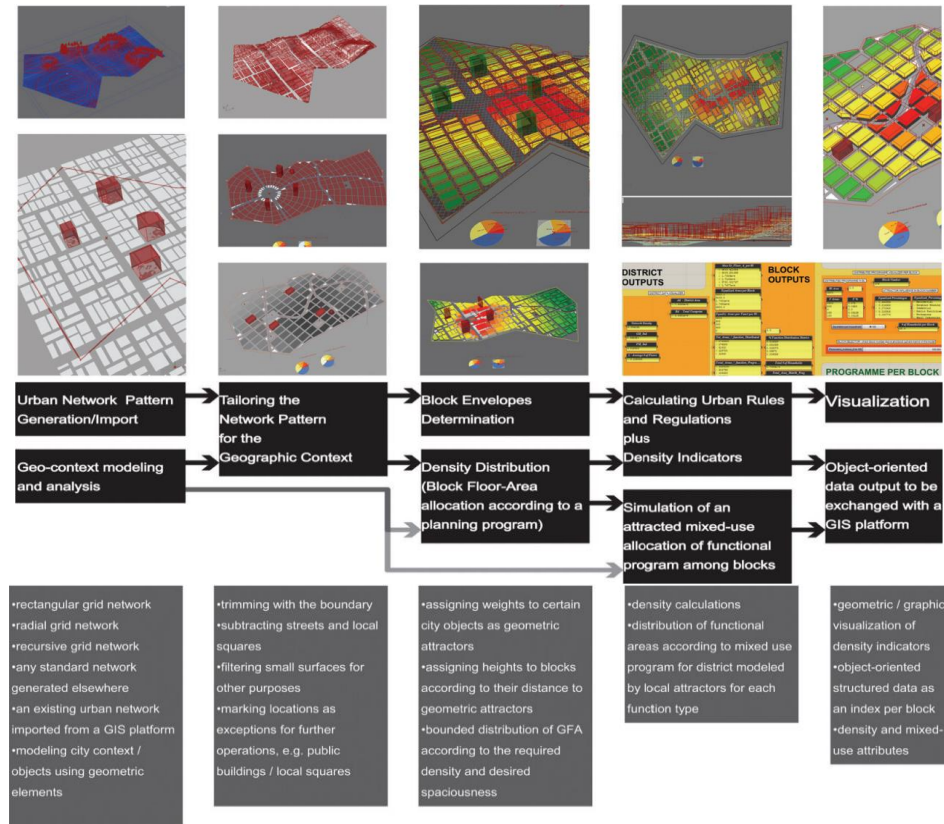


Figure 1: Parametric urban design process

Geometric features cut from the main geometry for having specific predefined conditions. The process can be replicated to create sets of geometry to which different generation rules can apply. The building height is managed by setting the maximum allowed number of floors. The number of floors is defined as a target number which is distributed through the grid as a simulation of land value. To simulate the effect of land value we defined the number of floors in a block as a function of the distance to a set of positive attractors – main square / main streets / the city center – and a set of negative attractors (repulsion effect) – site boundary. This function changes the number of floors depending on the resulting calculations [3]. And finally, the design process will be similar to figure 2-13



2: site limit or boundary



3: The main streets are designed manually by the designer



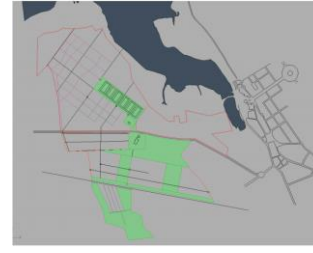
4: introducing focal points for radial grid option



5: definition of working areas



6 : main grids with main street subtraction.



7: areas to exclude pre-existences park areas and areas not suitable for construction



8: grids obtained from the previous operation



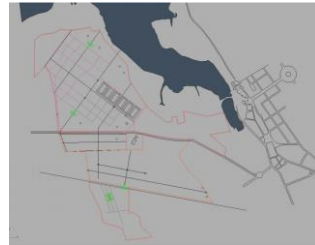
9: Subtraction of additional public spaces



10: location of public buildings or facilities



11: introduction of main squares (circle or polygon)



12: location of landmark buildings on top of axes



13: possible design

Tehran in Nasser al-Din Shah (Qajar period)

The city of Tehran has provided the legacy of earth's history as a capital of Iran more than two centuries. The historic center of Tehran (Nasser's fence) with an area of 2,250 hectares of approved historical texture, located in regions 11 and 12, has reached the official register of the Cultural Heritage Organization as the historical context. Tehran's historic texture, regardless of material values and physical values, is a messenger of spiritual values based on customs, habits and traditions. Therefore, preserving and reviving of this heritage should be an integral part of the plan to preserve and restore the historical heritage and texture of Tehran.



Figure 14: Today's map 11th and 12 regions of Tehran



Figure 15: Nasser Fence The gates of Tehran destroyed



Figure 16: Tehran Map of Qajar Time

The 11th and 12th regions of the Nasser Fence, which forms the traditional texture of Tehran and recorded in the cultural heritage, relate to the period of Nasser al-Din Shah. During this period, Tehran was expanding rapidly due to population growth. During this period, Nasser

al-Din Shah, considered Western patterns. He imitated the modernization of the Paris in the second half of the 19th century, and was equally used in Tehran [7].

These measures and reforms included: The construction of new neighborhoods in the suburbs and the creation of new streets in the old neighborhoods, making a network of wide and direct passages on the old city's texture, rebuilding of surrounding buildings with observance of the principles of mirroring, Fit the height of the walls of the streets along the street and the construction of a new urban network [8]. Also, the streets were paved, the old roads were made by using the style of the neoclassical streets of Europe, new streets were constructed in Tehran. The most important feature of them being wider, As a result, roads become straight. [9]. These wide and direct streets were places for commuting, business and Circulation. The first streets built European style by Nasser al-Din Shah were "Bob Homayoun", "Maryazkhane", "Naseriyeh", "Laleh Zar", "Cheragh Gaz" and "Baharestan" figure 17-22. Also square was made a central courtyard like many traditional squares of Iran[10]..



Figure 17: Baharestan



Figure 18: Cheragh Gaz



Figure 19: Bob Homayoun



Figure 20: Laleh Zar



Figure 21: Naseriyeh



Figure 22: Maryazkhane

Due to changes in the texture of Tehran in the period of Nasser-al-Din Shah, two types of residential construction can be made available to Tehran.

Houses that were located in the Midwest texture and were often on the sidelines of the main roads (broad alleys). These houses were made up of one floor and often two floors with tall and decorated walls and on the top of most entrance doors was small window.

Houses that were located in the western context of Tehran, in most cases were located on the main streets, broad alleys of streets and squares. In this housing style, the non-geometric form of land which were influenced by passage forms became regular and It affected illumination of houses and huge number of windows was opened to the streets. Walls, brickwork and houses were built in two floors, which was adapted from the neoclassical buildings. [7].



Figure 23: Houses of Qajar period

The destruction of Tehran's monuments

Over the past decade, Tehran's houses Instead of being listed on the country's historical monuments and protected by trustees and historic property owners, monuments have been destroyed by traders and disregarding of authorities [11]. Samples of destroyed buildings are in regions 11 and 12:



Figure 24: Delgosha house Figure 25: Qajari Sangolaj Figure 26: Sedaghat house Figure 27: jalalal Aldin

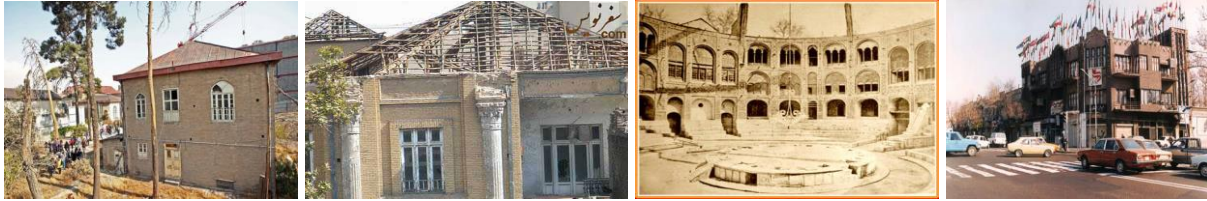


Figure 28: Dayijan Napelon Figure 29: Ameriha Figure 30: Tekye Dolat Figure 31: Parcham

Parameters for designing in this research

The design process in this research includes the following sections:

1. Providing design background
2. Put geometric inputs and information in the program and blend them with specific design patterns
3. Changing the design and manipulate the value assigned to the parameters
4. View program outputs and reset parameters

Since in this research, recovering urban texture of Tehran is done by patterns obtained from the traditional texture of ancient Tehran (the time of Nasser al-Din Shah). Therefore, it is necessary to study the traditional texture of Tehran to obtain the required parameters.

In this study, the parameters are classified into three categories:

1. Parameters related to Network
2. Parameters related to the elements of the configuration
3. Parameters related to the overall height of the texture and the skyline

Parameters related to Network

One of the key characteristics of recent studies on urban morphology is the use of networks to describe the built environment. In this perspective, the city is not seen as a collection of building blocks that may have geometrical regularities, ultimately architectural styles, but a network of interconnected open spaces created by those blocks – the urban grid [12]. The organic grid pattern should be regarded as an emergent pattern and therefore we did not use it as a designing pattern [3]. Therefore, we can use two key strategies: continuation (and not interruption) of the old grid; and following the organic lines of natural elements and environment. In contrasts are the modern master plans of the city[13].

Parameters related to the elements of the configuration

- Main streets: The width of the paths and their enclosures are the most important parameters associated with the main streets
- Focal points: The intersection of the main and secondary routes determines the position of the public spaces and sign elements and specific uses in the traditional

context. These elements include the entrance of an indoor house. We can also see a specific user in these areas like religious elements and etc. [14].

Parameters related to the overall height of the texture and the skyline

As stated above, the overall height of buildings was high, in order to imitate the neoclassical buildings and was built in one or two floors.

Design area and reason for choosing it

The suggested neighborhood for designing with the aim of recovering the old texture by using the parameters obtained from the Tehran (Nasser al-Din Shah) period is “Helel Ahmar” neighborhood where placed in the southwest of the 11th area(Nasseri fence). The area of the helel ahmar is 80 km2, formerly called Shiro Khorshid Street, which was renamed after the revolution. This neighborhood has been formed by historical monument existed in the past, such as the Qazvin Gate (Photo No.), Shahre No and Gomrok. Unfortunately all of these monuments are destroyed today. this neighborhood because of being in worn texture, is chosen for designing in this project [15].



Figure 32: Design area (left picture is this area in Nasser-al-din shah period, was garden, the middle and right pictures are today's (helal ahmar neighborhood))

The following, features of the traditional context of old Tehran In the form of parametric design parameters have been introduced.

Findings from comparative studies of parametric design criteria and old texture of Tehran for designing Helal ahmar neighborhood

- Check the elements related to the range of the site
- Check the type of network
- Check elements of the configuration
- Checking the total height of the texture and the skyline

Check the elements related to the range of the site

- **Investigating the elements determining the general neighborhood**

In old Tehran , The neighborhood area must be clear and the neighborhood had at least one element of identity. Like Helal ahmar neighborhood; identity element: Qazvin gate in Nasser al-Din Shah Period and today's: Razi cultural center



Figure 33: Qazvin gate



Figure 34: Razi cultural center

Check the type of network

- **The general pattern of the communication network**

The general pattern of network in that period was organic; the orientation of the units was affected by the form of passages and streets which came back to previous period of Nasser Alddin shah.



Figure 35 Tehran Map of Nasser alin shah Time

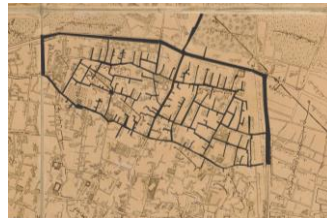


Figure 36 pattern of Paths of Tehran in Qajar

- **The position of the main and secondary communication paths and the effect of the main pathways in the overall context of the region**

In that period the main tracks connected the index points to one another, such as the connection of the Tehran old bazar to Qazvin gate in Nasser Aldin shah period



Figure 37: connection of the Tehran old bazar to Qazvin gate

- **Overall network orientation**

Overall network orientation was north- south

Check elements of the configuration

- **Geometry and location of houses**

The geometry of houses that asymmetrical quadrangle with the inner yard, was squeezed together. Houses were located on the main streets of the main streets - the broad alleys of streets and squares.



Figure 38 Compact texture of ancient Tehran

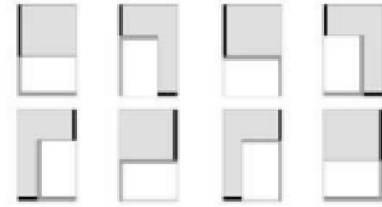


Figure 39 samples of geometry of houses



Figure 40 sample of houses with yard belongs Nasser Aldin shah time

- **Use parametric geometry in different parts of building**

Geometry is the basic knowledge of studying, measuring, and searching for the relation between forms, masses and spaces, and one of the most basic infrastructures of Islamic architecture. Material is valuable with helping of geometry and calculus it Creates a sacred space that the presence of God is felt in every corner of it. Geometric indices in Iranian Islamic architecture include semantic and structure. In other words, geometric patterns in Islamic art are used in the elements of the instruments and decorative forms. These patterns are derived from nature[16]. Figure 41 show parametric geometry in Sepahsalar mosque , one of the largest mosque in Tehran from Nasser aldin shah period.

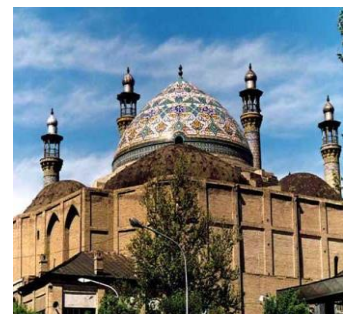


Figure41 Parametric geometry used in the interior and face of Sepahsalar mosque

- **The location of public spaces in local areas**

The intersection of the main and secondary passages was in the form of larger-scale spaces with landscaping and Courtyard (squares). Along the square, the (main core) buildings with higher altitudes and with the distance from the field, the buildings will be lowered.

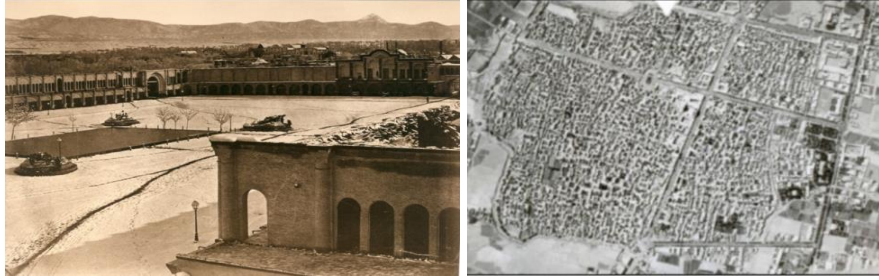


Figure 42: Courtyard (squares)

- **Investigating the effect of signs on public spaces and main routes: neighborhood structure and height of adjacent buildings**

Overall height of buildings was high and number of floors in the whole texture was one or two.

Checking the total height of the texture and the skyline

- **Investigating the density in relation to public spaces**

Significant and important elements were located around the public spaces and were visible in the skyline of the city. Such as Qazvin gate in Naeser al-Din Shah Period and today's: Razi cultural center

Designing Helal Ahmar neighborhood process

- **First step of designing (routes and type of network)**

To start designing, two factors are important: the main routes and type of network. In order to match with texture of Tehran in that period, the main lines must be follow the pattern of paths in Nassir al-Din Shah Period (Figure 36). This step is designed manually. The sub-routes are designed in such a way that the area has a combination of short and long paths and a large number of T-shaped and X shaped intersections and impasse. In the next step, it is possible to add paths or change the location and shape of the paths because with each change, the system will update the output data. In addition, as mentioned the intersection of the main and secondary lines creates cores. These cores can be larger-scale spaces with landscaping and Courtyard (squares).

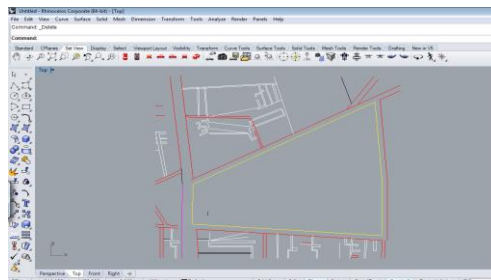


Figure43 :site limit or boundary in Rhino

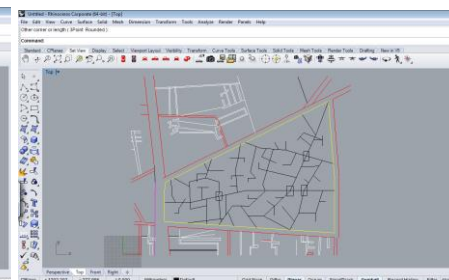


Figure44 : The main streets and cores are designed Manually by following pattern of Paths of Tehran in Qajar Time

After this process, completing the main and secondary routes, impacts and squares, are started by designer. Designer by using grasshopper plugin can add new routes in order to match with contemporary routes of Tehran.

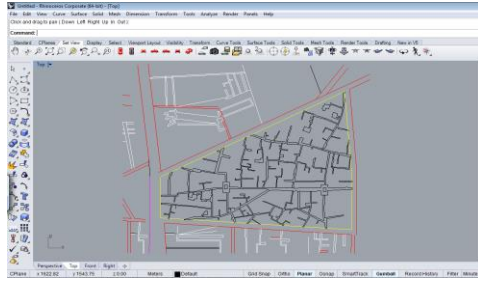


Figure 45 complete the main and secondary routes, impacts And squares

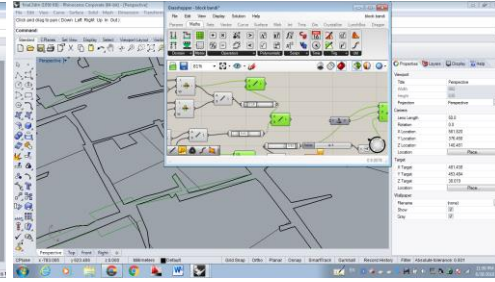


Figure 46 sample of algorithm for adding new routes in order to match with Contemporary routes of Tehran in grasshopper

- Second step of designing (creating designing area (large block) by grasshopper)

In this design we have 39 large blocks (39 items) for designing which each of them are divided to small block in terms of geometry of houses in (Nasser Aldin shah period), for designing residential and commercial place and etc.

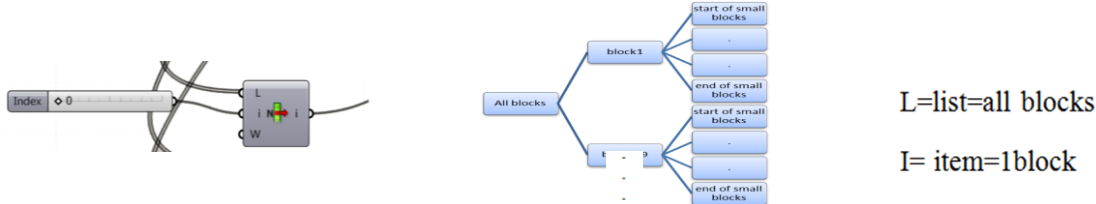


Figure 47 division blocks diagram

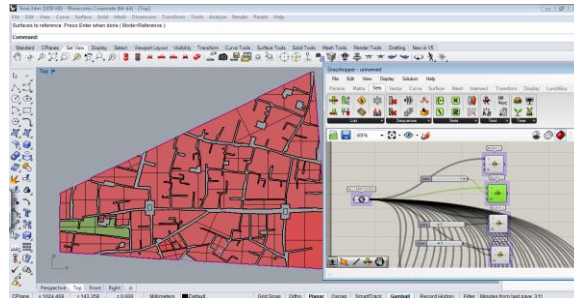


Figure 48 Creating large blocks by grasshopper (39 blocks)

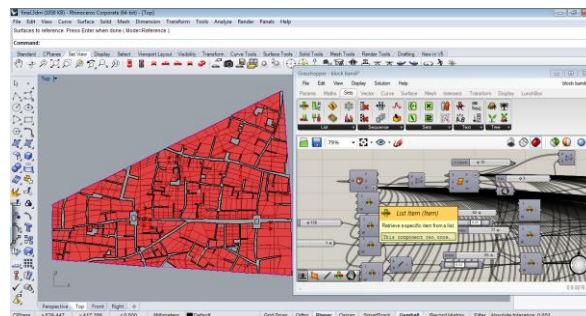


Figure 49 (39 blocks=39 items) each block is divided to small block in terms of the geometry (shape of residential houses in Nasser Aldin shah period

- Third step of designing

In this research, a parametric design has been used to generate three-dimensional patterns that can be changed during the process. One of the most important parameters is the height of buildings and density. In the parametric system, the height factor is considered as a

changeable parameter, which varies in the direction of the distances or proximity of the main nodes.

Due to the fluidity of the density in the design system, several types can be considered for the system. In this system, according to population density, the number of core and density changes simultaneously. As the population increases; we need to increase the height and the public spaces and service core in the system. Thus, by changing these two different things, you can create different types of system that can be modified in any of the following ways.

- Number of blocks
- Block sizes according to the surrounding paths
- Width and enclosure of paths
- The shape and size of the core
- Spacious public spaces

In all these species, with increasing distance from the core of the neighborhood, the density decreases with respect to the designated area. For example, for a population density of 90-120 per hectare, one of the proposed species can be single-core. In this project, a main core is considered as the square, which is considered as the old squares of Tehran during the time of Nasser al-Din Shah with the pattern of the central courtyard.

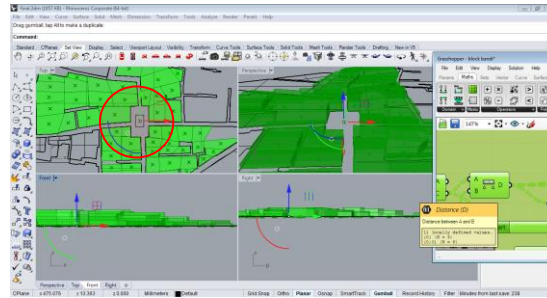


Figure 50 single core type, population density of 90-120 per hectare, Determine the range of altitude changes according to the distance from the core: with increasing distance from the core of the neighborhood, the density decreases

• Forth step of designing

After the third step and determined altitude of building in terms distance of core the position of the elements of the public and specific spaces and important uses is determined. Because of dividing blocks in small one we can choose each of them for different places. Some of blocks are chosen for residential place some of them for commercial, bank, mosque and landscaping. From the findings of the old urban texture of Tehran during Nasser al-Din Shah's time, it follows that residential blocks are located on the edge of the secondary or main passageways and the broad alleys, and main core as a square is built as a central courtyard. And round the square where placed public and commercial fields such as the bank and etc.

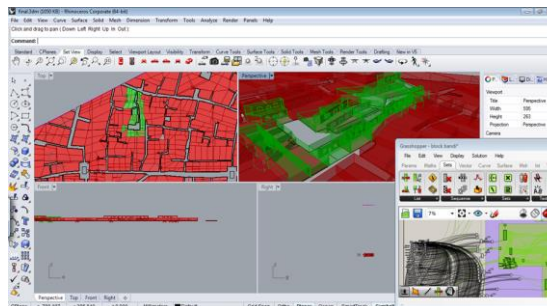


Figure 51: designing Alley, choosing one large block near Paths and some of small blocks (29 items) on it and designing them as houses by algorithms in grasshopper

In this figure(51) we can observe that each block can be designed for different Uses . The size of blocks the number of floors (one or two floor), the altitude of house, the number of windows, doors can be controlled and rehanged in every stage of designing by algorithms which designed by designer.

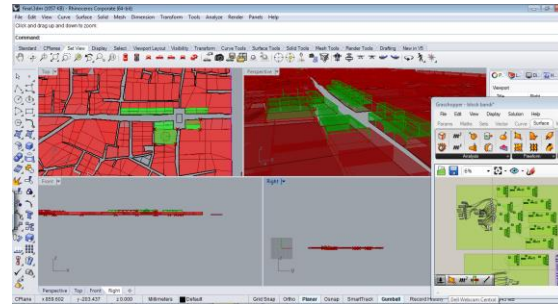


Figure 52: designing commercial fields such as the bank mosque and ..., around core, choosing small blocks (items) from large block around core (square)

As I mentioned in that time round the square where placed public and commercial fields such as the bank, mosque and etc.

- **Fifth step of designing (geometric patterns of buildings)**

At this stage of the design, materials and patterns used at the time of Nasser al-Din Shah, is designed. The dominant material on the buildings was brick and the geometric patterns used in the face of typical buildings such as mosques.

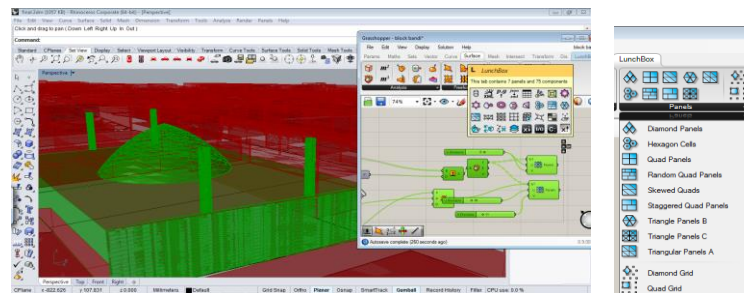


Figure 53 creating geometric pattern (brick and triangle pattern) on face of mosque by using lunch box panel in grasshopper

- **Final step of designing**

After design each block all of them are beaked (convert to mesh object in Rhinoceros).

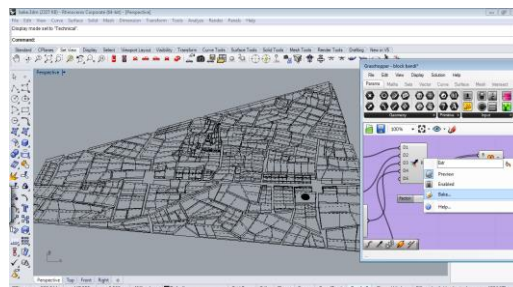


Figure 54 converting blocks designed by grasshopper to mesh in Rhino for rendering

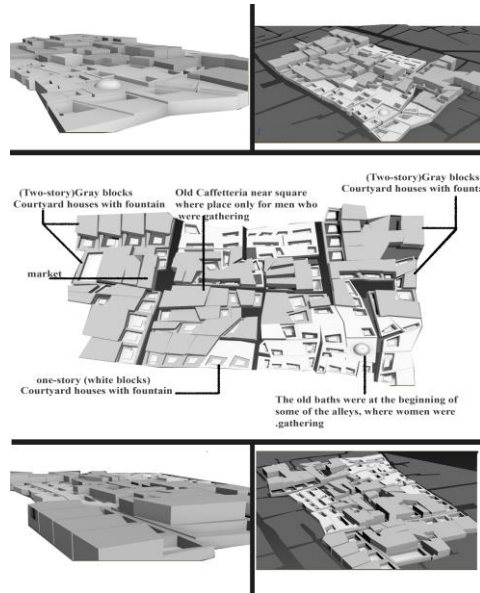


Figure 55: Final design in Rhinoceros and Grasshopper plugin

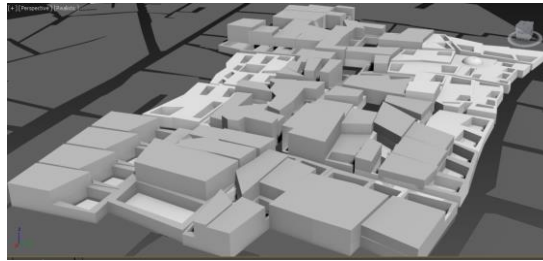


Figure 56 perspective of one part of Helal Ahmar neighborhood

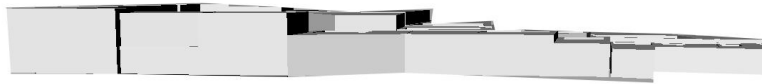


Figure 57 skyline in terms of distance of main core (square)

Using this design system makes the overall image and coordination of the limits of height and density in range. In the final design, a series of factors such as the distance from the core, the encirclement of routes and commercial buildings around the square affect the density. Therefore, it can be said that the mind of the designer in the traditional design method is not capable of creating such an image of the design scope. However, in the above method, a number of factors are affected, and in addition, by increasing the density of the future, the design can be updated. In addition to this, the system's ability to make data changes and update the layout in the following cases creates a flexible design. Change the value of the parameters associated with each of the components of the design such as:

- Change the number and dimensions of blocks in each range
- Redirect the main and the secondary paths
- Change the size and shape of public spaces and specific uses
- Ability to redefine the maximum allowable height in the range of change
- Enclosure of paths and height of adjacent buildings
- Change the location and shape of the particular user
- **The resulting design review from the parametric design process**

The form created in the process is largely In accordance with the geometry of the time frame of Nasser al-Din Shah, because main routes have been created along the paths of that time and as wide as them. Houses are located on one or two floors alongside the main roads. it is possible to create new blocks at each stage. Commercial spaces are located near the main core, such as those of that time and it is possible to control the distance between the core and the surrounding blocks. The factor of the height of the surrounding buildings, the distance to the public spaces, as well as the confinement of the routes, has been effective in determining the height of the buildings, and the effect of each of the above parameters is determined by the designer so According to the elevations of Nasser al-Din Shah's buildings, Along the square, the (main core) buildings with higher altitudes, and with the distance from the field, the buildings will be lowered. The overall height is in line with the texture of that time. The changeability of this process helps the designer to be able to temporarily move the main routes and important buildings as well as the elevation of buildings. Therefore, the advantage of the process used can be the variability of this amount to match the texture of that time.

Conclusions

From studies obtained by using parametric design, is deduced that the parametric design system is very dynamic and interactive. From changing design parameters and with the value assigned to the parameters, create different outputs for the design. Due to the obtaining of parameters from Tehran's traditional texture, the final form is largely in accord with the geometry of the organic texture and the texture of the old Tehran (Nasereddin Shah's time). Since the factor of the height of the buildings, the factors of distance to the public spaces and the enclosure of the paths were effective in determining the height of the buildings, the altitude change in the sky line was slowly taken. Finally, the capability of this process helps the designer to be able to relocate the position of the main routes and their important altitude elevations. As a result , in the traditional design system it is not possible to consider the impact of several factors simultaneously such that the system has the ability to change the design with regard to altitude changes. Thus, using this method and taking into account the urban elements in the traditional context (such as the distances of the neighborhood node, the texture of the paths and the placement of nodes and specific uses), the new texture will have an coordinate identity with the old Tehran texture and will have the same characteristics of ancient Tehran.

References

- [1] Duarte,Jose P,& Beirao,Jose (2011) Towards a Methodology for Flexible Urban Design: Designing with Urban Patterns and Shape Grammars **1**
- [2] Portugali J (2000) "Self- organization and City", Springer Verlag
- [3] Beirão, J., Nourian, P. & Mashhoodi, B., 2011. Parametric urban design: An interactive sketching system for shaping neighborhoods. In Proceedings of the Conference eCAADe 2011. eCAADe 2011. Ljubljana.225-224
- [4] Ruya,M (2009) Cohesion and Flexibility in urban design Process AMSTERDAM – Analysis by The Frame of Time and Scale(IFOU)
- [5] Beirao Jose (2012) CityMaker Designing Grammars for Urban Design
- [6] Beirão, J., Nourian, P. & Walderveen, B., 2011 Parametric 'Route Structure' Generation and Analysis an Interactive Design System Application from Urban Design
- [7] Jamat Ramezan M, Akbari Z, (2012) Texture developments in Tehran during the Qajar period Historical Sociology Volume 5 176-202
- [8] Qarib Fereydoun (1995) Tehran passages during the Qajar period. Proceedings of the Congress on Architectural and Urban History of Iran.
- [9] Sultanzadeh Hossein (1993) Entrance spaces in traditional Iranian architecture Cultural Heritage Organization

- [10] Habibi Mohsen 1999 “School of Esfahan Qajar government Tehran Style” Proceedings of the Second International Congress of Architecture and Urbanism of Iran
- [11] Website: <http://www.baladiye.com/fa/doc/report/35168>, 2018
- [12] Figueiredo L, Amorim L,(2007) “DECODING THE URBAN GRID: or why cities are neither trees nor perfect grids” istanbul
- [13] Karimi K, Motammed N, 2003, Urban planning of the city Isfahan in the past and present
- [14] Habib F, (2012) An analytical approach to domain concept in Iranian traditional cities vol 7
- [15] Website:
<http://farhangi.tehran.ir/Portals/0/Document/motaleat/%D9%87%D9%84%D8%A7%D9%84%20%D8%A7%D9%87%D9%85%D8%B1.pdf>
- [16] Nejad Ebrahimi, Ahad; Gharehbaghlou, Minou; Aliabadi, Morteza (2014). Parametric design pattern languageand geometric patterns in historical domes in Persian (ص)آ:loV architecture, Portugal

Stochastic homogenization of polymeric composites with randomly ellipsoidal reinforcement

***Damian Sokółowski¹ and †Marcin Kamiński¹**

¹Department of Structural Mechanics, Łódź University of Technology
Faculty of Civil Engineering, Architecture and Environmental Engineering
Al. Politechniki 6, 90-924 Łódź, Poland

*Presenting author: sokolowski.dmn@gmail.com

†Corresponding author: marcin.kaminski@p.lodz.pl

Abstract

The principle aim of this paper is determination of the basic probabilistic characteristics of homogenized stiffness tensor in particulate composites with ellipsoidal reinforcement and uncertain aspect ratio. A second objective is determination of susceptibility of this stiffness tensor to changes in this aspect ratio and type of reinforcing ellipsoid with spheroidal geometry, which include prolate and oblate; this is done for a wide range of ratio of main axes from 50:1 till 1:7. The homogenization problem is solved via the FEM with use of a cubic unitary single-particle Representative Volume Element (RVE) of the polymer and its 3D homogenization scheme is based on numerical determination of strain energy in the RVE under uniaxial and biaxial unitary stretch. Type and number of Finite Elements is optimized via the relative FEM error study so that its impact on probabilistic calculus is minimized and two different Finite Element types are used, i.e. 20-noded brick and 10-noded tetrahedral elements; strain energies are further used for determination of a homogenized orthotropic stiffness tensor. Probabilistic calculus is made with three independent approaches, the Iterative Stochastic Finite Element Method (ISFEM), the crude Monte-Carlo simulation and the semi-analytical method. A set of FEM solutions is available in probabilistic context upon application of the Response Function Method (RFM), where optimized polynomial approximations of the homogenized tensor components are recovered in the system MAPLE thanks to the Least Squares Method (LSM). RFM polynomial order is optimized via simultaneous maximization of correlation and minimization of LSM error and variance. A random aspect ratio is considered Gaussian, with mean value of 2 and coefficient of variation of aspect ratio smaller or equal to 0.15. We investigate numerically (1) if the resulting homogenized characteristics are also Gaussian, (2) how a change in aspect ratio affects the stiffness tensor in context of the probabilistic analysis and (3) if the homogenization increases the input uncertainty. We determine expected values, coefficients of variation, skewness and kurtosis for all available components of the effective elasticity tensor in 3D homogenization problems – all as functions of the coefficient of random dispersion for the input random parameter (aspect ratio of reinforcement).

Keywords: Stochastic Finite Element Method, probabilistic homogenization, particulate composites, ellipsoidal reinforcement.

Introduction

Homogenization technique serves for determination of the effective properties of all materials whose internal composition is characterized by some repeatable pattern. It is predominantly used in composites or advanced textiles. Properties considered by this method include mechanical [1], electrical [2], thermal [3] or coupled characteristics determined for linear or

non-linear [4][5] regime of such materials. Most of studies concern purely deterministic homogenization, but some of them also include randomness in their geometry, composition, reinforcement [6] or interphases [7]. They differ from the others at least by inclusion of a statistical scattering in the output that defines the bounds of properties for the considered materials together with some probability information. Then, they may possibly serve for determination of the material safety factors or the Probability Density Function of its properties and provide necessary safety margins into the design process. Some studies are also focused purely on determination of the Representative Volume Element (RVE) size [8] or discuss its effective generation techniques [9] so that the level of computational accuracy may be weighted by the available computational power. Homogenization of contemporary materials is only occasionally analytical as in [10]. This is because the RVEs are usually composed of several random phases with complex shapes or with multiple, randomly distributed inclusions or pores for which analytical calculus is simply unavailable. Instead, the RVE is modelled numerically, preferably with the Finite Element Method (FEM), and computed for the predefined load steps with specific boundary conditions. Only then these results are analytically converted and effective properties calculated. The FEM is leading for this application because of the ease of modelling and availability of commercial and non-commercial systems. Particulate composites are commonly homogenized on the basis of multi-particle RVEs [11] with particles of spherical as well as ellipsoidal shapes and different aspect ratios or orientation [12]. Such an approach ensures a more realistic representation of the composite than for the single-particle RVE but also includes an input randomness. This is because there is no certainty that, even most carefully chosen RVE, will exactly define the internal structure of each specimen. This is usually disregarded in current studies, where only one or at most several internal compositions are considered in the final computations and such an approach reduces their applicability to the mean characteristics with no information of dispersion at all. Lack of such works is evident and this is why randomness is introduced in this study in the form of an uncertain aspect ratio of the reinforcing particle. Additionally, properties of the matrix are also parametric with three mean values of its Young modulus and Poisson ratio. This is done on the basis of single particle RVE because of the computational intensity required for retrieval of the random characteristics, where each realization must be repeated multiple times with a variable input random parameter in the fashion very similar to the parametric design. These computations serve to check the probabilistic response of the homogenized (effective) stiffness tensor for the composite with the High Density Polyurethane (HDPU) matrix that is periodically reinforced with Carbon Black (CB) particles of ellipsoidal shape. Probabilistic calculations are made with three concurrent methods, which enables validation of each for the first four probabilistic moments and coefficients of the effective elasticity tensor.

Problem formulation and modeling details

Let us consider a periodic material composed of two phases, the matrix Ω_m and the reinforcement Ω_p in the form of ellipsoidal particles that remain both in their linear regime. Such a composite is approximated with a single particle Representative Volume Element (RVE) that permits retrieval of its mechanical properties, namely the effective stiffness tensor C_{ij}^{eff} . Once such a particle is placed centrally in this RVE, this stiffness tensor is orthotropic with 9 independent components, which are further reduced to 6 when two minor axes of the reinforcing particle have the same length (constitute a spheroid). They can be represented in the Voigt notation as C_{11}^{eff} , C_{22}^{eff} , C_{12}^{eff} , C_{13}^{eff} , C_{23}^{eff} , C_{44}^{eff} and C_{55}^{eff} , where the first pair of coefficients correspond to the uniaxial tensional stretches, second pair – to uniaxial shearing stretches and the last pair – to the biaxial tensional stretches. All the other components of the

effective stiffness tensor of such a material are null. Magnitude of this stiffness tensor for a strictly defined volume fractions of phases is lead by the aspect ratio A_r and orientation of the reinforcing particles. This study focuses on the first of these – the aspect ratio, which is considered here uncertain according to the Gaussian Probability Density Function (PDF)

$$p_b(x) = \frac{1}{\sigma(A_r)\sqrt{2\pi}} \exp\left(-\frac{(x - E[A_r])^2}{2\sigma^2(A_r)}\right); x \in \mathbb{R}. \quad (1)$$

This is done because the dispersion of geometry of reinforcement is considerable in the contemporary composites. Its basic probabilistic moments and coefficients could be computed successively from the below formulas. First, the expected value of this aspect ratio could be retrieved as [13]

$$E[A_r] = \int_{-\infty}^{+\infty} A_r p_w(x) dx = \mu(A_r) \equiv \frac{1}{M} \sum_{i=1}^M A_r^{(i)}. \quad (2)$$

It could be then used to compute variation of A_r

$$\text{Var}(A_r) = \int_{-\infty}^{+\infty} (A_r - E[A_r])^2 p_w(x) dx \equiv \frac{1}{M-1} \sum_{i=1}^M (A_r^{(i)} - E[A_r])^2, \quad (3)$$

and then converted to the coefficient of variation in a following way

$$\alpha(A_r) = \frac{\sqrt{\text{Var}(A_r)}}{E[A_r]} \quad (4)$$

With their use, the higher order coefficients of skewness β and kurtosis κ could be computed

$$\beta(A_r) = \frac{\mu_3(A_r)}{\sigma^3(A_r)} = \frac{E[A_r^3] - 3\mu(A_r)\sigma^2(A_r) - \mu^3(A_r)}{\sigma^3(A_r)}; \quad \kappa(A_r) = \frac{\mu_4(A_r)}{\sigma^4(A_r)} - 3, \quad (5)$$

Uncertainty of the aspect ratio A_r invokes uncertainty in the homogenized stiffness tensor, which is no more deterministic. Instead it could be defined by the first four probabilistic moments and coefficients, i.e. $E(C_{ij}^{eff})$, $\alpha(C_{ij}^{eff})$, $\beta(C_{ij}^{eff})$ and $\kappa(C_{ij}^{eff})$, where the expected value $E(C_{ij}^{eff})$ serves as the indicator of the most probable magnitude of C_{ij}^{eff} . They are all recovered in this study by triple computations consisting of the Iterative Stochastic Finite Element Method (ISFEM) with use of the statistically optimized 6th order [14], the crude Monte-Carlo simulation with 350 000 trials and the semi-analytical method with a direct differentiation of the characteristics from the response polynomial. An entire probabilistic calculus is based on the same response functions for each stiffness tensor coefficient, but calculations of the alternative methods are kept independent. This enables verification of the results of each of them and definition of a practical range of probabilistic moments. Discontinuous form of $E(A_r)$ and $\text{Var}(A_r)$ proposed on the right of eqn (1) and eqn (2) proves useful for the MCS, while the continuous one – in the ISFEM and the AM.

Mechanical properties of this linear composite include Young modulus and Poisson ratio. Reinforcing particle is defined by one set of parameters, $E_p = 10 \text{ GPa}$ and $\nu_p = 0.3$, while the matrix by three sets $[E_m, \nu_m] \in \{[4.4 \text{ MPa}, 0.374], [4.0 \text{ MPa}, 0.34], [3.6 \text{ MPa}, 0.306]\}$. Volume fraction of the matrix is set to 0.99 and of the particle - as 0.01 of the entire composite. Mesh of this composite (see Figure 1) is made of over 40 000 20-noded brick finite elements C3D20 with a second-order (full) stress approximation for all the 13 RVEs with modified aspect ratio of reinforcing particle. It was optimized via a FEM error study with 1st and 2nd order tetrahedral as well as hexahedral mesh of the RVE. This was done for three aspect ratios, the expected value of $A_r = \frac{2}{1}$, for the highly prolate particle with $A_r = \frac{7}{1}$ and also for the highly oblate particle with $A_r = \frac{1}{50}$ to ensure applicability of this mesh for an entire range of aspect ratios considered here. Figure 1 shows the three sections of RVEs for these three aspect ratios indicating the applied mesh and shapes of particles. Horizontal axis

x_2 on this graph corresponds to C_{22}^{eff} and the vertical one x_1 – to C_{11}^{eff} . Reinforcing particle is placed centrally to ensure an orthotropic form of the homogenized stiffness tensor C_{ij}^{eff} .

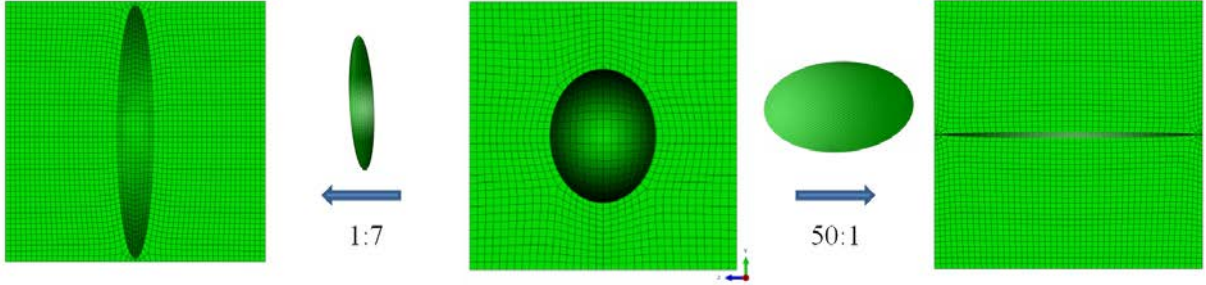


Figure 1. Discretization and range of reinforcing particle aspect ratio A_s .

Computational experiments are arranged in the following way. Firstly the FEM results are obtained for a set of 13 reinforcing particle aspect ratios $A_s \in \left\{ \frac{1}{7}, \frac{1}{6.5}, \frac{1}{6}, \frac{1}{4}, \frac{1}{2}, \frac{1}{1}, \frac{2}{1}, \frac{5}{1}, \frac{15}{1}, \frac{25}{1}, \frac{35}{1}, \frac{45}{1}, \frac{50}{1} \right\}$ and the same volume fraction of the reinforcing particle for the relevant unitary stretches. Secondly, strain elastic energies are retrieved from the RVE and they are used in homogenization to retrieve the effective orthotropic stiffness tensor C_{ij}^{eff} composed of 6 independent coefficients. A set of discrete results of C_{ij}^{eff} is then approximated with use of Response Function Method (RFM), where optimized polynomial approximations of the homogenized tensor components are recovered with the Weighted Least Squares Method (WLSM). At the end, triple probabilistic computations are performed in the symbolic computer algebra program MAPLE 2017® returning the first four probabilistic moments and coefficients of the homogenized stiffness tensor C_{ij}^{eff} . This procedure is repeated for three sets of mechanical properties of the matrix and probabilistic characteristics are plotted on the graphs with relation to the uncertain aspect ratio of the particle. Weighting type of WLSM is similar to Dirac function with weights $\mathbf{w} \in \{1, 1, 1, 1, 1, 1, 12, 1, 1, 1, 1, 1\}$ so that the expected value placed in the middle has the same weight as the remaining results. Volume fraction of the reinforcement is kept exactly the same for all the aspect ratios to exclude its influence on the results. A separate polynomial is chosen for all of the 18 sets of results composed of three groups, each with two stiffness coefficients C_{ij}^{eff} (6 in total). These groups are formed because their stiffness tensor coefficients correspond to stretches coming from the same morphology, i.e. 1. uniaxial tensional stretches (C_{11}^{eff} and C_{22}^{eff}), 2. biaxial tensional stretches (C_{12}^{eff} and C_{13}^{eff}) and 3. uniaxial shearing stretches (C_{44}^{eff} and C_{55}^{eff}). A sample response functions for C_{12}^{eff} and C_{13}^{eff} with Young modulus of the matrix $E_m = 4.0 \text{ MPa}$ and $\nu_m = 0.34$ is given below

$$F_{C_{12}^{eff}}^{WLSM} = 3.213713088792 \cdot 10^6 - 1638.14409991523 \cdot A_r + 169.16048601791 \cdot A_r^2 - 7.18211669828524 \cdot A_r^3 + 0.133618128634940 \cdot A_r^4 - 0.000907865453443940 \cdot A_r^5. \quad (6)$$

$$F_{C_{13}^{eff}}^{WLSM} = 3.20891922978851 \cdot 10^6 + 2970.83491907751 \cdot A_r + 38.3393705910314 \cdot A_r^2 - 3.25681924655126 \cdot A_r^3 + 0.0437294660806783 \cdot A_r^4 - 0.0000899175594518506 \cdot A_r^5. \quad (7)$$

An order of the approximating polynomial is optimized with simultaneous maximization of correlation and minimization of WLSM variance and error, where the precedence order is following: first correlation, second WLSM variance and third WLSM error. Polynomial approximations utilized for optimization ranged from full 1st order till full 12th order ones and

the unquestionably best fitting is ensured the 5th order polynomial. In most cases it ensured an exceptionally high correlation of at least 0.99, which proves smaller only for C_{12}^{eff} and C_{44}^{eff} but never lower than 0.89. There exists only one special case, i.e. C_{11}^{eff} , where the 7th order full polynomial is a little better than the 5th – correlation is higher on the 5th decimal number and its total error of WLSM is also smaller by 36%.

Results

A numerical example is targeted principally for a High Density Polyurethane (HDPU) reinforced with the Carbon Black and is chosen because of its relevance and applicability especially in the tire industry. Final results of the effective stiffness tensor C_{ij}^{eff} are summarized on the below graphs and include its first four probabilistic moments and coefficients, which are the expected value $E(C_{ij}^{eff})$, coefficient of variation $\alpha(C_{ij}^{eff})$, skewness $\beta(C_{ij}^{eff})$ and kurtosis $\kappa(C_{ij}^{eff})$. They are all shown as a function of the input coefficient of variation of particle aspect ratio $\alpha(A_r)$ for three different sets of mechanical properties of the linear matrix, i.e. for Young modulus of the matrix $E_m \in \{3.6, 4.0, 4.4\}$ MPa and Poisson ratio of $\nu_m \in \{0.374, 0.34, 0.306\}$, whose middle value represents properties of the High Density Polyurethane (HDPU). They are presented separately for 6 relevant stiffness tensor components placed in pairs on adjacent Figures (see for example Figure 2 and Figure 3). The graphs are ordered for an increasing probabilistic order and each characteristic is presented for three probabilistic methods. The first is the Iterative Stochastic Finite Element Method (ISFEM) with a caption of SPT, the second - Monte-Carlo simulation captioned by MCS and the third - semi-analytical approach with a caption of AM. These methodologies are differentiated by symbols on the below graphs and the sets of mechanical properties of the matrix are distinguished by colors.

The expected values $E(C_{ij}^{eff})$ are depicted on Figure 2 till Figure 7. They are generally unaffected by the increase of the input uncertainty with only a slight decrease below 1%. All three probabilistic methods return here perfectly coinciding results irrespectively to the input uncertainty and mechanical properties of the matrix. Obviously, together with an increase of these mechanical characteristics increase also all the expected values $E(C_{ij}^{eff})$. The rate of increase is much higher for the expectation (total increase of 155% – 212%) than for the stiffness tensor coefficients (total increase of 122%). This is, however, weighted by the volume fraction of the matrix in the entire composite equal to 0.99. Quite unexpectedly, the differences in relation of the homogenized stiffness tensor with the aspect ratio for the same groups of coefficients result only in marginal changes of the corresponding expectations (see for example Figure 6 and Figure 7). The highest expectations of over 8.1 kPa is reported for uniaxial tension coefficients C_{11}^{eff} and C_{22}^{eff} . The ones corresponding to biaxial stretches (C_{12}^{eff} and C_{13}^{eff}) are almost twice smaller and the shearing ones – 3.5 times smaller.

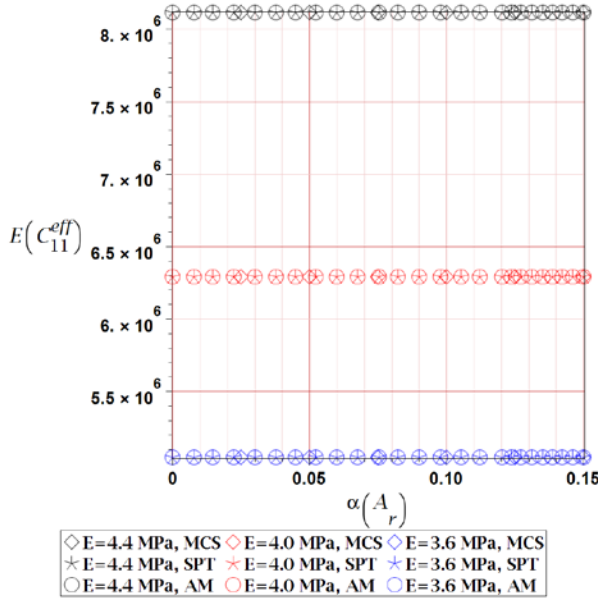


Figure 2. Expected value $E(C_{11}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

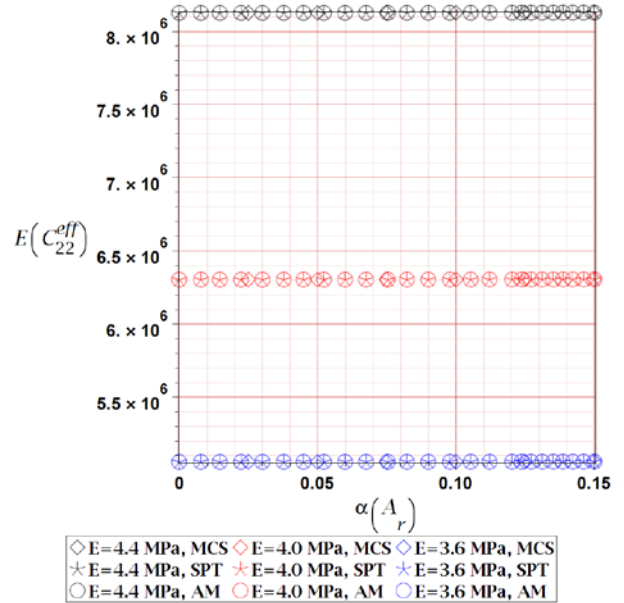


Figure 3. Expected value $E(C_{22}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

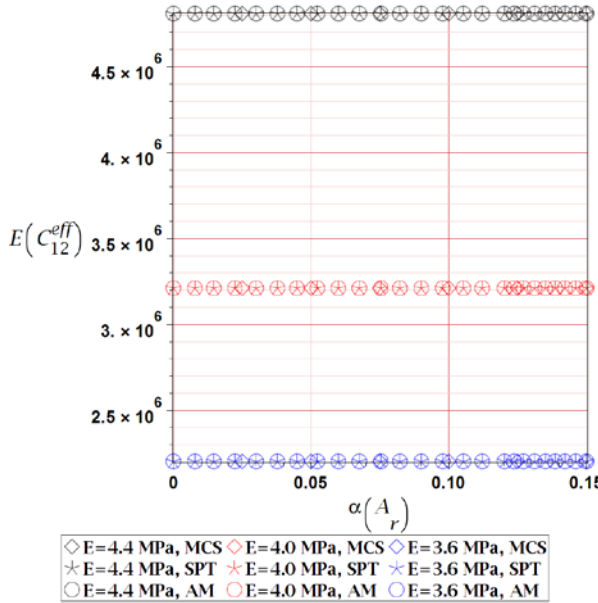


Figure 4. Expected value $E(C_{12}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

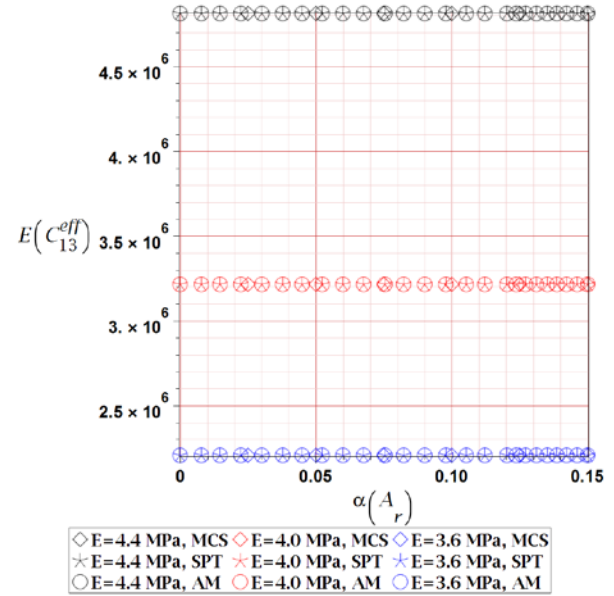


Figure 5. Expected value $E(C_{13}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

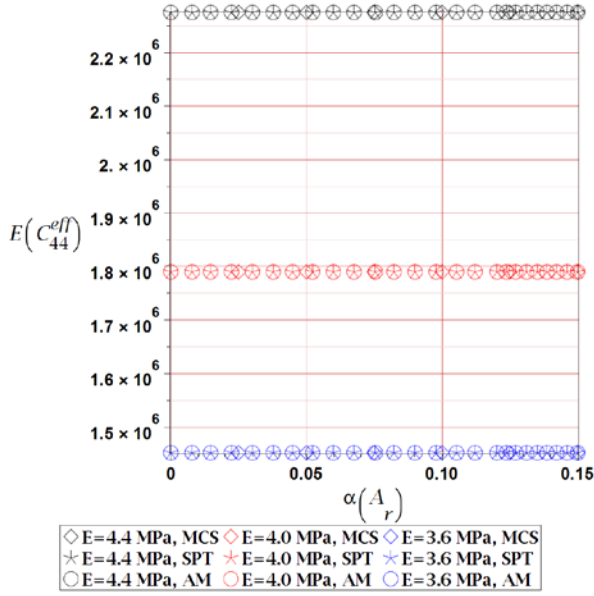


Figure 6. Expected value $E(C_{11}^{eff})$ for an increasing matrix Young modulus E w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

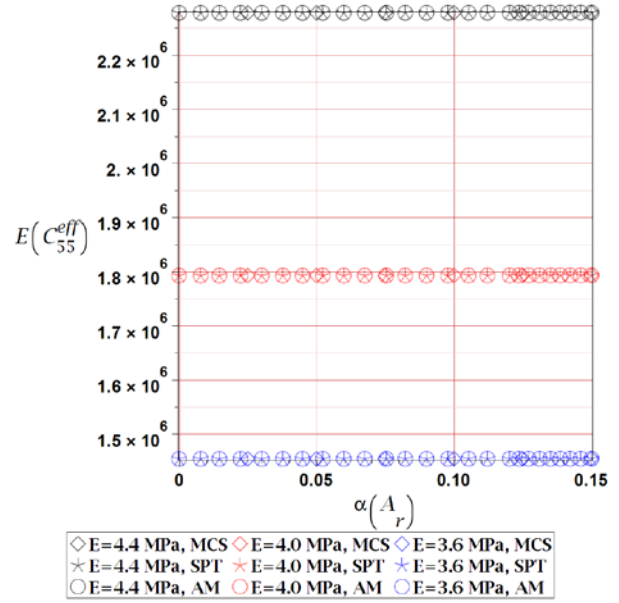


Figure 7. Expected value $E(C_{11}^{eff})$ for an increasing matrix Young modulus E w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

Coefficients of variation $\alpha(C_{ij}^{eff})$ are depicted on Figure 8 till Figure 13 and are at least 20 times smaller than the input ones. This means that homogenization considerably decreases the uncertainty. They almost linearly increase together with $\alpha(A_r)$, perfectly agree for all the three probabilistic methods and decrease together with an increase of mechanical properties of the matrix. Unlike the $E(C_{ij}^{eff})$, $\alpha(C_{ij}^{eff})$ differ significantly for the groups of components up to 60 times, which is reported for $\alpha(C_{44}^{eff})$ vs. $\alpha(C_{55}^{eff})$. The highest coefficient of variation is reported for C_{13}^{eff} and the lowest – for C_{12}^{eff} .

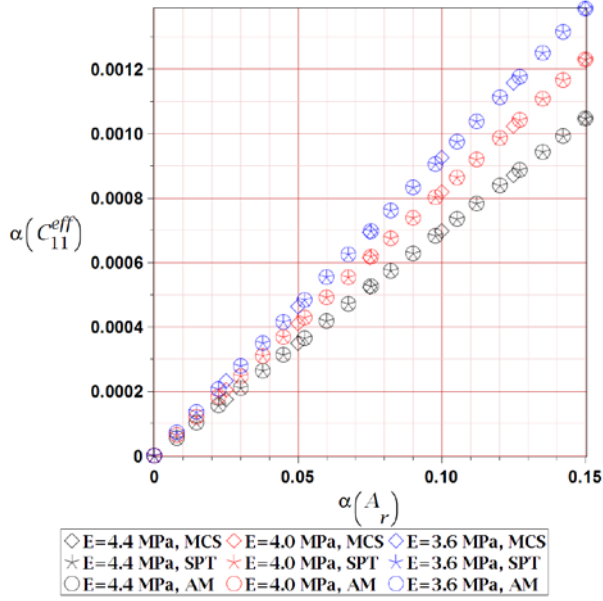


Figure 8. Coefficient of variation $\alpha(C_{11}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

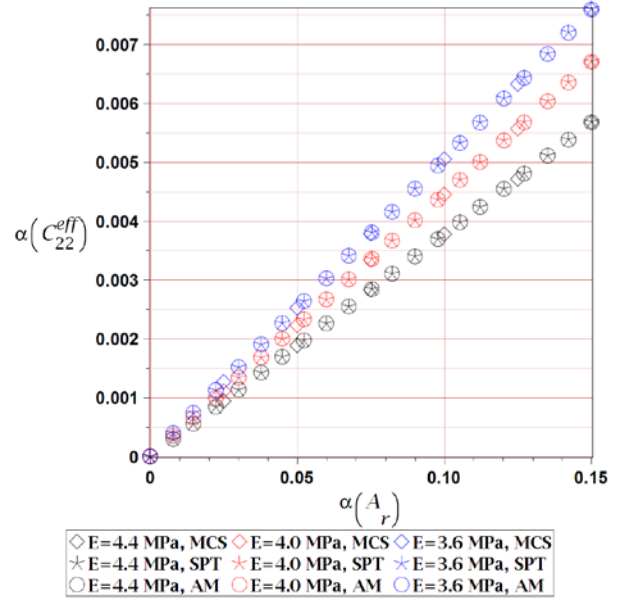


Figure 9. Coefficient of variation $\alpha(C_{22}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

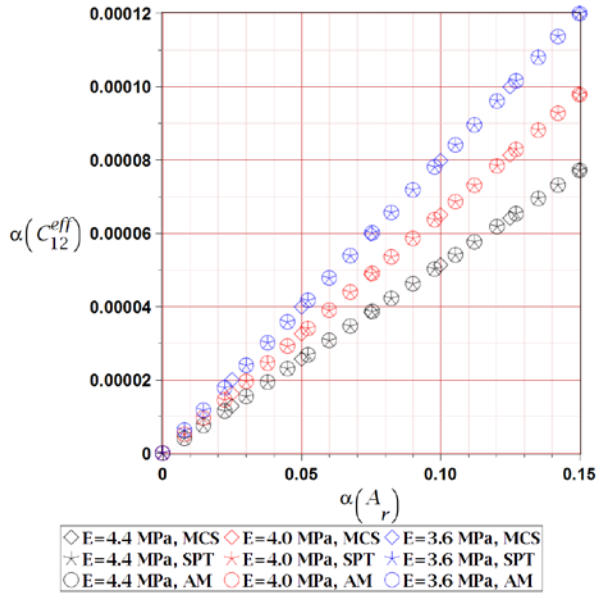


Figure 10. Coefficient of variation $\alpha(C_{12}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

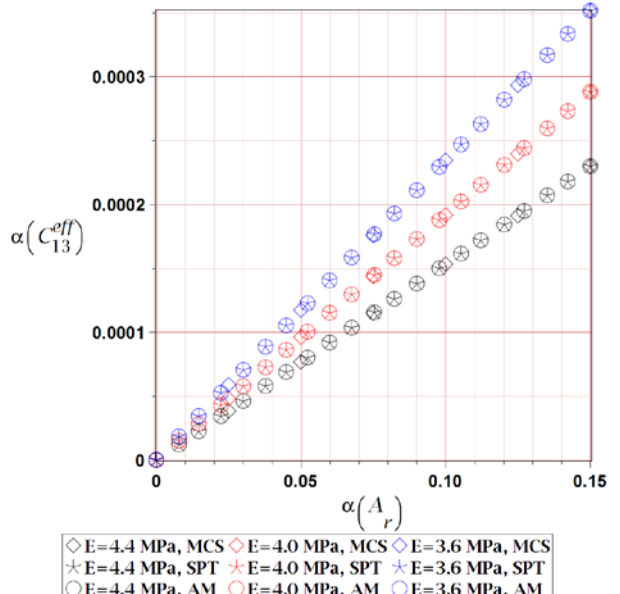


Figure 11. Coefficient of variation $\alpha(C_{13}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

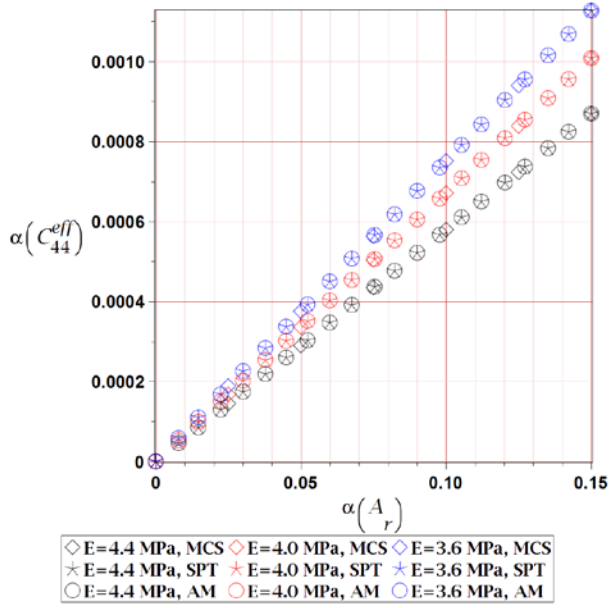


Figure 12. Coefficient of variation $\alpha(C_{44}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

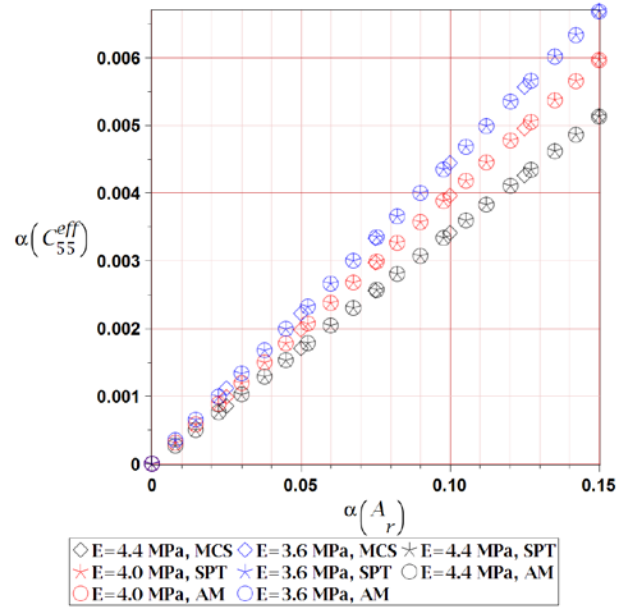


Figure 13. Coefficient of variation $\alpha(C_{55}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

Skewness of the homogenized stiffness tensor $\beta(C_{ij}^{eff})$ are presented on Figure 14 till Figure 19. It starts from 0 for a null uncertainty, almost linearly increases with an increase of $\alpha(A_r)$ and reach maximum magnitude between 0.02 and 0.3 for $\alpha(A_r) = 0$. All skewnesses are positive and in majority of them, a small concavity is visible. They differ for particular components by an order of magnitude so that in each stretch group one skewnesses is at least 10 times higher than the other (see Figure 14 and Figure 15). The SPT and AM return here perfectly the same result, while the MCS shows quite a significant scatter for all stiffness tensor components having small magnitude, i.e. C_{11}^{eff} , C_{13}^{eff} and C_{44}^{eff} . Quite importantly, a change in mechanical properties of the matrix does not have an impact on the skewness – the differences are marginal and much below a single percent.

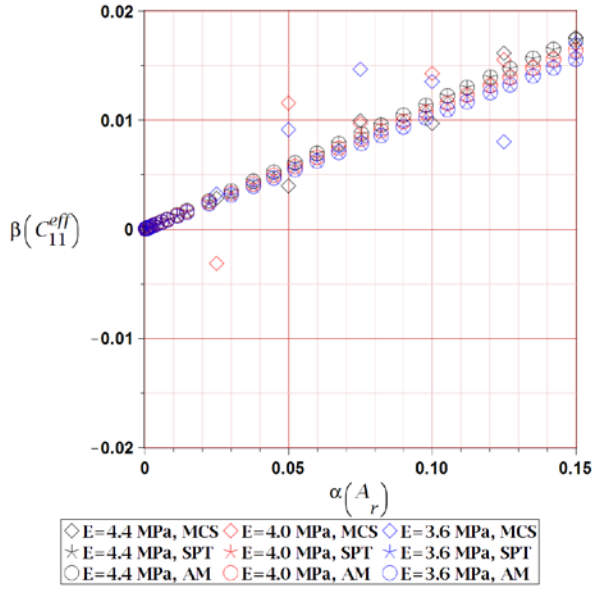


Figure 14. Skewness $\beta(C_{11}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

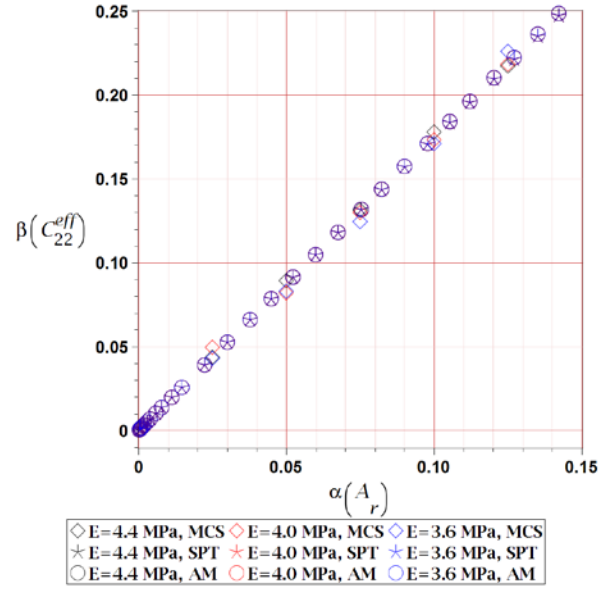


Figure 15. Skewness $\beta(C_{22}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

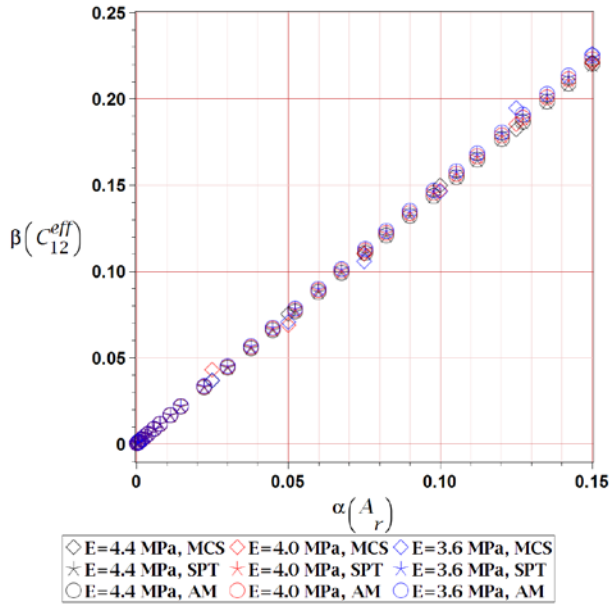


Figure 16. Skewness $\beta(C_{12}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

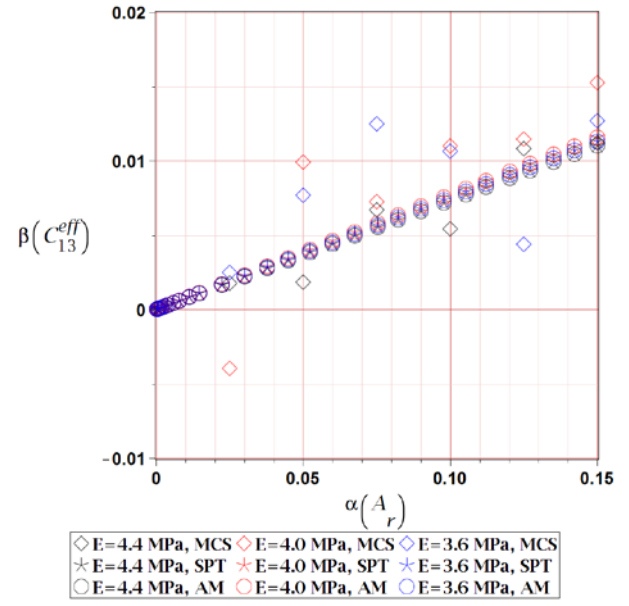


Figure 17. Skewness $\beta(C_{13}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

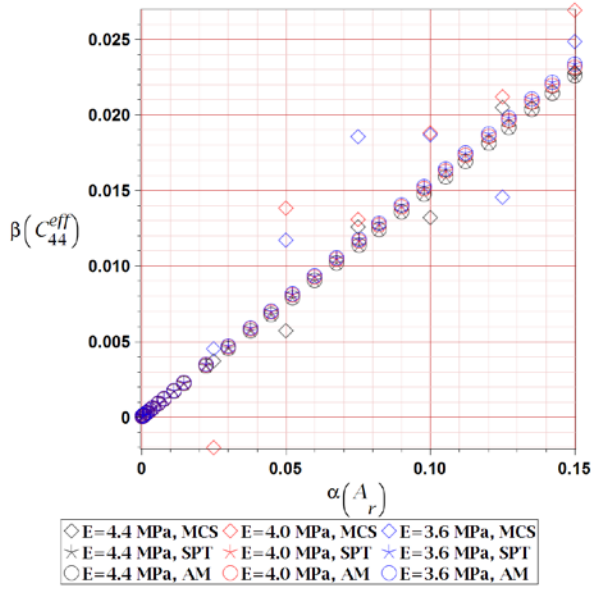


Figure 18. Skewness $\beta(C_{44}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

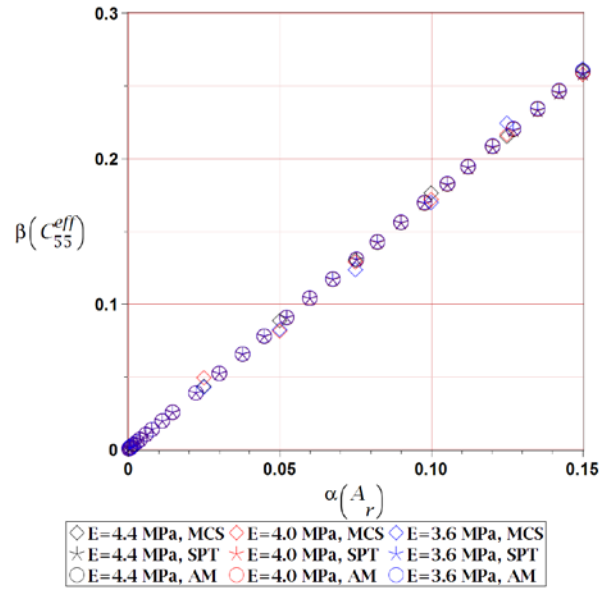


Figure 19. Skewness $\beta(C_{55}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

Graphs of kurtosis $\kappa(C_{ij}^{eff})$ are assembled in Figure 20 till Figure 25 and presented in relation to $\alpha(A_r)$. They are all positive and increasing with an increase of the input uncertainty in a convex manner. The only exception for this rule is $\kappa(C_{13}^{eff})$, where the AM returns a small convexity. Similarly to the skewness, kurtoses of coefficients from the same groups of stretches differ in magnitude several times and the MCS shows a large scatter for coefficients having small kurtoses. The three probabilistic methods lose perfect convergence for kurtosis and, by this, it is not uniquely defined with $\alpha(A_r)$. The SPT serves for the lower limit of approximation, the MCS – for the higher limit (when no MCS scatter occurs) and the AM usually stays in the middle. This discrepancy increases together with an increase of $\alpha(A_r)$ and reaches maximally 75% for C_{22}^{eff} . Difference is generally smaller between AM and SPT especially for kurtoses with small magnitude, i.e. $\kappa(C_{11}^{eff})$, $\kappa(C_{13}^{eff})$ and $\kappa(C_{44}^{eff})$ and for $\kappa(C_{11}^{eff})$ it does not exist. Tuning of mechanical properties of the matrix does not have an influence on kurtosis of the stiffness tensor. Differences are only visible for the MCS, but this method proves to be unreliable for determination of $\kappa(C_{ij}^{eff})$ because of a large scatter reported for half of coefficients.

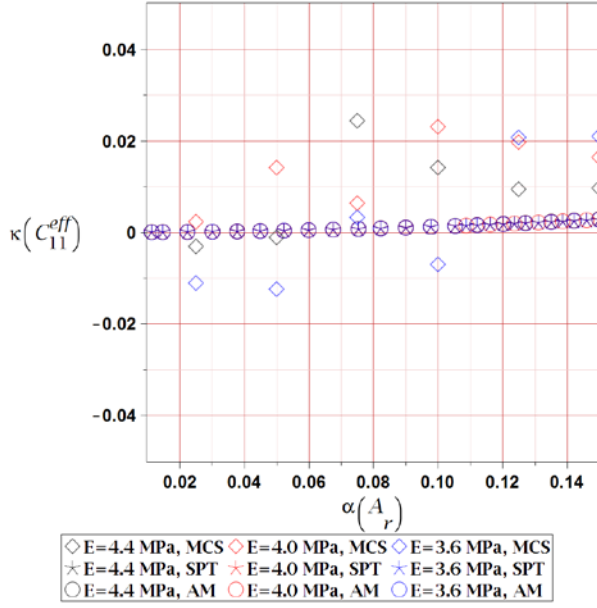


Figure 20. Kurtosis $\kappa(C_{11}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

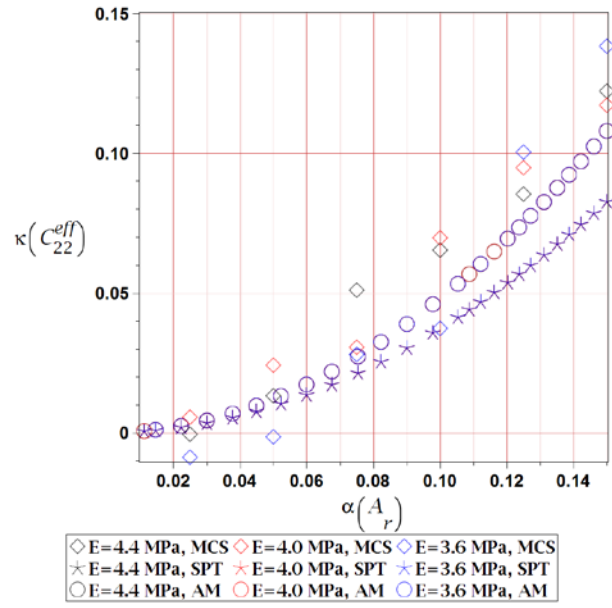


Figure 21. Kurtosis $\kappa(C_{22}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

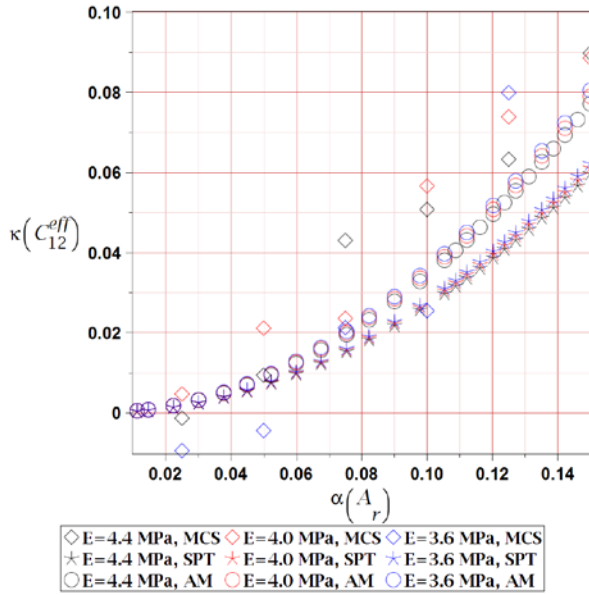


Figure 22. Kurtosis $\kappa(C_{12}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

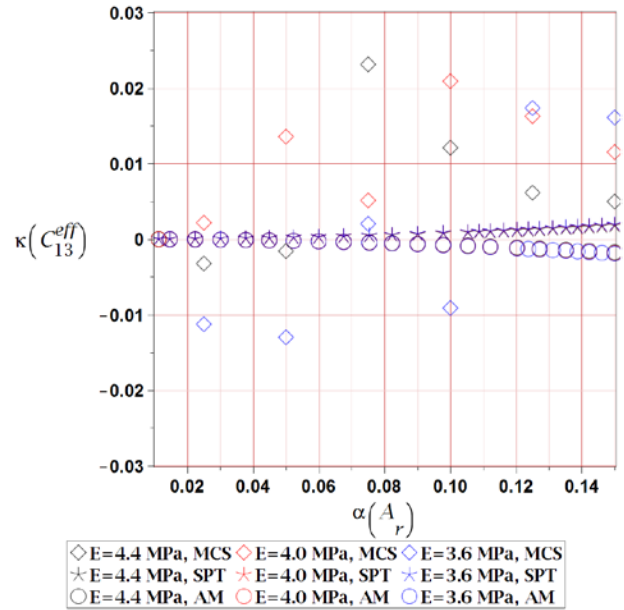


Figure 23. Kurtosis $\kappa(C_{13}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

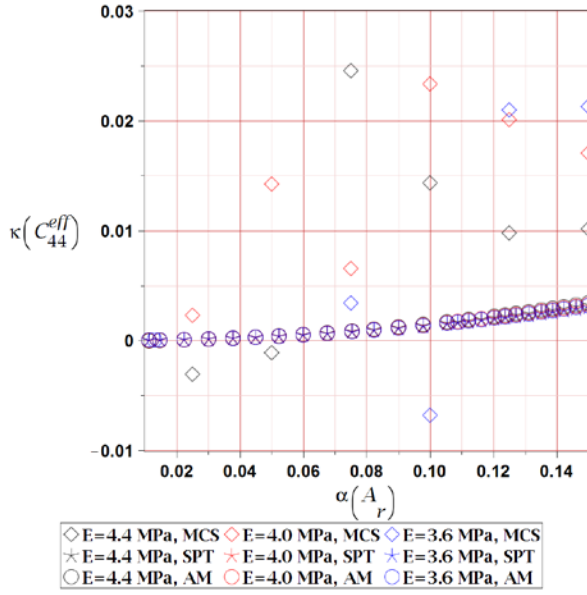


Figure 24. Kurtosis $\kappa(C_{44}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

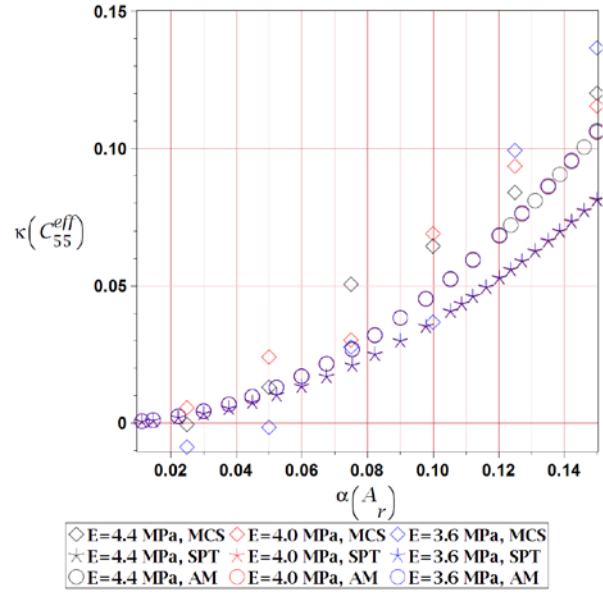


Figure 25. Kurtosis $\kappa(C_{55}^{eff})$ for an increasing matrix Young modulus E_m w.r.t. coefficient of variation of particle aspect ratio $\alpha(A_r)$.

Conclusions

This work presents a successful study of probabilistic characteristics of the homogenized random orthotropic effective stiffness tensor C_{ij}^{eff} coming from a single-particle RVE. Randomness in this study comes from an uncertain aspect ratio A_r of ellipsoidal reinforcing particle and it is based on a high range of $A_r \in \left[\frac{1}{7}, \frac{50}{1}\right]$. This is done for increasing mechanical characteristics of the matrix, i.e. Young modulus E_m and Poisson ratio ν_m in the ranges of $E_m \in \{3.6, 4.0, 4.4\} \text{ MPa}$ and $\nu_m \in \{0.374, 0.34, 0.306\}$. Middle values of these characteristics aim to represent properties of the HDPU in its linear regime. Methodology applied for retrieval of these random characteristics is triple. Firstly, a framework of the Iterative Stochastic Finite Element Method is applied, secondly a crude Monte Carlo simulation is utilized and finally the semi-analytical method is used. All these methods are based on polynomial approximations of the discrete results of stiffness tensor components coming from homogenization of RVE modelled and computed in the FEM.

This study proves that for a Gaussian aspect ratio the resulting stiffness tensor C_{ij}^{eff} is not exactly Gaussian and has a non-zero skewness and kurtosis. It also shows that the homogenization dampens the initial uncertainty so that the resulting stiffness tensor has lower variation than the input one and that the increase in initial uncertainty only marginally decreases the expected values of stiffness tensor. An increase of mechanical properties of the matrix has a visible effect only on the coefficients of variation of the homogenized stiffness tensor $\alpha(C_{ij}^{eff})$. Other probabilistic characteristics are not affected by this variable, at least within the considered limits. The coefficient of variation, skewness and kurtosis of C_{ij}^{eff} all have a major difference in magnitude for the corresponding stretches coming from the same type of boundary conditions, i.e. for uniaxial and biaxial tension as well as uniaxial shear. This is because of the different course of relation of these coefficients relative to the particle aspect ratio A_r . The randomness for uniaxial tension coefficients is higher in the axis where

prolate particle had its main axis and biaxial tension as well as uniaxial shear is much more random for axes where oblate particle had its main axes. It is worth to mention that $\alpha(C_{ij}^{eff})$, $\beta(C_{ij}^{eff})$ and $\kappa(C_{ij}^{eff})$ are all increasing with an increase of an input uncertainty so that the homogenized stiffness tensor is less Gaussian with an increase of this input uncertainty. Agreement of the three probabilistic methods is perfect for $E(C_{ij}^{eff})$ and $\alpha(C_{ij}^{eff})$ and the Monte-Carlo simulation diverges from other methods for $\beta(C_{ij}^{eff})$ and $\kappa(C_{ij}^{eff})$. This is especially visible for stiffness tensor coefficients which have a small magnitude of probabilistic characteristics, where this method has a major scatter. The ISFEM and semi-analytical approach lose convergence solely for $\kappa(C_{ij}^{eff})$, but for majority of C_{ij}^{eff} they still return a similar result, which is always smooth and without any local effects.

Further ISFEM computational studies will concern numerical simulation of the RVE with matrix and an interphase in its hyper-elastic regime for particulate composites with spherical reinforcement. An additional interesting extension of this work would be inclusion of the random interphase in-between the two current composite constituents and verification of its influence on the behavior of this composite.

References

- [1] Dinartz, F. and Sabar, H. (2017) New micromechanical modeling of the elastic behavior of composite materials with ellipsoidal reinforcements and imperfect interfaces, *International Journal of Solids and Structures* **108**, 254–262.
- [2] Zhang, Y., Dai, K., Tang, J., Ji, X. and Li, Z. (2010) Anisotropically conductive polymer composites with a selective distribution of carbon black in an in situ microfibrillar reinforced blend, *Materials Letters* **64**, 1430–1432.
- [3] Levin, V. and Markov, M. (2016) Effective thermal conductivity of micro-inhomogeneous media containing imperfectly bonded ellipsoidal inclusions, *International Journal of Engineering Science* **109**, 202–215.
- [4] Lavergne, F., Sab, K., Sanahuja, J., Bornert, M., Toulemonde, C. (2016) An approximate multiscale model for aging viscoelastic materials exhibiting time-dependent Poisson's ratio, *Cement and Concrete Research* **86**, 42–54.
- [5] Avazmohammadi, R. and Castañeda, P. (2013) Tangent Second-Order Estimates for the Large-Strain, Macroscopic Response of Particle-Reinforced Elastomers, *Journal of Elasticity* **112**, 139–183.
- [6] Kamiński, M. and Sokołowski, D. (2016) Dual probabilistic homogenization of the rubber-based composite with random carbon black particle reinforcement, *Composite Structures* **140**, 783–797.
- [7] Sokołowski, D. and Kamiński, M. (2018) Computational homogenization of carbon/polymer composites with stochastic interface defects, *Composite Structures* **183**, 434–449.
- [8] Saavas, D., Stefanou, G., Papadarakakis, M. (2016) Determination of RVE size for random composites with local volume fraction variation, *Computer Methods in Applied Mechanics and Engineering* **305**, 340–358.
- [9] Sohn, D. (2018) Periodic mesh generation and homogenization of inclusion-reinforced composites using an element-carving technique with local mesh refinement, *Composite Structures* **185**, 65–80.
- [10] Parnell, W. J., (2016) The Eshelby, Hill, Moment and Concentration Tensors for Ellipsoidal Inhomogeneities in the Newtonian Potential Problem and Linear Elastostatics, *Journal of Elasticity* **125**, 231–294.
- [11] Sabina, F. J., Gandarrilla-Pérez, C. A., Otero, J. A., Rodríguez-Ramos, R., Bravo-Castillero, J., Guinovart-Díaz R. and Valdiviezo-Mijangos, O. (2015) Dynamic homogenization for composites with embedded multioriented ellipsoidal inclusions, *International Journal of Solids and Structures* **69–70**, 121–130.
- [12] Parsaee, A., Shokrieh, M. M. and Mondali, M. (2016) A micro–macro homogenization scheme for elastic composites containing high volume fraction multi-shape inclusions, *Computational Materials Science* **121**, 217–224.
- [13] Kamiński, M. (2013) *The Stochastic Perturbation Technique for Computational Mechanics*, Wiley, Chichester.
- [14] Kamiński, M. (2015) On the dual iterative stochastic perturbation-based finite element method in solid mechanics with Gaussian uncertainties, *International Journal for Numerical Methods in Engineering* **104**, 1038–1060.

Numerical simulation of crack propagation interacting with microdefects using adaptively refined XFEM

***Z.H. Teng¹, † D.M. Liao¹, S.C. Wu², Z. B. Zhang¹, T. Chen¹, F. Sun¹**

¹ State Key Lab. of Material Processing and Die & Mould Technology, Huazhong University of Science & Technology, China

² State Key Lab. of Traction Power, Southwest Jiaotong University, China

*Presenting author: tengzihao@hust.edu.cn

†Corresponding author: liaodunming@hust.edu.cn

Abstract

The presence of the microdefects (microcracks, voids, inclusions) in the vicinity of the macrocrack tip have a significant effect on the crack propagation. Understanding the influence of crack shielding and amplification of microdefects on the SIFs (stress intensity factors) of the macrocrack is critical to accurately simulate crack propagation and predict structural life. The modeling of macrocrack growth involving multiple microdefects is inconvenient due to the different scales and the extension of cracks. In this paper, a multi-level, adaptively refined mesh near the macrocrack tip where the microdefects exist is formulated by the combination of virtual node polygonal element shape function and quadtree meshes. In the framework of XFEM (extended finite element method), the crack growth problem of different scales is unified and solved in one set of meshes. Based on the above approach, we numerically investigated the influence of different kinds of microdefects on the macrocrack propagation. The effectiveness and accuracy of the proposed method are verified by static cracking examples containing microdefects. Then, the influences of parameters such as microdefect position and size on the macrocrack SIFs and the propagating path are studied. The numerical results can provide a basis for component safety assessment.

Keywords: Adaptively mesh refinement, XFEM, Crack propagation, microdefects, virtual node polygonal element

Introduction

Many engineering structures have a large number of microdefects, such as microcracks, voids, inclusions and dislocations. Experimental studies have shown [1-5] that the presence of microdefects has a significant effect on the propagation of macrocrack. Due to the difference in the position and shape of microdefects, it may cause an amplification or shielding effect to the main crack. Accurate analysis of the influence of microdefects on the macrocrack is critical to assessing the service life of the structures. In the last century, analytical methods were used to

study the influence of microdefects on the macrocrack under certain conditions [6-10]. However, analytical methods have some limitations and it is difficult to deal with complex forms of microdefects.

In recent decades, numerical simulation methods have been widely used to deal with the interaction of the macrocrack and microdefects. There are two main methods commonly used, namely the continuum mechanics approaches and the discrete approaches [11, 12]. In the former method, the regions containing microdefects were represented as inclusions with effective mechanical constants. However, this method may lead to a decrease in computational accuracy, and in particular, the local stress concentration caused by microdefects can not be accurately modelled. While in discrete methods, microdefects are explicitly modelled, which can better analyze the interaction of the macrocrack and microdefects. Based on the widely used FEM, a lot of researches have been done on the interaction of the macrocrack and different types of microdefects [13-16]. For crack propagation problems with complex structures involving multiple microdefects, the conventional FEM is not effective since the remeshing process can not be avoided. To overcome the difficulties of remeshing, several methods have been developed, such as the boundary element method [17], meshless method [18], extended finite element method(XFEM) [19]. Among them, the XFEM has gained the most attention due to the feature that the discontinuities can be independent of the mesh.

In our previous work [20], a dynamic multi-level adaptive mesh refinement method and the corresponding VP-XFEM for crack propagation problems were proposed. Compared with the traditional XFEM and meshless methods, this method can obtain better accuracy, convergence, and computational efficiency. Based on the proposed mesh refinement method, the minimum mesh size can reach 1/100 of the original size when the refinement level reaches 7. With this method, multi-level mesh refinement is only performed on the local region where microdefects exist, and the problems at different scales can be unified into a set of mesh without additional processing. Therefore, this paper will extend this algorithm to study the influence of microdefects on macrocrack propagation.

The current research mainly focuses on the effect of microcracks on static macrocrack, little attention is paid to the interaction of growing macrocrack and microdefects [21]. Researches that comprehensively consider the effects of microcracks, voids and inclusions on macrocrack propagation are quite limited. Therefore, the VP-XFEM algorithm is used in this paper to consider the effects of microcracks, microvoids and inclusions on macrocrack propagation.

Numerical formulation

In this section, we briefly describe the composition of VPM shape functions and their continuity at hanging nodes. According to the properties of the VPM shape function, a multi-level adaptive mesh refinement method for microdefects and growing macrocrack is proposed. On this basis, the specific form of VP-XFEM is derived.

VPM shape function and its properties

Polygon elements have gained wide applications in the field of computational methods due to

their strong flexibility. The VPM uses polynomials as the approximation function based on the partition of unity (PU), which permits the direct usage of the Hammer integral and the Gauss integral as employed in standard FEM. This feature makes VPM easy to be applied to XFEM.

Assuming that a polygon domain Ω whose boundary is enclosed by n nodes $\{P_1, P_2, \dots, P_n\}$ in sequence. By introducing the centroid of the polygon P_k (virtual node), the domain can be divided into n non-overlapping virtual sub-triangles $\{T_1, T_2, \dots, T_n\}$. By considering a point P_l with the coordinate (x, y) , which is located in the virtual sub-triangle T_i , the VPM shape function has the following form

$$\Phi_l(\mathbf{x}) = W_{I,i}(\mathbf{x}) \left[(\delta_{il} + \delta_{jl}) \phi_{VP,i}(\mathbf{x}) + \phi_{VP,k}(\mathbf{x}) \phi_l(\mathbf{x}) \right] + W_{II,i}(\mathbf{x}) \phi_{VP,i}(\mathbf{x}) \quad (1)$$

where $W_{I,i}$ and $W_{II,i}$ are the weight functions of the constant strain triangular element (CST) and the least-squares method (LSM) in the i -th virtual sub-triangle; $\phi_{VP}(\mathbf{x})$ and $\phi_{VP}(\mathbf{x})$ are shape functions based on the area coordinates of CST and LSM, respectively. The specific formulations can be found in Ref. [22]. It is worth mentioning that the virtual node P_k is only used for the calculation of the shape function and will not introduce additional degrees of freedom.

VPM shape function continuity test at hanging nodes

Taking the quadrilateral elements as an example (see Fig. 1), one of the initial four large elements is divided into four small elements ①, ⑤, ⑥, and ⑦. Thus the elements ② and ③ have the hanging nodes of d and e . For the VPM, all elements are considered as polygonal elements with n nodes, while n is changeable. In this manner, there are no hanging nodes any more. Fig. 1 (b) shows the segmentation of polygonal elements when using the VPM shape functions. The handling of the elements of ② and ③ is just in the same way as other elements, except that one more virtual sub-triangle is generated. Fig. 1 (b) and (c) show the VPM shape functions at nodes d and e , respectively. It is clearly implied that the VPM shape functions possess a good continuity even near the hanging nodes.

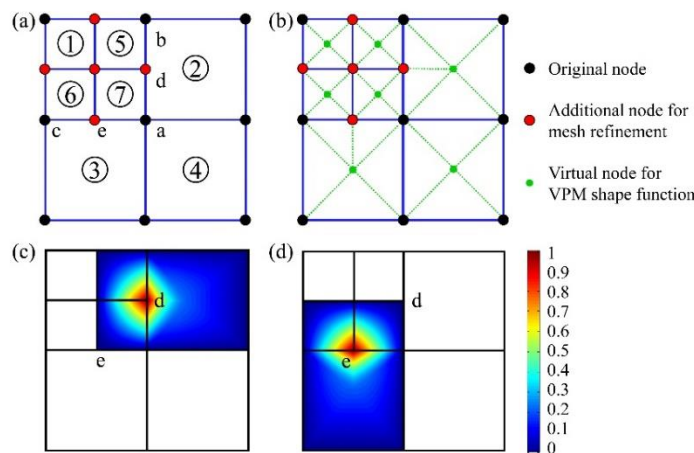


Fig. 1 Mesh refinement of quadrilateral element and the continuity test of VPM shape functions: (a) mesh refinement and the nodes distribution; (b) element division for the calculation of VPM shape functions; (c) the VPM shape function of node d ; (d) the VPM shape function of node e .

Based on the properties of VPM shape functions, an adaptive mesh refinement method for the discontinuities and the corresponding VP-XFEM are proposed by the combination of VPM and XFEM. Since VP-XFEM uses the polynomial shape functions of polygonal elements, it is distinctly different from other methods. The resultant conforming shape functions as well as special integration for sub-elements are not required.

Adaptively mesh refinement

In order to accurately simulate the effects of microdefects on the macrocrack, it is necessary to set fine meshes at the regions containing microdefects or near the crack tip of the macrocrack. Since the size of the microdefects is much smaller than the size of the macrocrack (only 1/100 or even 1/1000), when the quadtree structure is adopted, the refinement levels will reach 8 to 10. Due to the difference in the refinement levels, a lot of hanging nodes are generated in the transition zone between the coarse mesh and the fine mesh. On the other hand, the area to be refined will constantly change because of the crack growth. Therefore, it is also necessary to consider dynamic mesh refinement and coarsening. The previous section shows that VPM is well suited for dealing with non-uniform meshes with hanging nodes. All of the meshes are treated as polygon elements and handled in the same way.

Based on the properties of the VPM shape function, a dynamic adaptive multi-level mesh refinement and coarsening method is proposed in this section. Two sets of meshes are introduced to deal with the refinement process. The initial background mesh with coarse elements (called the base mesh here) is stored by the first meshing set and keeps the same throughout crack growth steps. The second meshing set is created dynamically with growing crack tips. In our previous work [\[20\]](#), the mesh refinement method for the crack propagation problem has been described in detail, but the microvoids and inclusions were not considered. Therefore, the mesh refinement method is further improved for the circular discontinuities (inclusion or hole interface).

Fig. 2 shows a single-step mesh refinement with three circular discontinuities and one main crack. In order to clearly show the refined mesh, only the level 3 refinement is shown in this example, but the process of refining level 8 and above is exactly the same. For crack propagation problem, multiple sub-steps need to be calculated, and each sub-step needs to perform a single-step mesh refinement. Fig. 3 presents the flowchart of the multi-step dynamic mesh refinement around the discontinuities with VP-XFEM. In each sub-step, meshes around the discontinuities are refined based on the initial input mesh. Then in the next sub-step, the crack information is updated, and the mesh refinement is performed again based on the new positions of the crack tips as well as the initial input mesh. At this time, the elements refined in the previous sub-step around the crack tips do not need to be refined due to the moving of the crack tips, so these elements can be maintained as the initial coarse elements. In this way, the dynamic changing of the refined area around the crack tips can be realized, which fulfills the mesh coarsening naturally. Here, a schematic diagram of the refined meshes for the calculation of crack growth containing multiple discontinuities is shown in Fig. 4.

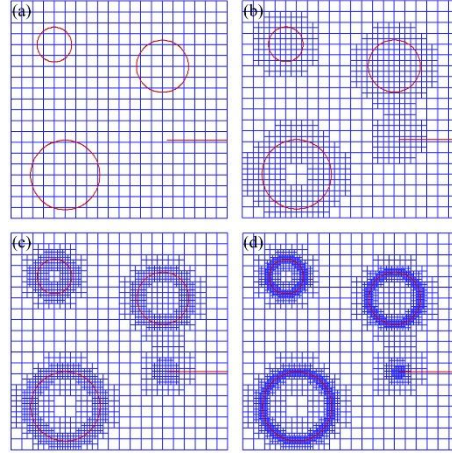


Fig. 2 The mesh refinement around the discontinuities: (a) initial mesh and the location of the discontinuities; (b) one-level refined mesh; (c) two-level refined mesh; (d) three-level refined mesh.

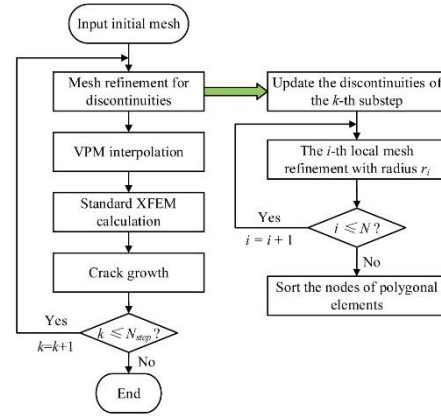


Fig. 3 The flowchart for the multi-step dynamic mesh refinement during a typical crack growth simulation in heterogeneous material.

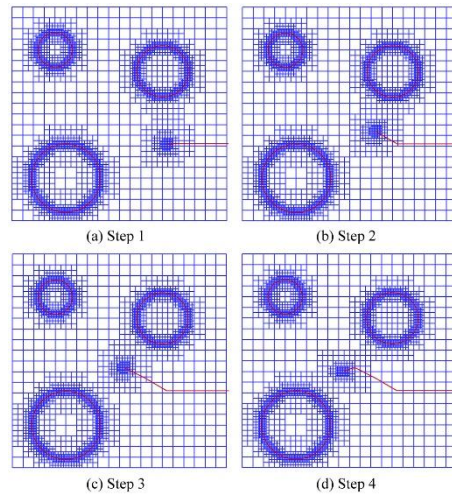


Fig. 4 The refined meshes during a typical crack growth simulation in heterogeneous material with VP-XFEM.

VP-XFEM approximation

By introducing the VPM shape functions into the XFEM displacement approximation with cracks, holes and inclusions [23-26], the VP-XFEM approximation for discontinuities in 2-D can be written as

$$\begin{aligned}
 \mathbf{u}^h(\mathbf{x}) = & \sum_{i=1}^n \Phi_i(\mathbf{x}) \mathbf{u}_i + \sum_{i \in n_r} \Phi_i(\mathbf{x}) [H(\mathbf{x}) - H(\mathbf{x}_i)] \mathbf{a}_i \\
 & + \sum_{i \in n_A} \Phi_i(\mathbf{x}) \sum_{j=1}^4 [\beta_\alpha(\mathbf{x}) - \beta_\alpha(\mathbf{x}_i)] \mathbf{b}_i^j \\
 & + \sum_{i \in n_i} \Phi_i(\mathbf{x}) [\varphi(\mathbf{x}) - \varphi(\mathbf{x}_i)] \mathbf{c}_i \\
 & + \sum_{i \in n_h} \Phi_i(\mathbf{x}) [\psi(\mathbf{x}) - \psi(\mathbf{x}_i)] \mathbf{d}_i
 \end{aligned} \tag{2}$$

The nodes associated with elements completely cut by the crack are enriched by the Heaviside function $H(\mathbf{x})$. $H(\mathbf{x})$ takes the value +1 on one side of the crack and -1 on the other side of the crack. $\beta_\alpha(\mathbf{x})$ is the crack tip enrichment function, which is used to enrich the nodes associated with elements partially cut by the crack. In the polar coordinate system r and θ of the crack tip, $\beta_\alpha(\mathbf{x})$ is defined as below

$$\beta_\alpha(r, \theta) = \left\{ \sqrt{r} \sin \frac{\theta}{2}, \quad \sqrt{r} \cos \frac{\theta}{2}, \quad \sqrt{r} \sin \frac{\theta}{2} \sin \theta, \quad \sqrt{r} \cos \frac{\theta}{2} \sin \theta \right\} \tag{3}$$

The nodes of elements cut by the inclusions and material interfaces are enriched by $\varphi(\mathbf{x})$ which is defined as [25]

$$\varphi(x) = \sum_{i \in n_i} N_i(x) |\phi_i| - \left| \sum_{i \in n_i} N_i(x) \phi_i \right| \tag{4}$$

where ϕ_i is the level set function, $N_i(x)$ is the standard FEM shape function. The nodes of elements cut by the holes are enriched by $\psi(\mathbf{x})$ which takes the value of 1 outside the hole and 0 inside the hole.

Numerical results and discussion

In this section, several numerical examples are given to study the effects of microdefects on the propagation of macrocrack. All the examples are solved by VP-XFEM. The material properties of the plate in all examples is set as $E = 50$ kPa, $\nu = 0.3$. The steady crack propagation of linear elastic material under plane strain condition is considered. In order to ensure the accuracy of the calculated SIFs at the crack tip, the radius of the integration region is four times the size of

the mesh near the crack tip when using the interaction integral.

Influence of a microcrack on the macrocrack propagation

In this section, the influence of the position and shape of microcracks on the macrocrack propagation is studied. The computational model in Fig. 5 is considered. The length of the microcrack is $l=1$ mm, the distance between the macrocrack tip and the center of the microcrack is $s=1.25$ mm. The upper and lower boundaries of the plate were subjected to a load of $\sigma=1$ kPa. Studies have shown that [11], the inclination angle φ of the microcrack and its angle θ with the macrocrack tip have a great influence on the SIFs at the macrocrack tip. With the change of θ and φ , the microcrack can either cause crack shielding or crack amplification. Therefore, the angle $\theta=60^\circ, 90^\circ, 120^\circ$ are considered, and for each case, the inclination angle φ is increased with increment of 30° , $\varphi=0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$. For the VP-XFEM, an initial mesh of 19×19 is used. The refinement level is 8 and the refinement radius α_r is 0.15. The growth of microcrack is not considered in this example. The growth increment of the macrocrack is set to $\Delta a_0=0.1$ mm, and the total number of sub-steps is 20.

The variation of SIFs at the macrocrack tip during the crack growth process is shown in Fig. 6. When θ is 60° , φ has a great influence on the value of K_I at the macrocrack tip. When θ is 90° , φ has some influence on the K_I in the early stage of crack propagation, but the influence becomes very small in the later stage. When θ is 150° , we can see that the influence of φ on K_I is really small as the curves are basically overlapping. It can be concluded that only at the front region of the macrocrack tip, the inclination angle φ of the microcrack has a great influence on the SIFs.

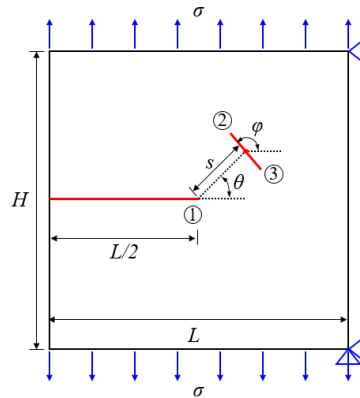


Fig. 5 Sketch of macrocrack growth problem in presence of a microcrack (Unit: mm).

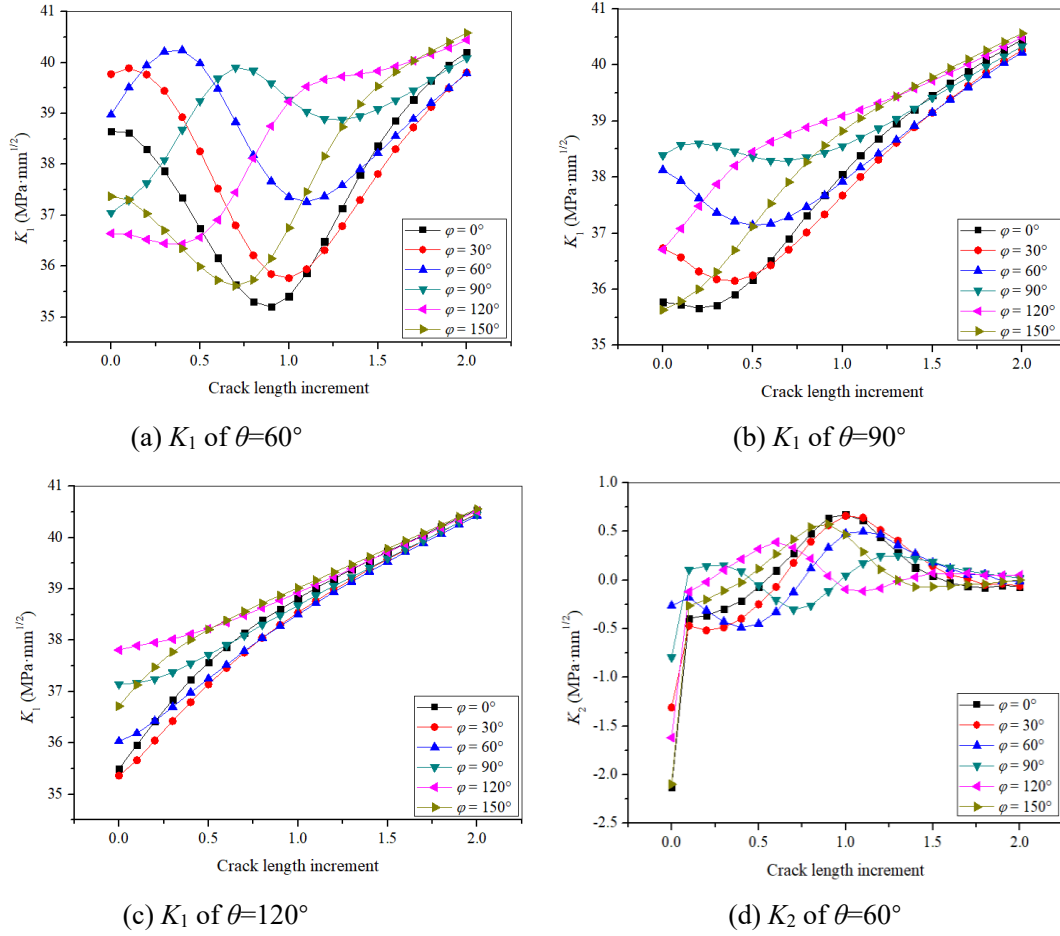


Fig. 6 The variation of SIFs at the macrocrack tip during the crack growth process in presence of a microcrack.

Influence of a microvoid on the macrocrack propagation

In order to investigate the effect of the microvoid on the macrocrack propagation, the numerical example in Fig. 8 is considered. The boundary conditions and material properties are consistent with the first example. The diameter d of the microvoid is kept constant at 2 mm. Different a values (1.86mm, 2.33mm, 3.00mm) are chosen and θ gradually increases from 0° to 90° in increments of 30° . The parameters of the VP-XFEM are set as follows: the initial mesh is 19×19 , the refinement level is 8, and the refinement radius is 0.15. The increment of crack propagation is $\Delta a_0 = 0.15$ mm, and the total number of sub-steps is 20. When the crack intersects the microvoid, the calculation is stopped.

Fig. 8 shows the variation of K_1 at the macrocrack tip during the crack growth process at different values of a . It can be seen that the value of a does not change the tendency of the microvoid to affect the macrocrack during its propagation process. While microvoid have different effects on macrocrack at different values of θ . When the values of θ are 30 and 60 degrees, the microvoid always enhances the K_1 value at the macrocrack tip. But for $\theta=90^\circ$, the microvoid causes a shielding effect. When $\theta=60^\circ$, the microvoid causes a amplification effect in the early stage of crack propagation, after reaching a certain length, the amplification effect

changes to a shielding effect.

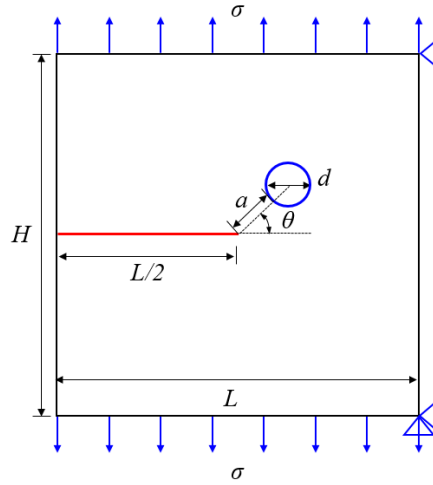


Fig. 7 Sketch of macrocrack growth problem in presence of a microvoid (Unit: mm).

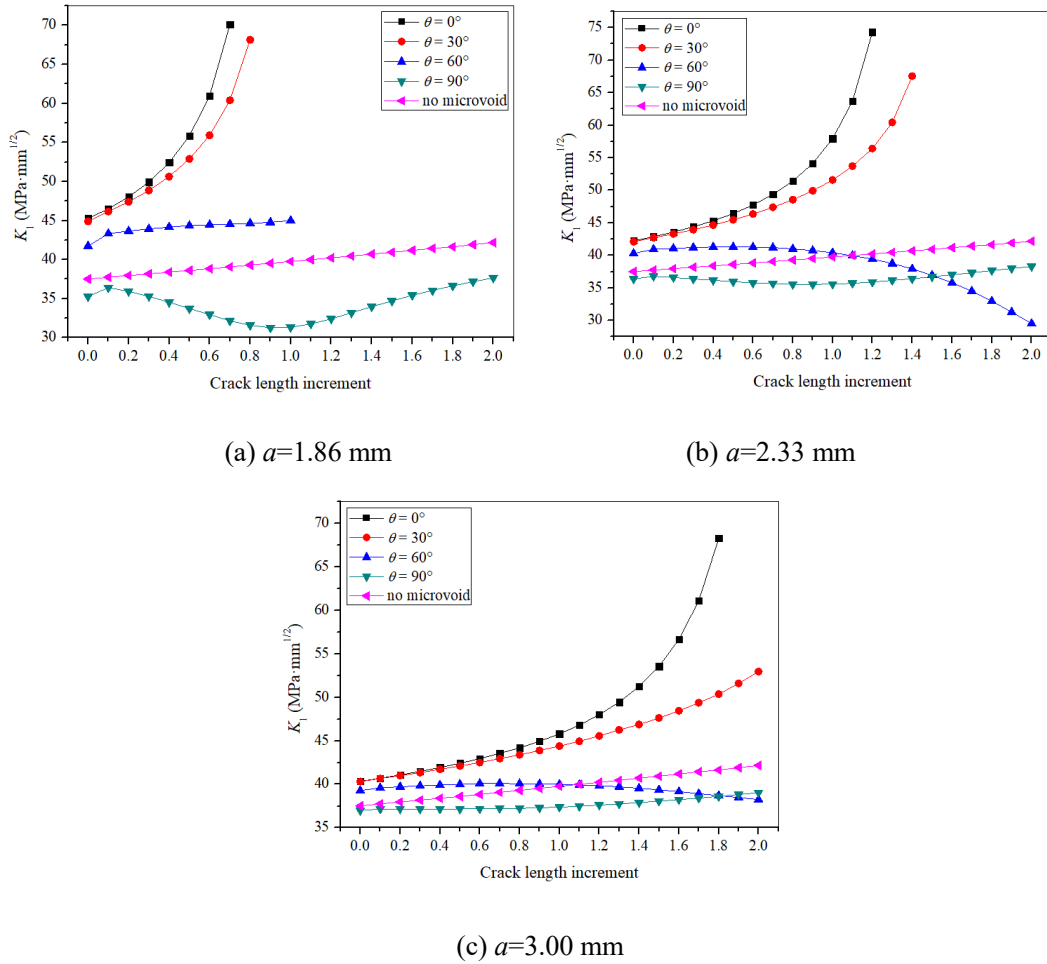


Fig. 8 The variation of K_I at the macrocrack tip during the crack growth process in presence of a microvoid.

Influence of a microinclusion on the macrocrack propagation

In order to study the influence of microinclusion on macrocrack propagation, consider the example shown in Fig. 9. The size and material of the plate, the location of the macrocrack, and the boundary conditions remain the same as in the first numerical example. A circular microinclusion of diameter d is placed near the macrocrack tip. The distance between the center of the microinclusion and the macrocrack tip is s , and the angle of the line connecting the crack tip and the center of microinclusion is θ . In order to investigate the influence of different materials and different inclusion sizes on the macrocrack propagation, the center of the microinclusion is fixed as $s=3.0$ mm, $\theta=30^\circ$. Inclusion in two sizes and two materials are considered, $d_1=2.0$ mm, $d_2=1.6$ mm, and $E_1=25$ kPa, $E_2=100$ kPa. The parameters of the VP-XFEM are set as follows: the initial mesh is 19×19 , the refinement level is 8, and the refinement radius is 0.15. The increment of crack propagation is $\Delta a_0=0.2$ mm, and the total number of sub-steps is 20.

Fig. 10 shows the variation of SIFs at the macrocrack tip during the crack growth process. It can be seen that the larger the inclusion size, the greater the effect on the macrocrack when the elastic modulus of the inclusion is the same. For inclusions of the same size, soft microinclusion causes an amplification effect to the K_1 of macrocrack in the early stage of crack propagation. When the macrocrack expands to a certain length, soft microinclusion causes an shielding effect. While the case of hard microinclusion is just the opposite. For the variation of K_2 , it shows a similar pattern. Soft microinclusion will attract the macrocrack to growth in its direction. While for hard microinclusion, the crack extends away from its direction.

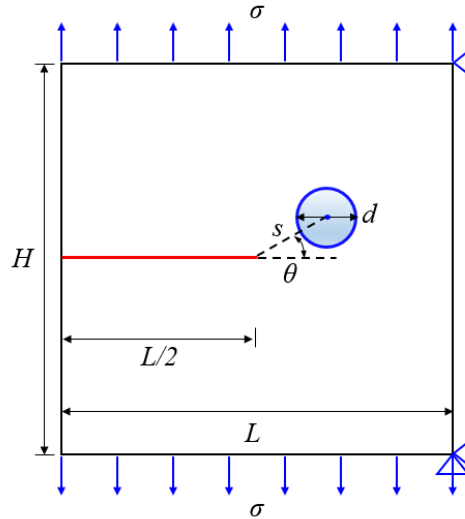


Fig. 9 Sketch of macrocrack growth problem in presence of a microinclusion (Unit: mm).

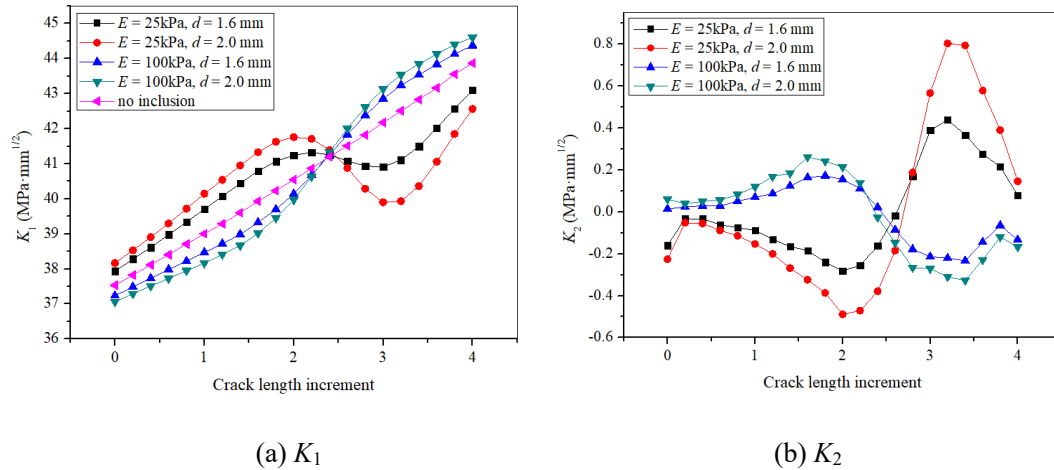


Fig. 10 The variation of SIFs at the macrocrack tip during the crack growth process in presence of a microinclusion.

Conclusions

In this paper, a multi-level, adaptively refined mesh near the macrocrack tip where the microdefects exist is formulated by the combination of virtual node polygonal element shape function and quadtree meshes. The influence of the nearby microdefects on the macrocrack propagation have been numerically investigated by using VP-XFEM. The results show that, the location and geometric parameters of the nearby microdefects have great influence on the propagation of the macrocrack. Both of the amplification and shielding effect can be seen in different kinds of microdefects with different parameters.

References

- [1] Song SH, Bae JS. Fatigue crack initiation and propagation from hole defects. *Experimental Mechanics*, 1998, 38(3): 161-166.
- [2] Tsay LW, Shan YP, Chao YH, Shu WY. The influence of porosity on the fatigue crack growth behavior of Ti-6Al-4V laser welds. *Journal of Materials Science*, 2006, 41(22): 7498-7505.
- [3] Seo KJ, Choi BH, Lee JM, Shin SM. Investigation of the mixed-mode fatigue crack growth of a hot-rolled steel plate with a circular microdefect. *International Journal of Fatigue*, 2010, 32(7): 1190-1199.
- [4] Wu SC, Yu C, Yu PS, et al. Corner fatigue cracking behavior of hybrid laser AA7020 welds by synchrotron X-ray computed microtomography. *Materials Science and Engineering: A*, 2016, 651: 604-614.
- [5] Yang RS, Xu P, Yue ZW, et al. Dynamic fracture analysis of crack-defect interaction for mode I running crack using digital dynamic caustics method. *Engineering Fracture Mechanics*, 2016, 161: 63-75.
- [6] Rose LRF. Microcrack interaction with a main crack. *International Journal of Fracture*, 1986, 31(3): 233-242.

- [7] Gong SX. An exact formulation for the microcrack-finite main crack interaction. *International Journal of Fracture*, 1994, 66: R51-R56.
- [8] Meguid SA, Wang XD. On the dynamic interaction between a microdefect and a main crack. *Proc. R. Soc. Lond. A*, 1995, 448(1934): 449-464.
- [9] Tamuzs VP, Petrova VE. Modified model of macro–microcrack interaction. *Theoretical and applied fracture mechanics*, 1999, 32(2): 111-117.
- [10] Tamuzs VP, Petrova VE. On macrocrack–microdefect interaction. *International applied mechanics*, 2002, 38(10): 1157-1177.
- [11] Petrova V, Tamuzs V, Romalis N. A survey of macro-microcrack interaction problems. *Applied Mechanics Reviews*, 2000, 53(5): 117-146.
- [12] Liu G, Zhou D, Bao Y, et al. Multiscale analysis of interaction between macro crack and microdefects by using multiscale projection method. *Theoretical and Applied Fracture Mechanics*, 2017, 90: 65-74.
- [13] Meguid S A, Gaultier P E, Gong S X. A comparison between analytical and finite element analysis of main crack-microcrack interaction. *Engineering fracture mechanics*, 1991, 38(6): 451-465.
- [14] Biner S B. A FEM analysis of crack growth in microcracking brittle solids. *Engineering fracture mechanics*, 1995, 51(4): 555-573.
- [15] Soh A K, Yang C H. Numerical modeling of interactions between a macro-crack and a cluster of micro-defects. *Engineering fracture mechanics*, 2004, 71(2): 193-217.
- [16] Li X, Li X, Jiang X. Influence of a micro-crack on the finite macro-crack. *Engineering Fracture Mechanics*, 2017, 177: 95-103.
- [17] Portela A, Aliabadi M, Rooke D. The dual boundary element method: effective implementation for crack problem. *Int. J. Numer. Meth. Eng.*, 1991, 33: 1269-1287.
- [18] Belytschko T, Lu YY, Gu L. Crack propagation by element-free Galerkin methods. *Eng. Fract. Mech.*, 1995, 51: 295-315.
- [19] Moës N, Dolbow J, Belytschko T. A finite element method for crack growth without remeshing. *Int. J. Numer. Meth. Eng.*, 1999, 46: 131-150.
- [20] Teng Z H, Sun F, Wu S C, et al. An adaptively refined XFEM with virtual node polygonal elements for dynamic crack problems. *Computational Mechanics*, 2018(1–3):1-20.
- [21] Liu G, Zhou D, Guo J, et al. Numerical simulation of fatigue crack propagation interacting with micro-defects using multiscale XFEM. *International Journal of Fatigue*, 2018, 109:70-82.
- [22] Tang XH, Wu SC, Zheng C, Zhang JH. A novel virtual node method for polygonal elements. *Applied Mathematics and Mechanics-English Edition*, 2009, 30(10): 1233-1246.
- [23] Moës N, Dolbow J, Belytschko T. A finite element method for crack growth without remeshing. *Int J Numer Meth Eng*, 1999, 46: 131-150.

- [24] Sukumar N, Chopp DL, Moes N, Belytschko T. Modeling holes and inclusions by level sets in the extended finite element method. *Comput Methods Appl Mech Eng*, 2001, 190: 6183–200.
- [25] Moës N, Cloirec M, Cartraud P, et al. A computational approach to handle complex microstructure geometries. *Computer methods in applied mechanics and engineering*, 2003, 192(28-30): 3163-3177.
- [26] Jiang S, Du C, Gu C, Chen X. XFEM analysis of the effects of voids, inclusions and other cracks on the dynamic stress intensity factor of a major crack. *Fatigue & Fracture of Engineering Materials & Structure*, 2014, 37: 866-882.

A Semi-Lagrangian Method Based on MAC and Its Application in Fluid Simulation of Casting Filling Process

*Y.L. Tang, † D.M. Liao, T. Chen, S.Y. Fan

State Key Lab. of Material Processing and Die & Mould Technology, Huazhong University of Science & Technology, China

*Presenting author: tangyulong@hust.edu.cn

†Corresponding author: liaodunming@hust.edu.cn

Abstract

In this paper, a marker-and-cell based semi-Lagrangian method, which combines the advantages of Lagrangian and Eulerian perspective, is introduced to simulate the metal flow during casting filling process. This method treats fluid in terms of a vast collection of particles with some physical properties that move around, naturally bypass the unique complex free interface tracking problem in Eulerian perspective. For the interior of the fluid, the efficient Projection method is used to solve the incompressible Navier-Stokes equations. On the other hand, the diffusion equation is solved on the background grid by mapping the information carried by the particles to the grid, thus avoiding the inconvenience of field description in Lagrangian perspective. And the level set method is introduced to describe boundary. In order to study the application of this method in the casting field, two examples are given with this method and the widely used VOF method. The results show that this method can produce sharp and discontinuous fluid interfaces, which is closer to the casting process with high filling speed.

Keywords: marker-and-cell, semi-Lagrangian method, casting filling process, fluid simulation

Introduction

In the field of casting, computational fluid dynamics is mainly used to simulate the metal filling process and analyze defects that may occur during the filling process, such as cold shut^[1] and oxide inclusion^[2]. The flow front of the molten metal plays a very important role in analyzing the filling defects. The mainstream free surface tracking methods in casting simulations use a fixed grid. Methods like VOF (Volume-of-Fluid)^[3], Level set^[4], and CLSVOF (Coupled Level Set and Volume-of-Fluid)^[5] are widely used. This kind of methods uses continuous functions to indicate the free surface, in a pure Eulerian perspective. Traditionally, the main difficulty in using these methods has been the maintenance of a sharp boundary^[6]. Methods that use separate, boundary-fitted grids for each phase^[7-9] may offer potentially the highest accuracy, are rarely used in casting simulations because of their complexity.

Another class of methods is the particle-based “meshfree” methods, such as MPS (Moving Particles Semi-implicit)^[10], SPH (Smoothed Particle Hydrodynamics)^[11], DPD (Dissipative Particle Dynamics)^[12,13]. The absence of a mesh allows Lagrangian simulations, in which the particles can move according to their own state. However, when calculating the force of particles, we have to find the location of neighboring particles. To avoid being n-body problems, some complex data structures such as the multi-grid technique should be applied. Besides, a casting is usually very complex in geometry that requires many particles. The relatively low computing efficiency makes them very difficult to be applied to casting simulation.

The MAC (marker-and-cell) method^[14] is a semi-lagrangian method that between the fixed grid and mesh-free methods. It uses a fixed grid to describe the interior of the fluid, and marker particles to indicate fluid configuration. With a structured grid used, it not only has the advantage of fixed grids in solving NS equations but also has the advantage of mesh-free methods in advancing fluid front. However, it is rarely used in casting simulations although it is very popular in the field of computational fluid dynamics. In this paper, the MAC method, coupled with the level set approach to describe the complex boundary, is introduced to simulate the casting filling process. Through the actual calculation case, the potential applications and limitations of this method in the casting filling simulation process are discussed.

Description of the Method

Outline of procedure

The governing equations for viscous incompressible flows are the continuity and the Navier-Stokes equations as follows^[15]:

$$\nabla \cdot \vec{u} = 0 \quad (1)$$

and

$$\frac{D\vec{u}}{Dt} = -\nabla\varphi + \nu\Delta\vec{u} + \vec{f} \quad (2)$$

Where $\varphi = P/\rho$ is the ratio of pressure to density, it is usually referred to as “pressure” simply, and the triangular symbol Δ represents the Laplacian. Using the efficient Projection procedure^[16], a Poisson equation for pressure can be obtained:

$$\Delta\varphi^{n+1} = \frac{\nabla \cdot \vec{u}^*}{\delta t} \quad (3)$$

$$\vec{u}^* = \vec{u}^n + \delta t \left(\nu \Delta \vec{u}^n + \vec{f}^n - \vec{u}^n \cdot \nabla \vec{u}^n \right) \quad (4)$$

The superscript “n” and “n+1” denotes the cycle number, for example, φ^{n+1} is the pressure value of cycle n+1. Solving Eq.(3) and substitute the pressure back to Eq. (2), then the velocity field of cycle n+1 is obtained:

$$\vec{u}^{n+1} = \vec{u}^* - \delta t \cdot \nabla \varphi^{n+1} \quad (5)$$

The overall procedure is based on the classical Marker-and-Cell method developed by Harlow and Welch^[14]. For a calculation cycle, there are six steps as follow:

- (1) Compute pressure. Solve the Poisson equation of pressure based on the velocity and liquid level set field.
- (2) Update velocity. Update the velocity field by substituting back the new pressure field to the Navier-Stokes equation.
- (3) Convection. The marker particles are moved according to the velocity components in their vicinities, with temperature interpolated from background cells. Adjustments are made when particles across cell boundaries.
- (4) Tracking the free surface. Update the liquid level set value according to the positions of the marker particles.
- (5) Compute temperature. Mapping the particle temperature into cells, and exchange the cell

temperature according to the thermal diffusion.

(6) Re-distribute particles. Check the particle distribution and re-distribute particles when necessary, or put new particles into the inlet region according to the inlet boundary conditions.

The marker particles introduced into this calculation are only for the purpose of indicating fluid configuration, more specifically, helping to compute the liquid level-set value. With the liquid level-set value computed, the free surface of the liquid could be described more precisely.

Lagrangian Interpolation for Free Surface

When a free surface does not pass through the center of the cell, the second derivative of pressure φ in Eq. (3) should be taken carefully, the traditional differential process for the uniform grid will not be accurate enough. As a brief description, see Fig. 1. Let Φ denotes the level set value, the free surface with $\Phi = 0$ passes between point x_{k-1} and point x_k , and the applied pressure is φ_a . As x_{k-1} lies out the liquid, it cannot be applied to decide the grad of φ at point k because of discontinuity. Instead, we use x_a , x_k and x_{k+1} to construct a second order Lagrangian interpolation for φ :

$$\varphi(x) = \frac{(x - x_k)(x - x_{k+1})}{(x_a - x_k)(x_a - x_{k+1})} \varphi_a + \frac{(x - x_a)(x - x_{k+1})}{(x_k - x_a)(x_k - x_{k+1})} \varphi_k + \frac{(x - x_a)(x - x_k)}{(x_{k+1} - x_a)(x_{k+1} - x_k)} \varphi_{k+1} \quad (6)$$

and take its second derivative with respect to x gives

$$\frac{d^2 \varphi}{dx^2} = \frac{2}{(x_a - x_k)(x_a - x_{k+1})} \varphi_a + \frac{2}{(x_k - x_a)(x_k - x_{k+1})} \varphi_k + \frac{2}{(x_{k+1} - x_a)(x_{k+1} - x_k)} \varphi_{k+1} \quad (7)$$

On the right side of Eq. (7), only x_a is to be decided, which can be estimated by level set value of its adjacent cell centers:

$$x_a = \frac{x_{k-1}\Phi_k - x_k\Phi_{k-1}}{\Phi_k - \Phi_{k-1}} \quad (8)$$

The Lagrangian interpolation as shown by Eq. (7) and Eq. (8) can be easily extended to two-dimensional or three-dimensional cases.

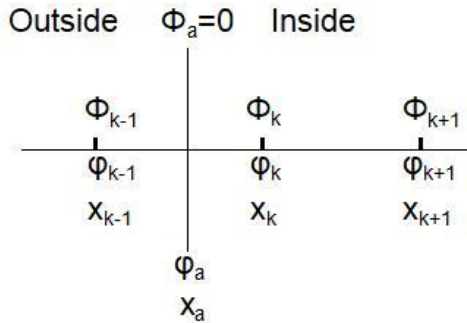


Fig. 1. One-dimensional schematic for free surface passes between cell centers

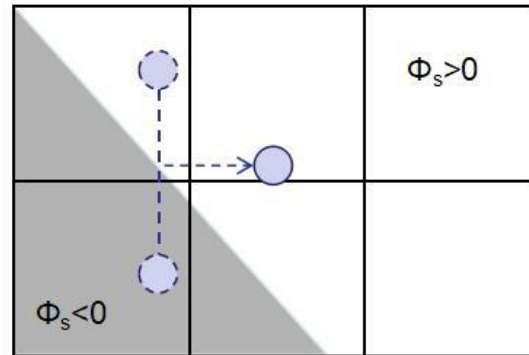


Fig. 2. Using level set to define oblique wall. The value inside the wall is negative while positive outside the wall.

Level set for the wall

The geometry of the casting is usually very complicated, involving many bevels and complex surfaces, which cannot be properly expressed using a simple uniform grid. For example, as Fig. 2 shows, an oblique wall blocks a dropping particle and deflects it to the right. If a simple uniform grid is used, because no oblique wall is defined, the particle will be reflected back up, not to the right. To handle this problem, one can preserve a level set field Φ_s for the solid wall, which is defined at the cell corner rather than the cell center. Φ_s is usually a signed distance field to the wall surface that computed when meshing the geometry. At the convection step [step (3)], whenever a particle runs to a new position \vec{x}^* , compute its solid level set value $\Phi_s(\vec{x}^*)$. If this value is negative (\vec{x}^* inside the solid), a reflection operation should be taken:

$$\vec{x}^{n+1} = \vec{x}^* - (1 + \alpha) \Phi_s(\vec{x}^*) \frac{\nabla \Phi_s(\vec{x}^*)}{|\nabla \Phi_s(\vec{x}^*)|} \quad (9)$$

In Eq. (9), the collision coefficient α should be between 0 and 1. When $\alpha = 1$ a completely elastic collision strategy is applied, and $\alpha < 1$ applies an inelastic collision at the normal direction of the wall surface.

Examples and Applications

Collapse of a water column

The collapse of a water column was calculated with the parameters listed in table 1, and the geometry is depicted in Fig. 3. This geometry refers to the article of Koshizuka and Oka^[10], because they had carried out an experiment^[17]. In the experiment, a removable board supports the initial water column, and then it is pulled up within 0.05s and collapse starts. In our calculation, the pulling process of the board is neglected. In addition, we have also calculated this case with a VOF (Volume of Fluid) method using the same discrete procedure, and the same grid resolution. The results are shown in Fig. 4, cells with fluid are visualized rather than particles, in order to make a better comparison to the VOF method.

As shown in Fig. 4, at the beginning of the collapse, both methods have similar results. At $t=0.24s$, the results of the two calculations are slightly different. The fluid calculated by the MAC based method flows a bit faster than the VOF based method. It seems that the fluid calculated by VOF based method exhibits a greater viscosity. And by $t=0.50s$, the fluid interface calculated by the two methods is very different. From the MAC based method, we can see that the water hits the wall and breaks into pieces, forms very sharp and discontinuous interfaces. From the VOF based method, the water has smooth and continuous interfaces, which is not so real when referred to the experiment results.

Table. 1. Calculation parameters

Parameters	Value
Horizontal grid resolution	292
Vertical grid resolution	146
Particles per cell	4
Particle radius	$1.732 \delta x$
Particle collision coefficient	$\alpha = 0$
Courant condition	$u_{\max} \delta t / \delta x \leq 0.5$
Maximum time step	$\delta t \leq 10^{-3} s$

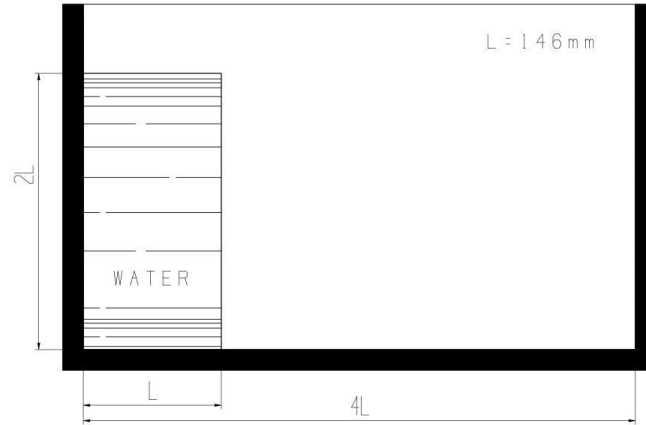


Fig. 3. Geometry of collapse of a water column

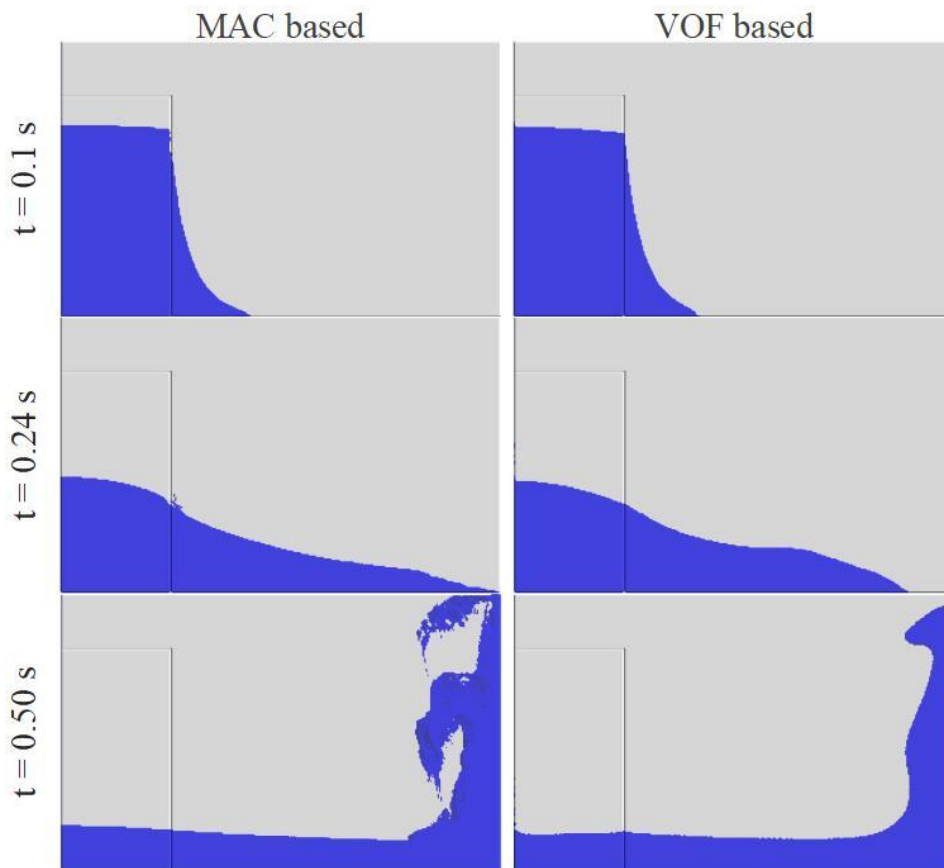


Fig. 4. Collapse of a water column calculated with MAC based method and VOF based method

Filling of a complex casting

The filling process of a practical iron casting scheme was calculated to study the application of this method. The simulation results were compared to results that calculated by a VOF program. Fig. 5(a) shows the geometric model of the investment casting system. The overall dimensions of the system are 200mm×100mm×500mm. There are four main sprues, each with 8 castings evenly distributed. The grid resolutions are 100×50×250, and there are a total of 1.25 million cells on the solution domain. In each cell, only 6 particles are placed for the sake of efficiency. Table. 2 shows the calculation parameters.

Table. 2. Parameters needed for the iron casting

Parameters	Value
Liquid iron density (kg/m^3)	6800
Liquid iron dynamic viscosity ($\text{Pa}\cdot\text{s}$)	4.98e-3
Liquid iron specific heat ($\text{kJ}/(\text{kg}\cdot\text{K})$)	0.82
Liquidus temperature ($^{\circ}\text{C}$)	1500
Solidus temperature ($^{\circ}\text{C}$)	1429
Latent heat (kJ/kg)	250
Acceleration of gravity (m/s^2)	(0, 0, -9.8)
Inlet temperature ($^{\circ}\text{C}$)	1700
Inlet velocity (m/s)	0.15

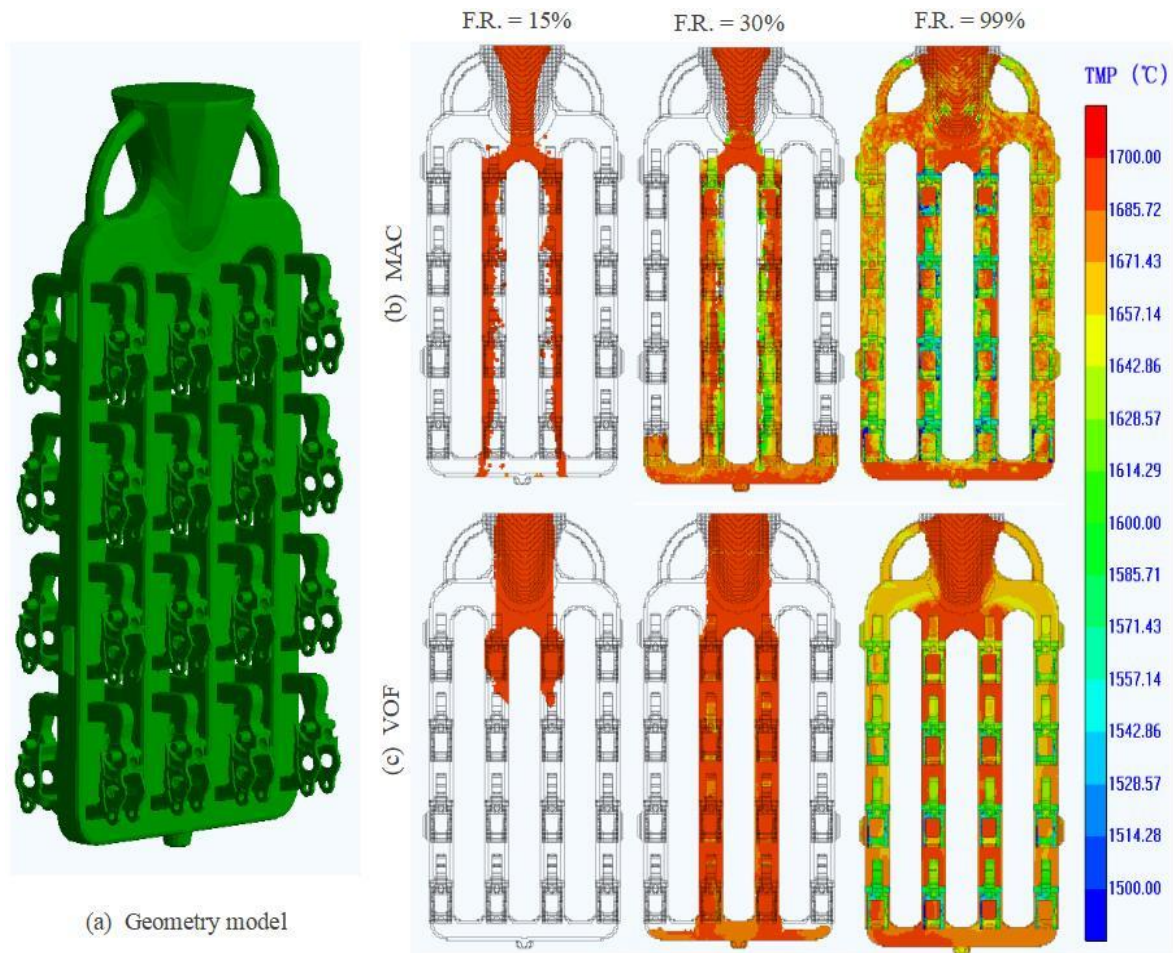


Fig. 5. Simulation of a complex casting filling process

Fig. 5(b) shows the results given by a MAC based method and Fig. 5(c) shows the results given by a VOF based method. After flowing into the casting system from the top of the pouring cup, the molten metal accelerates down until it hits the wall, and then diverts into two streams and continues to accelerate downward along the two inner sprues. Note that the molten metal will not fill the sprues at this moment, because the sum of the section areas of the inner sprues is larger. When these two streams reach the bottom of the system, they spread out symmetrically along the runner, and then enter the two outside sprues under the effect of pressure. At last, the castings start to be filled from bottom to top.

Both methods give the correct results, but the VOF based method gives a coarser stream in the early stage of the filling process as if the molten metal suffers a large surface tension, which is not so real. The MAC based method gives a better-looking stream. For the temperature field, VOF based method gives a smoother distribution, and the MAC based method gives a mushy local temperature distribution, seems that the flow is more turbulent. This phenomenon may be caused by the re-distributing process. The particles have a tendency to gather together after a long run, and the re-distributing process must be carried out even though extra errors are introduced.

It takes 125 minutes to complete this simulation by the MAC based method using an Intel Core i7-6700HQ CPU, with an 8.0GB RAM. While the VOF based method takes only 98 minutes. That is to say, keeping track of the particles, takes not only more space, but also more time. This is a shortcoming of the MAC based method. However, with the particles tracked, something more complicated could be described, such as the oxide inclusions, the rising of slags in the mushy phase, which requires further research.

Conclusion

In this paper, the marker-and-cell method is introduced to simulate the casting filling process. The level set technique is used to enable the description of the complicated free surface and wall boundary, in a simple uniform grid. This MAC based method tracks the fluid surface directly by the marker particles, requires not only more storage space but also more calculation time than the VOF based method. However, it produces very sharp and discontinuous interfaces easily, which is a really pleasing property that suits the casting process with high flowing speed. Besides, with the particles tracked, some difficult problems like the oxide inclusions and the sand washing problems could become easier to be described.

Reference

- [1] Vazquez V, Juarez-Hernandez A, Mascarenas A, et al. Cold shut formation analysis on a free lead yellow brass tap. *Engineering Failure Analysis* 2010; 17(6): 1285-1289.
- [2] Cao L, Sun F, Chen T, et al. Quantitative prediction of oxide inclusion defects inside the casting and on the walls during cast-filling processes. *International Journal of Heat and Mass Transfer* 2018; 119: 614-623.
- [3] Hirt CW, Nichols BD. Volume of fluid (VOF) method for the dynamics of free boundaries. *J. Comput. Phys* 1981; 39: 201.
- [4] Sussman M. A level set approach for computing solutions to incompressible two-phase flow [D]. Los Angeles: University of California, 1994.
- [5] Sussman, M, Puckett EG. A coupled level set and volume-of-fluid method for computing 3D and axisymmetric incompressible two-phase flows. *Journal of computational physics* 2000; 162(2): 301-337.
- [6] Tryggvason G, Bunner B, Esmaeeli, A, et al. A front-tracking method for the computations of multiphase flow. *Journal of Computational Physics* 2001; 169(2): 708-759.
- [7] Ryskin, G, Leal LG. Numerical solution of free-boundary problems in fluid mechanics. Part 1. The finite-difference technique. *Journal of Fluid Mechanics* 1984; 148: 1-17.
- [8] Ryskin G, Leal LG. Numerical solution of free-boundary problems in fluid mechanics. Part 2. Buoyancy-driven motion of a gas bubble through a quiescent liquid. *Journal of Fluid Mechanics* 1984; 148: 19-35.
- [9] Takagi S. Three-dimensional deformation of a rising bubble. In *Proc. German-Japanese Symp. on Multi-Phase Flow* 1994; 499.
- [10] Koshizuka S, Oka Y. Moving-particle semi-implicit method for fragmentation of incompressible fluid. *Nuclear science and engineering* 1996; 123(3): 421-434.
- [11] Gingold RA, Monaghan JJ. Smoothed particle hydrodynamics: theory and application to non-spherical

- stars. Monthly notices of the royal astronomical society 1977; 181(3): 375-389.
- [12] Hoogerbrugge PJ, Koelman J. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. EPL (Europhysics Letters) 1992; 19(3): 155.
 - [13] Koelman J, Hoogerbrugge PJ. Dynamic simulations of hard-sphere suspensions under steady shear. EPL (Europhysics Letters) 1993; 21(3): 363.
 - [14] Harlow, FH, Welch JE. Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. The physics of fluids 1965; 8(12): 2182-2189.
 - [15] Teman R. Navier- Stokes Equations [M]. 3rd Ed, Amsterdam: North-Holland, 1984:30.
 - [16] Chorin, AJ. Numerical solution of the Navier-Stokes equations. Mathematics of computation 1968; 22(104): 745-762.
 - [17] Koshizuka S, Oka Y, Tamako H. A particle method for calculating splashing of incompressible viscous fluid. American Nuclear Society, Inc., La Grange Park, IL (United States) 1995.

A positional FEM formulation applied to 2D dynamic nonlinear analysis of structures and mechanisms with improved frictional internal sliding connections

***Tiago Morkis Siqueira¹ and Humberto Breves Coda¹**

¹Department of Structural Engineering, São Carlos School of Engineering, University of São Paulo, São Carlos, SP, Brazil.

*Presenting and corresponding author: morkis@usp.br

Abstract

In this work we extend a total Lagrangian formulation applied to the dynamical analysis of plane frames containing sliding connections (prismatic and cylindrical joints) to include frictional dissipation. An improvement in the friction force model is proposed to smooth the force transition from rest to motion states, allowing the proper modelling of residual displacements at the joints. Friction dissipation is added to the total mechanical energy for the achievement of the equations of motion by the Principle of Stationary Total Energy. The resulting nonlinear equations are solved by the Newton-Raphson method. Some examples are presented to show the formulation effectiveness.

Keywords: Friction model; Sliding connection; Nonlinear dynamics; Lagrange multiplier.

Introduction

In the dynamical analysis of structures and mechanisms conservative systems simplifications are frequently assumed. However, real bodies present dissipation due to several sources. The frictional dissipation effect, in particular, is important to be considered when relative motion from parts of the body exists. This is the case of sliding connections, such as prismatic and cylindrical joints, that by introducing translational movement among body members allows friction forces to develop along their surfaces contact.

The friction phenomena itself has a very complex nature that mathematical models try to describe, with more or less accuracy, depending on which aspects of the friction force the proposed expressions intent to consider. The models become more detailed and representative at the cost of more parameters. A comprehensive surveys on friction models can be found in [1–3]. In the literature, friction models can be classified generally in dynamic or static [3], whether the force is, respectively, time dependent or not. Thus, static models, as opposed to dynamical models, dismiss the introduction of state variables to the problem, rendering a straightforward description of the force expression. Still, static models have difficulties in describing the discontinuity of the friction force at null speed, which results in instabilities in the numerical solution. Several models try to circumvent this problem [4–9] commonly assuming null friction at null speed, which is not a good approximation when relative motion is intermittent, or require additional parameters for the transition between motion and rest states.

In this work, we propose a modification on a classic static friction model to be employed in sliding connections of plane frame finite elements by positional formulation. The improved model is based on the Coulomb friction considering the Stribeck curve and viscous effect. To reduce the abrupt transition between rest and motion states, an interpolation of the static friction value to the resultant force is employed in a quasi-null relative speed interval. Thus, the proposed model intends to represent the force transition in a smooth way, allowing the description of residual displacements when the final stop stage is achieved, which is important to ensure high precision movements reproduction in structures and mechanisms.

The framework used to model the dynamical system [10,11] is a fully nonlinear finite element approach for large deformations based on a total Lagrangian description of the solids which uses positions as the main degrees of freedom. The Saint-Venant-Kirchhoff constitutive model is adopted to define the plane frame elastic strain energy using the Green-Lagrange strain and the second Piola-Kirchhoff stress tensor. Since in this technique velocity and acceleration are referred to a Lagrangian inertial reference frame, the Newmark approximation is applied to integrate time. The sliding connections, as prismatic and cylindrical joints, are introduced in the total energy of the system by means of Lagrange multipliers [11]. Moreover, friction dissipation is added to the energy expression to allow finding the equations of motion (comprising the frictional effect) by the Principle of Stationary Total Energy. The resulting nonlinear system is solved by the Newton-Raphson method.

This work is organized as follows. First brief aspects of the nonlinear plane frame element need to be presented followed by the kinematical constraints that the sliding connections impose. Then, the dynamical equilibrium is obtained. Known the system parameters, the friction force can be introduced in its variational form and the improved model is presented. Time integration and system solution follows this explanation. Lastly, examples are shown to demonstrate the developed formulation. Dyadic notation is preferred throughout this text due its brevity; however, index notation is also used to clarify particular aspects when necessary.

Nonlinear finite element kinematics

The plane frame finite element employed is presented thoroughly elsewhere [10,11], however, to develop the present work some aspects need to be briefly stated. As the finite element behaviour is represented by a total Lagrangian description, its strain field needs to be obtained as a function of the initial and current configurations of the solid, restricted to a finite number of degrees of freedom.

In the positional approach of the FEM, instead of nodal displacements, the parameters of the discretized plane frame are its positions (coordinates) and the cross section angle (Fig. 1). The deformation function, \vec{f} , depicted in Fig. 2, can be written indirectly as function of the non-dimensional space and nodal parameters by mappings from the non-dimensional space to the initial configuration, \vec{f}^0 , as:

$$\begin{aligned} f_1^0 &= x_1 = \phi_\ell(\xi) X_1^\ell + \frac{h_0}{2} \eta \cos[\phi_\ell(\xi) \theta_\ell^0] \\ f_2^0 &= x_2 = \phi_\ell(\xi) X_2^\ell + \frac{h_0}{2} \eta \sin[\phi_\ell(\xi) \theta_\ell^0] \end{aligned} \quad (1)$$

and to the current configuration, \vec{f}^1 , as:

$$\begin{aligned} f_1^1 &= y_1 = \phi_\ell(\xi)Y_1^\ell + \frac{h_0}{2}\eta \cos[\phi_\ell(\xi)\theta_\ell] \\ f_2^1 &= y_2 = \phi_\ell(\xi)Y_2^\ell + \frac{h_0}{2}\eta \sin[\phi_\ell(\xi)\theta_\ell] \end{aligned} \quad (2)$$

where \bar{x} and \bar{y} represents any point on the domain of a finite element in the initial and current configuration, respectively. The coordinates for both directions $i = 1, 2$ of each node ℓ along the reference line in the initial and current configurations are X_i^ℓ and Y_i^ℓ , respectively. The initial nodal value of the cross section angle is θ_ℓ^0 and after deformation is denoted as θ_ℓ . In addition, the cross section height is h_0 , ξ is the non-dimensional space variable in the direction of the reference line and η follows the height direction. The shape functions $\phi_\ell(\xi)$ are obtained by Lagrange polynomials of any order.

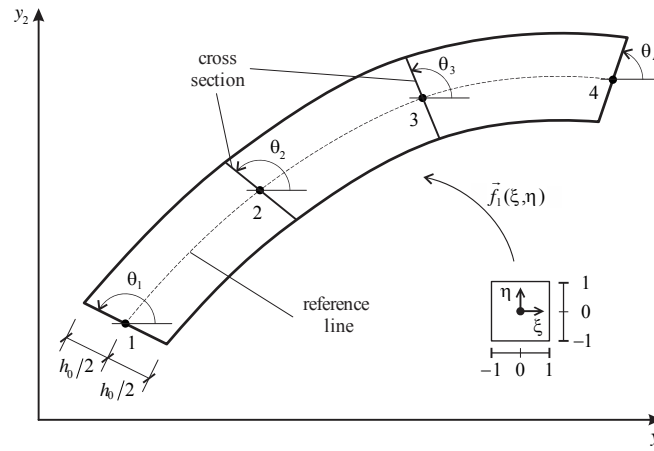


Fig. 1. Current configuration mapping for a cubic approximation

The deformation function can be written as a composition of the previous mappings, eq. (1) and (2), as presented by [12,13], as:

$$\bar{f} = \bar{f}^1 \circ (\bar{f}^0)^{-1} \quad (3)$$

Since only the gradient \mathbf{A} of the deformation function, but not the function itself, is necessary to obtain the strain field [14], we can write:

$$\mathbf{A} = \text{Grad}(\bar{f}) = \mathbf{A}^1 \cdot (\mathbf{A}^0)^{-1} \quad (4)$$

where,

$$A_{ij}^0 = \frac{\partial f_i^0}{\partial \xi_j} \quad \text{and} \quad A_{ij}^1 = \frac{\partial f_i^1}{\partial \xi_j} \quad (5)$$

During the iterative solution strategy both \mathbf{A}^0 and \mathbf{A}^1 are numerical values calculated at the integration points resulting in a purely numerical procedure.

Since the Saint-Venant-Kirchhoff constitutive law is employed, the Green-Lagrange strain tensor \mathbf{E} have to be calculated. This objective measure is given, for instance, by [14]:

$$\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I}) = \frac{1}{2}(\mathbf{A}^t \cdot \mathbf{A} - \mathbf{I}) \quad (6)$$

where \mathbf{I} is the second order identity tensor and $\mathbf{C} = \mathbf{A}^t \cdot \mathbf{A}$ is the right Cauchy-Green stretch tensor.

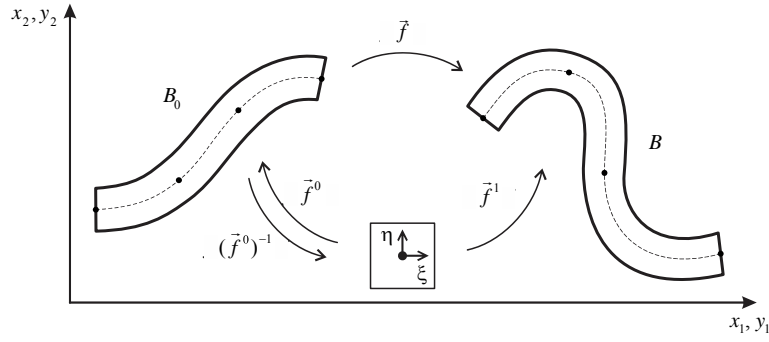


Fig. 2. Deformation mapping

As there is no relation between the cross section angle and the slope of the reference line, the frame kinematic can be regarded as Reissner's. It should be mentioned that the cross section dimensions are maintained the same during motion, thus, to avoid volumetric locking, the constitutive equation is relaxed in order to exclude transverse expansions.

Kinematical constraints due to sliding connections

To develop the friction force it is required first to describe how the constraint equations for the sliding connections are defined, particularly concerning the curvilinear position, which is a new variable introduced in the equations of motion, and at which this force will act. Here we summarize the description of the connections as a prismatic or a cylindrical joint. More details can be obtained in [11].

Sliding connections are the ones that constrain relative translations between parts of the body. Fig. 3 illustrates both joints and their plane representation. In either case, a sliding node, at which the joint exists, is constrained to move over a trajectory comprised of path elements. The difference between the prismatic and the cylindrical joint is the relative rotation, which is allowed only by the last one.

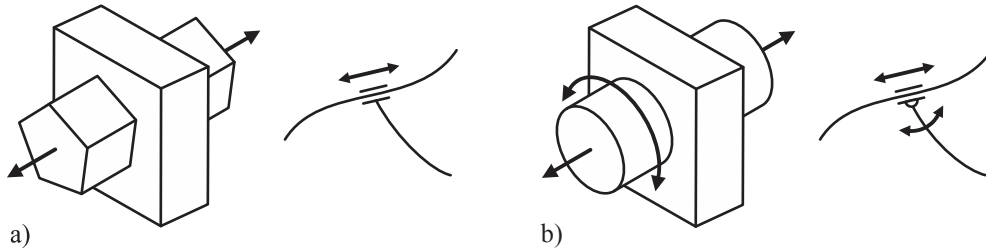


Fig. 3. Sliding connections and its plane representation: a) prismatic and b) cylindrical joints

Fig. 4 depicts the case of a prismatic joint, belonging to node \hat{P} , and its path contact point \bar{P} . The connection is free to move along the path $s(\xi)$ defined by path finite elements, which, although not used in this work, may have an arbitrary roughness profile $\vec{r}(s)$. The notation $(\bar{\bullet})$ is used to identify variables related to path elements and $(\hat{\bullet})$ is used for sliding elements. The new variable $s_p = s(\xi_p)$ that defines the curvilinear position and the cross section orientation of the path point is also illustrated.

The constraint equations, \vec{c} , can be written for both types of joints as a single expression:

$$c_i = \hat{Y}_i^P - \phi_\ell(\xi_P) \bar{Y}_i^\ell - \Delta \theta_P^0 \delta_{i3} - r_i(s_P)(1 - \delta_{(i)3}) = 0_i \quad (7)$$

where: i is the direction ($i=1,2,3$ for prismatic joints and $i=1,2$ for cylindrical joints); δ_{ij} is the Kronecker delta; $\Delta\theta_p^0 = \hat{\theta}_p^0 - \bar{\theta}_p^0$ is the difference of cross sections orientations at the initial configuration, which must be constant during the sliding process of a prismatic joint to maintain a fixed relative angle; and the components of the roughness profile, obtained by its height function $\|\vec{r}(s)\|$, are given by:

$$\begin{aligned} r_1(s_p) &= \|\vec{r}(s_p)\| \cos[\phi_\ell(\xi_p)\bar{\theta}_\ell] \\ r_2(s_p) &= \|\vec{r}(s_p)\| \sin[\phi_\ell(\xi_p)\bar{\theta}_\ell] \end{aligned} \quad (8)$$

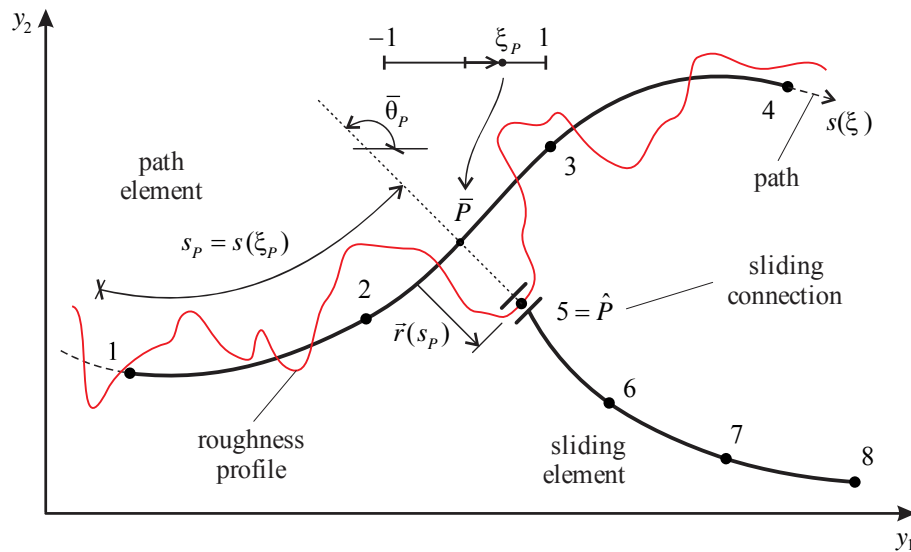


Fig. 4. Sliding connection over an arbitrary path (depicted for a prismatic joint)

It is noteworthy that the curvilinear variable $s(\xi)$ represents an arch-length function defined by the non-dimensional coordinate ξ and the path element coordinates.

Unconstrained equations of motion

Using the Law of Conservation of Energy, the dynamical equilibrium of a conservative system is obtained by its total energy Π_0 as:

$$\Pi_0 = \Pi - \mathcal{Q} \quad (9)$$

where Q represents the dissipation of a ‘larger’ system of total energy Π . Eq. (9) can be rewritten as:

$$\Pi = \Pi_0 + \mathcal{Q} \quad (10)$$

or, making explicit the energy parcels of the new larger conservative system:

$$\Pi = \mathcal{U} - \mathcal{P} + \mathcal{K} + \mathcal{Q} \quad (11)$$

where \mathcal{U} is the stored elastic strain energy, \mathcal{P} is the potential of conservative external forces and \mathcal{K} is the kinetic energy of the body.

Following Lanczos and others [15–17], it is not always possible to write down closed expressions for dissipative parcels but only its infinitesimal change. Thus, the equations of

motion are stated from the variation of the energies present in eq. (11), which is understood as the Principle of Stationary Total Energy:

$$\delta\Pi = \delta\mathcal{U} - \delta\mathcal{P} + \delta\mathcal{K} + \delta\mathcal{Q} = 0 \quad (12)$$

in which the symbol δ means variation.

The total energy can be stated by writing the known expressions of the energies in eq. (11) as function of the current configuration nodal parameters of the discretized body, grouped in the vector \vec{R} , as:

$$\Pi(\vec{R}) = \int_{V_0} u(\mathbf{E}(\vec{R})) dV_0 - \vec{F} \cdot \vec{R} - \int_{s_0} \vec{q} \cdot \vec{y} ds_0 + \frac{1}{2} \int_{V_0} \rho_0 \dot{\vec{y}} \cdot \dot{\vec{y}} dV_0 + \mathcal{Q}(\vec{R}) \quad (13)$$

where the specific strain energy u depends on the strain state \mathbf{E} of the body, eq. (6), which is function of the nodal parameters \vec{R} , as defined by the gradient of the deformation function in eq. (4).

As mentioned before, the Saint-Venant-Kirchhoff constitutive relation is employed due to its simplicity and good representation of large displacements on solids that remain in the small to moderate strain regimen, which comprehends the majority of the usual applications in engineering. For the plane frame utilized, its specific energy is given as:

$$u = \frac{\mathbb{E}}{2} (E_{11}^2 + E_{22}^2) + \mathbb{G} (E_{12}^2 + E_{21}^2) \quad (14)$$

where \mathbb{E} is the longitudinal elastic parameter that approaches the Young modulus for small strains. The shear elastic modulus is $\mathbb{G} = \mathbb{E}/[2(1+\nu)]$, being ν a constant that reproduces the Poisson ratio for small strains. The second Piola-Kirchhoff stress tensor is easily obtained by the energy conjugacy property as:

$$\mathbf{S} = \frac{\partial u}{\partial \mathbf{E}} \quad (15)$$

Still in eq. (13), \vec{F} and \vec{q} are the concentrated and distributed conservative external loading, respectively. The initial length of the frame reference line is s_0 . The material mass density in the initial configuration, of volume V_0 , is ρ_0 . The material points' velocity is denoted using the over-dot as $\dot{\vec{y}}$. External damping dissipation, proportional to the velocity in its differential form (Rayleigh damping), is introduced as:

$$\frac{\partial \mathcal{Q}}{\partial \dot{\vec{y}}} \cdot \delta \dot{\vec{y}} = \int_{V_0} c_\rho \rho_0 \dot{\vec{y}} \cdot \delta \dot{\vec{y}} dV_0 \quad (16)$$

in which c_ρ is a proportionality constant.

The equations of motion (geometric nonlinear dynamical equilibrium) are obtained by the development of the variations in eq. (13). In a compact form, the equilibrium can be written as:

$$\vec{F}^{\text{int}} - \vec{F} + \mathbf{M} \cdot \ddot{\vec{R}} + \mathbf{D} \cdot \dot{\vec{R}} = \vec{0} \quad (17)$$

where: $\vec{F}^{\text{int}} = \text{Grad}(\mathcal{U})$ is the internal force vector; \vec{F} collects all the external loads; \mathbf{M} is a constant mass matrix; $\mathbf{D} = c_\rho \mathbf{M}$ is the external damping matrix; and $\dot{\vec{R}}$ and $\ddot{\vec{R}}$ are the velocity and acceleration vectors of the nodal parameters. More details about the development of the variations of eq. (13) can be obtained in [10,11].

Constrained equations of motion

The dynamical equilibrium stated by eq. (17) is called unconstrained since no restraints, such as the ones from the sliding connections, are considered. The literature presents several consolidated methodologies to impose constraints such as in [15,18] on mechanical and structural applications or in [19–21] which deal with general optimization problems. Here, we employ the well-known Lagrange multiplier method along with the Principle of Stationary Total Energy to impose the sliding restrictions. In what regards the later introduction of friction dissipation, the multipliers are of great value since in Mechanics they might be understood as the contact forces between bodies, an essential information for the friction model.

The Principle of Stationary Total Energy is extended for the case of holonomic constraints by modifying the total energy through the introduction of a new potential \mathcal{C} , referred as the constraint potential, as:

$$\Pi = \mathcal{U} - \mathcal{P} + \mathcal{K} + \mathcal{Q} + \mathcal{C} \quad (18)$$

When using Lagrange multipliers the expression of the new potential is simply given by:

$$\mathcal{C} = \vec{\lambda} \cdot \vec{c} \quad (19)$$

where $\vec{\lambda}$ represents the vector of multipliers, which are new variables of the system. Eq. (19) indicates the presence of a multiplier for each constraint equation in \vec{c} . It is worth mentioning that the constraint potential is null at the solution, therefore, the total energy is not altered.

Knowing the expression of \mathcal{C} , the first variation of the constrained energy, eq. (18), is:

$$\delta\Pi = \delta\mathcal{U} - \delta\mathcal{P} + \delta\mathcal{K} + \delta\mathcal{Q} + \delta\mathcal{C} = 0 \quad (20)$$

which, neglecting friction, can be developed in a similar fashion as the unconstrained case leading to the constrained equations of motion (constrained geometric nonlinear dynamical equilibrium), expressed in a compact form as:

$$\vec{F}^{\text{int}} - \vec{F} + \mathbf{M} \cdot \ddot{\vec{Y}} + \mathbf{D} \cdot \dot{\vec{Y}} + \vec{F}^c = \vec{0} \quad (21)$$

in which \vec{F}^c represents the restriction forces arriving from the constraint potential. As the multipliers are new variables, the variation of \mathcal{C} is organized in the following force vector, which separates the parameters \vec{Y} (including s_p) and the multipliers:

$$\delta\mathcal{C} = \delta\vec{Y} \cdot \nabla\vec{c} \cdot \vec{\lambda} + \delta\vec{\lambda} \cdot \vec{c} = \left\{ \delta\vec{Y} \quad \delta\vec{\lambda} \right\} \cdot \left\{ \begin{matrix} \nabla\vec{c} \cdot \vec{\lambda} \\ \vec{c} \end{matrix} \right\} = \left\{ \delta\vec{Y} \quad \delta\vec{\lambda} \right\} \cdot \vec{F}^c \quad (22)$$

where the tensor $\nabla\vec{c}$ represents the Jacobian matrix of the constraint vector. In order to shorten this presentation, the derivatives of the constraint equation for the sliding connections, eq. (7), can be found in reference [11].

Friction force on the sliding connection

The friction force is included in the system directly in the Principle of Stationary Total Energy as a dissipative potential. As mentioned previously, dissipative potentials are introduced in their differential form since closed expressions might be unknown, as is the case for the dissipated friction energy \mathcal{Q}_f . However, the variation of this potential can be written as the work done by the friction force \vec{F}^f on its displacement trajectory \vec{d} as:

$$\delta\mathcal{Q}_f = \vec{F}^f \cdot \delta\vec{d} \quad (23)$$

To develop eq. (23), parameters that describes the force displacement must be chosen. For that, the coordinates of the sliding node and its path contact point could be picked. However, since in the previous formulation the curvilinear position s_p is already used as an intrinsic variable, the displacement along the trajectory is simply the scalar expression $d = s_p - s_p^0$, being s_p^0 an arbitrary initial value, and its variation is $\delta d = \delta s_p$. As the friction force acts tangentially to the trajectory, with its value given by F_f , the dissipative parcel is introduced directly in the curvilinear position as:

$$\delta Q_f = F_f \delta s_p \quad (24)$$

To organize the equations of motion system, we make $\vec{\Lambda}^f = \{F_f\}$, the previous equation is rewritten as:

$$\delta Q_f = \{\delta s_p\} \cdot \vec{\Lambda}^f \quad (25)$$

Considering the correspondence of the friction force vector $\vec{\Lambda}^f$ to the system variables, the equations of motion are restated to include frictional dissipation as:

$$\vec{F}^{\text{int}} - \vec{F} + \mathbf{M} \cdot \ddot{\vec{r}} + \mathbf{D} \cdot \dot{\vec{r}} + \vec{F}^c + \vec{\Lambda}^f = \vec{0} \quad (26)$$

Improved friction model

From the manner that the friction force was introduced in the formulation any expression can be readily applied without changing the equilibrium equation. In this work, we employ the Coulomb model considering the Stribeck effect and viscous friction. Fig. 5 shows the overall behaviour of the friction force with the relative velocity among bodies. This model considers the stick-slip effect, which is the difference between the friction force at rest (static friction) and at motion (kinetic friction), by the Stribeck curve, using the most usual expression proposed by Bo e Pavelescu [22]. A linear model represents the viscous friction, which occurs if lubricant layers are present on the surfaces.

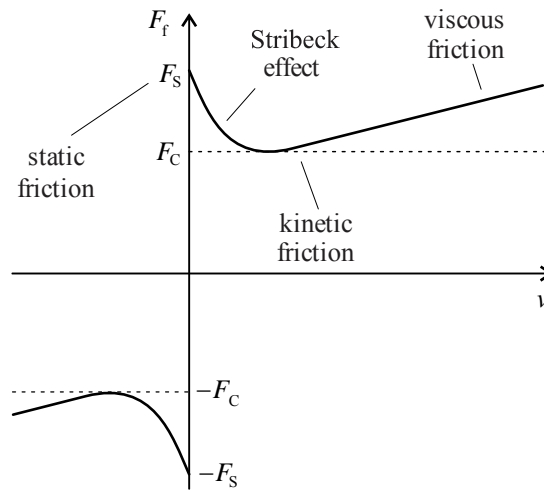


Fig. 5. Friction model representation

The mathematical expression for the friction force due this model is written as:

$$F_f = \begin{cases} \left[F_C + (F_S - F_C) e^{-(|v|/v_\sigma)^{\delta_\sigma}} \right] \text{sgn}(v) + \eta v & \text{if } v \neq 0 \\ \min(F_S, F_R) \text{sgn}(F_R) & \text{if } v = 0 \end{cases} \quad (27)$$

with the static and kinetic friction forces, respectively, given by:

$$F_S = \mu_s F_N \quad \text{and} \quad F_C = \mu_k F_N \quad (28)$$

where, μ_s and μ_k are, respectively, the static and kinetic friction coefficients and F_N the absolute value of the contact force normal to the trajectory at the joint contact point. In addition, η is the viscous friction coefficient and $v = \dot{s}_p$ is the joint relative velocity tangential to the path (directly obtained from the curvilinear position). The Stribeck parameters are its decay velocity v_σ and power δ_σ . The sign function is represented by $\text{sgn}(\bullet)$.

For null relative velocity, second condition in eq. (27), the tangential resultant force F_R acting on the connection is required for comparison with the static friction value. This evaluation verifies if there is tendency of motion in case the resultant force is greater than the static friction force, or not, otherwise. However, in the transition from motion to rest, the shift from one force to another is done abruptly using their smallest value, which, as concern numerical simulations, may create instabilities in the system solution and the need to use very small time steps.

For this reason, a linear interpolation between the values of the static friction force and the resultant force is proposed for the stabilization of the friction force response when there is a range $[-v_0, v_0]$ of quasi-null velocities, as depicted in Fig. 6.

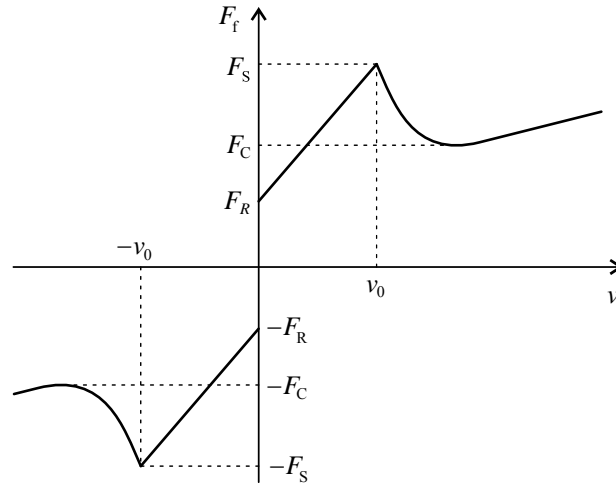


Fig. 6. Improved friction model

The improved friction model is written as:

$$F_f = \begin{cases} \left[F_C + (F_S - F_C) e^{-(|v|/v_\sigma)^{\delta_\sigma}} \right] \text{sgn}(v) + \eta v & \text{if } |v| > |v_0| \\ -F_S \text{sgn}(F_R) & \text{if } |v| \leq |v_0| \text{ and } |F_R| \geq F_S \\ \frac{F_S - |F_R|}{|v_0|} v + F_R & \text{if } |v| \leq |v_0| \text{ and } |F_R| < F_S \end{cases} \quad (29)$$

where v_0 is the quasi-null speed limit. One should note that F_C and F_s are always positive since they are obtained from the absolute value of the normal force, eq. (28), thus, the sign of the friction force in eq. (29) depends on the values and signs of the relative velocity and resultant force.

In the proposed approach, F_R is not a constant value, but depends upon the system own force state at a given time instant, which can even be null, if applicable. Therefore, the system response can be stabilized by means of a smooth transition from the motion state to rest state and vice versa. In addition, by taking into account the resulting force, the proposed improved friction model is capable to represent residual displacements of the sliding connection.

It should be noted that the quasi-null speed limit value v_0 depends on the adopted time step, or, inversely, the chosen time step has to be able to represent the movement when in the quasi-null velocity interval. For better convergence of the iterative solution method, the recommended value of the limit velocity should be close to the relative stop speed of the bodies but not too small to allow the smooth transition among forces at rest.

Known the coefficients of the model, which depends on the materials that make the sliding connection and its path, the forces required to calculate the friction force have to be related to the variables that describe the joint. The normal force vector \vec{F}^N is found from the component of the Lagrange multipliers vector due to the translational constraints, $\vec{\lambda} = \{\lambda_1, \lambda_2\}$, at the normal direction of the path at the contact point, defined by the normal vector \vec{N}^P , as:

$$\vec{F}^N = \frac{\vec{\lambda} \cdot \vec{N}^P}{\|\vec{N}^P\|} \frac{\vec{N}^P}{\|\vec{N}^P\|} \quad (30)$$

and its absolute value, actually used in the calculation, is:

$$F_N = \|\vec{F}^N\| = \frac{|\vec{\lambda} \cdot \vec{N}^P|}{\|\vec{N}^P\|} \quad (31)$$

In the plane case, the components of the normal vector are obtained from the tangent vector of the path finite element at the contact point, $\vec{T}_i^P = \phi_{\ell, \xi}(\xi_P) \vec{Y}_i^\ell$ ($i=1,2$), as $\vec{N}_1^P = -\vec{T}_2^P$ and $\vec{N}_2^P = \vec{T}_1^P$.

The resultant force, equal to the inertial force at the sliding node, is obtained directly from the equilibrium equation (26), considering only the sliding node degrees of freedom (positions and curvilinear variable), as:

$$\vec{F}^R = \vec{F} - \vec{F}^{\text{int}} - \vec{F}^c \quad (32)$$

or, as to identify the terms referred to the degrees of freedom:

$$\begin{Bmatrix} F_{\bar{Y}_1^P}^R \\ F_{\bar{Y}_2^P}^R \\ F_{s_P}^R \end{Bmatrix} = \begin{Bmatrix} F_{\bar{Y}_1^P} \\ F_{\bar{Y}_2^P} \\ F_{s_P} \end{Bmatrix} - \begin{Bmatrix} F_{\bar{Y}_1^P}^{\text{int}} \\ F_{\bar{Y}_2^P}^{\text{int}} \\ 0 \end{Bmatrix} - \begin{Bmatrix} F_{\bar{Y}_1^P}^c \\ F_{\bar{Y}_2^P}^c \\ F_{s_P}^c \end{Bmatrix} \quad (33)$$

where \vec{F} represents all the external loads, \vec{F}^{int} the internal force of the sliding element and \vec{F}^c the connection constraint force. Subscripts \bar{Y}_1^P , \bar{Y}_2^P and s_P refer to the sliding node position degrees of freedom and the curvilinear position, respectively. In the definition of eq.

(32), being a quasi-null velocity case, the velocity-proportional external damping was neglected. The friction force is also not present since its value is already considered indirectly through the constraint force at the curvilinear position direction. As the tangential value of the resultant force F_R is required, the tangent vector is used to decompose the Cartesian terms as:

$$F_R = \frac{\left(F_{\bar{Y}_1^P} - F_{\bar{Y}_1^P}^{\text{int}} - F_{\bar{Y}_1^P}^c\right)\bar{T}_1^P + \left(F_{\bar{Y}_2^P} - F_{\bar{Y}_2^P}^{\text{int}} - F_{\bar{Y}_2^P}^c\right)\bar{T}_2^P}{\|\bar{T}^P\|} + F_{s_p} - F_{s_p}^c \quad (34)$$

As expected from the physical significance of the multipliers as contact forces, we have $F_{\bar{Y}_1^P}^c = \lambda_1$ and $F_{\bar{Y}_2^P}^c = \lambda_2$. This result can be obtained by developing the constraint force given in eq. (22) for the constraint equation in (7).

Time integration and nonlinear system solution procedure

For the time discretization and nonlinear system solution, the equations of motion (26) are written for a specific time instant as:

$$\bar{g}(\bar{Y}_{t+1}, \bar{\lambda}_{t+1}) = \bar{F}_{t+1}^{\text{int}} - \bar{F}_{t+1} + \mathbf{M} \cdot \ddot{\bar{Y}}_{t+1} + \mathbf{D} \cdot \dot{\bar{Y}}_{t+1} + \bar{F}_{t+1}^c + \bar{\Lambda}_{t+1}^f = \bar{0} \quad (35)$$

where \bar{g} is the residual of the Newton method (or mechanical unbalanced vector), null when \bar{Y}_{t+1} and $\bar{\lambda}_{t+1}$ are a solution of the system of equations. One can note that $\bar{\lambda}_{t+1}$ only appear in the terms \bar{F}_{t+1}^c and $\bar{\Lambda}_{t+1}^f$.

Since the description of the solid is made by a total Lagrangian approach, the inertial force is obtained using a constant mass matrix which allows the adoption of the Newmark approximations for the material velocity and acceleration vectors, see, for instance, the discussion in [23–25]. Those approximations for a time step Δt , with its usual parameters β and γ , are given by:

$$\bar{Y}_{t+1} = \bar{Y}_t + \Delta t \dot{\bar{Y}}_t + \Delta t^2 \left[\left(\frac{1}{2} - \beta \right) \ddot{\bar{Y}}_t + \beta \ddot{\bar{Y}}_{t+1} \right] \quad (36)$$

$$\dot{\bar{Y}}_{t+1} = \dot{\bar{Y}}_t + \Delta t (1 - \gamma) \ddot{\bar{Y}}_t + \gamma \Delta t \ddot{\bar{Y}}_{t+1} \quad (37)$$

Substituting both previous expressions in eq. (35) we arrive at:

$$\begin{aligned} \bar{g}(\bar{Y}_{t+1}, \bar{\lambda}_{t+1}) = & \bar{F}_{t+1}^{\text{int}} - \bar{F}_{t+1} + \left(\frac{\mathbf{M}}{\beta \Delta t^2} + \frac{\gamma \mathbf{D}}{\beta \Delta t} \right) \cdot \bar{Y}_{t+1} + \bar{F}_{t+1}^c \\ & + \bar{\Lambda}_{t+1}^f - \mathbf{M} \cdot \bar{T}_t + \mathbf{D} \cdot \bar{R}_t - \gamma \Delta t \mathbf{D} \cdot \bar{T}_t = \bar{0} \end{aligned} \quad (38)$$

in which \bar{T}_t and \bar{R}_t represents the dynamical contribution of the previous time step as:

$$\bar{T}_t = \frac{\bar{Y}_t}{\beta \Delta t^2} + \frac{\dot{\bar{Y}}_t}{\beta \Delta t} + \left(\frac{1}{2\beta} - 1 \right) \ddot{\bar{Y}}_t \quad \text{and} \quad \bar{R}_t = \dot{\bar{Y}}_t + \Delta t (1 - \gamma) \ddot{\bar{Y}}_t \quad (39)$$

For the friction force calculation, eq. (29), the Newmark expressions are employed to approximate the tangential velocity $v = \dot{s}_p$. Consequently, this force is entirely defined in a time step $t+1$ by the Lagrange multipliers, curvilinear position and current path finite element nodal positions.

Eq. (38) is clearly nonlinear for the variables $\{\vec{Y}_{t+1}, \vec{\lambda}_{t+1}\}$, thus, a usual first order Taylor expansion can be employed to obtain the Newton method as:

$$\begin{Bmatrix} \Delta \vec{Y}_{t+1} \\ \Delta \vec{\lambda}_{t+1} \end{Bmatrix} = -(\mathbf{H}_{t+1})^{-1} \cdot \vec{g}(\vec{Y}_{t+1}^0, \vec{\lambda}_{t+1}^0) \quad (40)$$

in which, the correction $\{\Delta \vec{Y}_{t+1}, \Delta \vec{\lambda}_{t+1}\}$ is obtained from the trial solution $\{\vec{Y}_{t+1}^0, \vec{\lambda}_{t+1}^0\}$ and the Hessian matrix given by:

$$\mathbf{H}_{t+1} = \nabla \vec{g}_{t+1} = \mathbf{H}_{t+1}^e + \mathbf{H}_{t+1}^c \quad (41)$$

The Hessian related to the energy potentials due to individual finite elements is called \mathbf{H}^e and its expressions can be found in [10,11]. The Hessian matrix due the constraint potential of the sliding connections is written as:

$$\mathbf{H}^c = \frac{\partial \vec{F}^c}{\partial \{\vec{Y}, \vec{\lambda}\}} = \begin{bmatrix} \vec{\lambda} \cdot \nabla (\nabla \vec{c}) & \nabla \vec{c} \\ (\nabla \vec{c})^t & \mathbf{0} \end{bmatrix} \quad (42)$$

where, $\nabla (\nabla \vec{c})$ is a third order tensor that can be understood as the set of Hessian matrices due to each constraint equation c_i , and $\mathbf{0}$ is the null matrix.

It must be stressed that, however achieved a value for s_p in the solution process, it is not sufficient to update \vec{F}^c and the Hessian matrix as the function $\xi_p = \xi(s_p)$ is not explicitly written. The solution of this stage is done by adopting a least square method to find the non-dimensional coordinate from the converged values of the path element and the sliding node as described in detail by [11]. Given the numerical value of the non-dimensional variable in the dimensionless space, the transitions among path elements is straightforward when its value exceeds the space domain.

Examples

Some examples are presented to show the capabilities of the proposed formulation regarding the correct description of the friction force and its structural effects. In all simulations, the Newmark parameters for the average acceleration in the time step were adopted, $\beta = 0.25$ and $\gamma = 0.50$, which do not introduce numerical damping in the solution.

Axial vibration with friction dissipation

To validate the improved friction model we employ the structure depicted in Fig. 7 a) which consists of a bar with length $L = 1.0\text{m}$ submitted to an initial displacement $d = 1.0\text{mm}$ at its left extremity (proportionally distributed over its extension). A cylindrical joint exists at the same end, which is free to move over a finite element with locked degrees of freedom to simulate a rigid support. A vertical force $P = 2000\text{N}$ is applied to manifest frictional effects on the connection.

Discretizing the bar with one two-noded (linear) frame element results the equivalent mass-spring system shown in Fig. 7 b). Adopting a squared cross-section with $b_0 = h_0 = 0.1\text{m}$ and Young modulus $\mathbb{E} = 2 \cdot 10^8\text{Pa}$, the axial spring stiffness is $k = \mathbb{E}b_0h_0 / L = 2 \cdot 10^6\text{N/m}$. The

shear modulus is half the value of the Young modulus. The equivalent system mass $m = 5.066 \text{ kg}$ is lumped at the joint node. Knowing all the system parameters, the mass-spring natural frequency is given by $\omega_n = \sqrt{k/m} = 628.38 \text{ rad/s}$, and its oscillation period is $T_n = 2\pi / \omega_n = 0.01 \text{ s}$. For this reason, the adopted time increment is $\Delta t = 10^{-4} \text{ s}$.

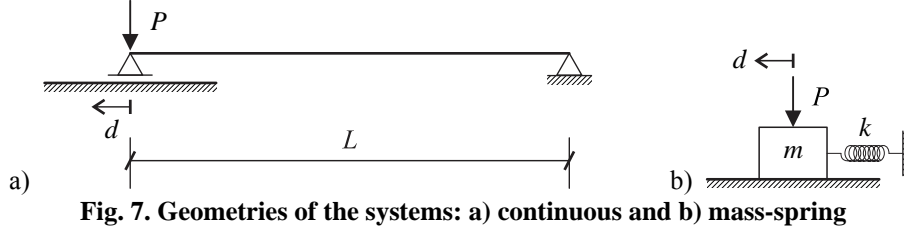


Fig. 7. Geometries of the systems: a) continuous and b) mass-spring

The sliding connection displacement for the frictionless case is shown in Fig. 8 where the harmonic oscillation with expected period and amplitude values are reached. Also in Fig. 8, two cases of friction are simulated: one with dry friction only and the other that adds the viscous friction term. Adopted dry friction parameters are $\mu_s = 0.05$, $\mu_k = 0.03$, $v_\sigma = 0.1 \text{ m/s}$ and $\delta_\sigma = 2$, for the viscous case $\eta = 100 \text{ Ns/m}$. The quasi-null speed limit was chosen as $v_0 = 2 \cdot 10^{-3} \text{ m/s}$. This mass-spring system subjected only to Coulomb kinetic friction has analytical solution presented in [26]. In spite of the reference solution have been proposed for a simpler case, one can verify in a similar manner that the friction dissipation did not altered the system oscillation period. Moreover, the decay envelope for the dry case is liner whereas when adding the viscous term the envelope changed to an exponential tendency, as is expected from its similarity to a one degree of freedom mass-spring-damper system.

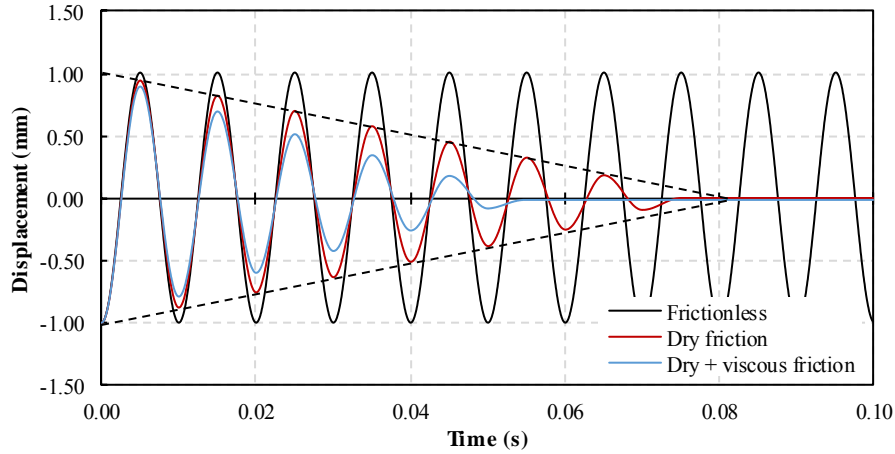


Fig. 8. Displacement of the equivalent mass-spring system

With the proposed improved friction model, the residual displacement is correctly captured as illustrated in Fig. 9. This displacement occurs when the spring restitution force, i.e., its internal force, and the friction force become balanced outside the bar undeformed configuration. This effect can only be properly represented since the resultant force is calculated in the friction model.

Given the existence of residual displacements, the friction force also has a residual value as shown in Fig. 10. For the case with viscous friction the force value at rest is -30.93 N . For

this small displacement analysis, from the spring stiffness one can obtain the residual displacement as $F_f / k = -1.5465 \cdot 10^{-2} \text{ mm}$, which is exactly the simulation value in Fig. 9.

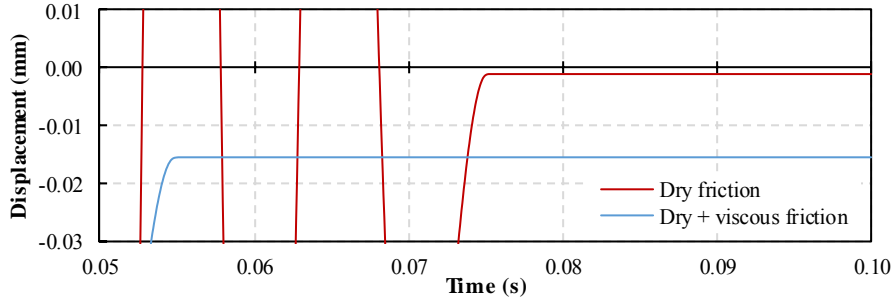


Fig. 9. Residual displacement for the equivalent mass-spring system

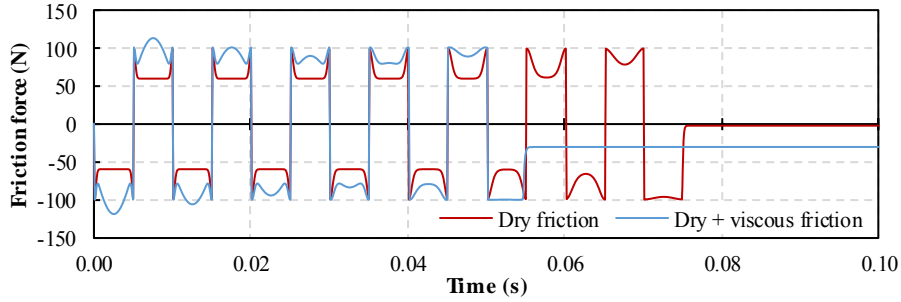


Fig. 10. Friction force for the equivalent mass-spring system

The evolution of the strain energy (S.E.) and the kinetic energy (K.E.) are also interesting parameters to be observed (Fig. 11). For the frictionless case the sum of those energies is constant throughout the analysis. We highlight that, although there is an external load applied, no energy is associated to it since there is no displacement in its direction. When friction is introduced in the system the energies sum decay with time, faster for the viscous friction case than to the dry one as is expected from the higher friction values obtained (Fig. 10). Due the existence of residual displacements, there is also a residual energy as presented in the detail of Fig. 11 in log scale. The rest energies sum of the viscous friction case is obtained in the simulation equal to 0.2392 mJ . The same value can be found from the spring strain energy $kd^2/2$, being d the residual displacement, revealing that its residual energy value is due only to the bar deformation.

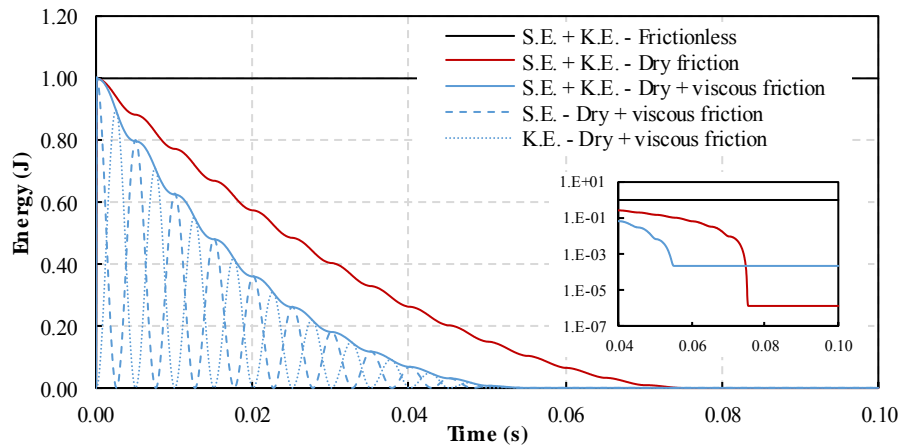


Fig. 11. Energy time history of the equivalent mass-spring system

Increasing the spatial discretization to 10 cubic finite elements, and adopting a mass density of $\rho_0 = 1250 \text{ kg/m}^3$, one achieves the same oscillation period for the mass-spring system. To analyse this discretized continuous system, all previous parameters were kept but the quasi-null speed limit $v_0 = 3 \cdot 10^{-2} \text{ m/s}$ and the time step $\Delta t = 1.25 \cdot 10^{-5} \text{ s}$. The time increment was chosen to allow a good representation of the passage of the axial displacement wave in the domain of each finite element. This wave has velocity $c = \sqrt{E b_0 h_0 / \rho_0} = 4000 \text{ m/s}$. Fig. 12 shows the displacement results for both friction cases and the frictionless one. As expected, the oscillation period, amplitude and decay type is similar to the equivalent system.

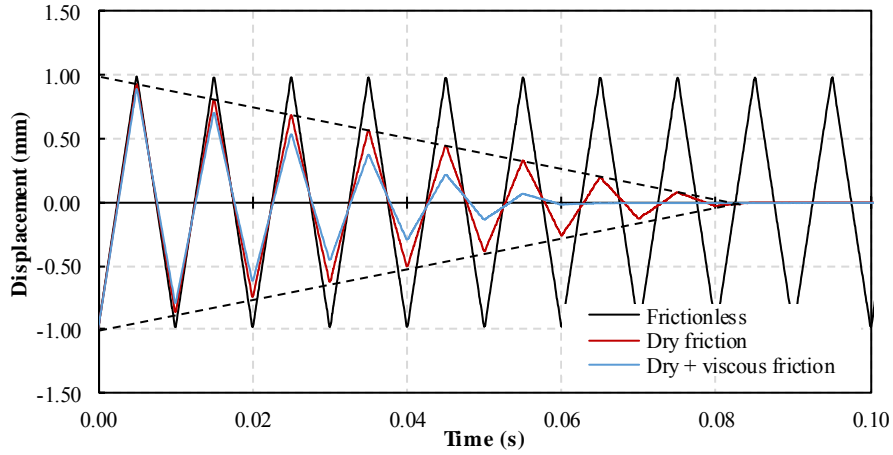


Fig. 12. Sliding connection displacement for the discretized continuous system

The improved friction model was able to represent the residual displacement of the sliding connection for this continuous system as well (Fig. 13). However, due to the continuity of the bar and the only source of dissipation to be due to the joint friction, the remainder of the bar keeps vibrating as shown for the bar mid-point displacement history in Fig. 14. The last result shows the existence of a stationary wave of axial displacement between the extremity nodes, which are at rest, one due the boundary condition and the other due the friction force.

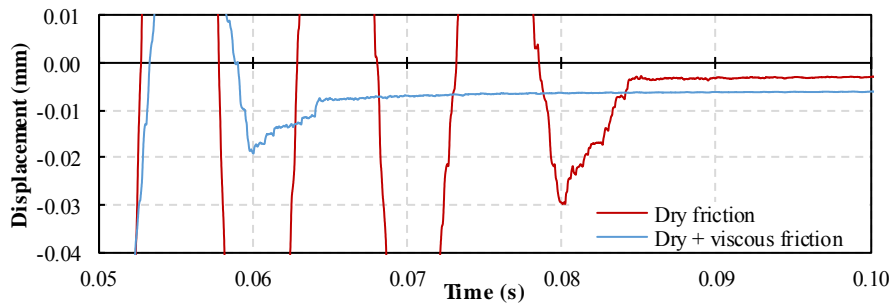


Fig. 13. Sliding connection residual displacement for the discretized continuous system

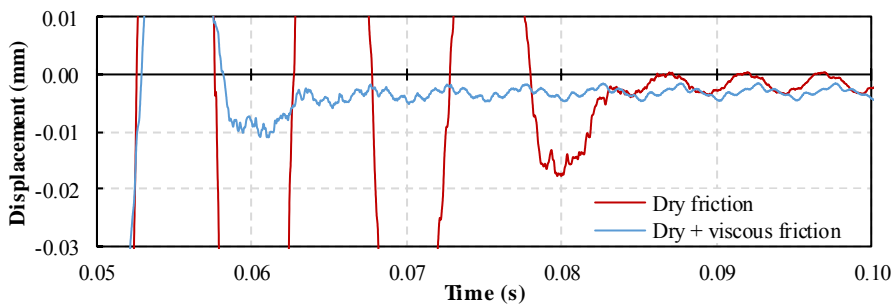


Fig. 14. Bar mid-point displacement for the discretized continuous system

Fig. 15 depicts the friction force time history for both cases. Oscillations on the value of the force during the perceptible displacements agree to the expected response of the discretized continuous system and occur due to higher vibrations modes that appear from the temporal and spatial resolutions adopted to represent properly the problem. We note that the employed model was able to capture the friction force reduction when the sliding connection is at rest and, due to the residual wave, its value shows a steady-state response that balances out the resultant force arriving from the rest of the body. Lastly, we present the energy time history for the continuous case for all the cases studied (Fig. 16). The energies sum decays in the presence of friction similarly to the mass-spring system.

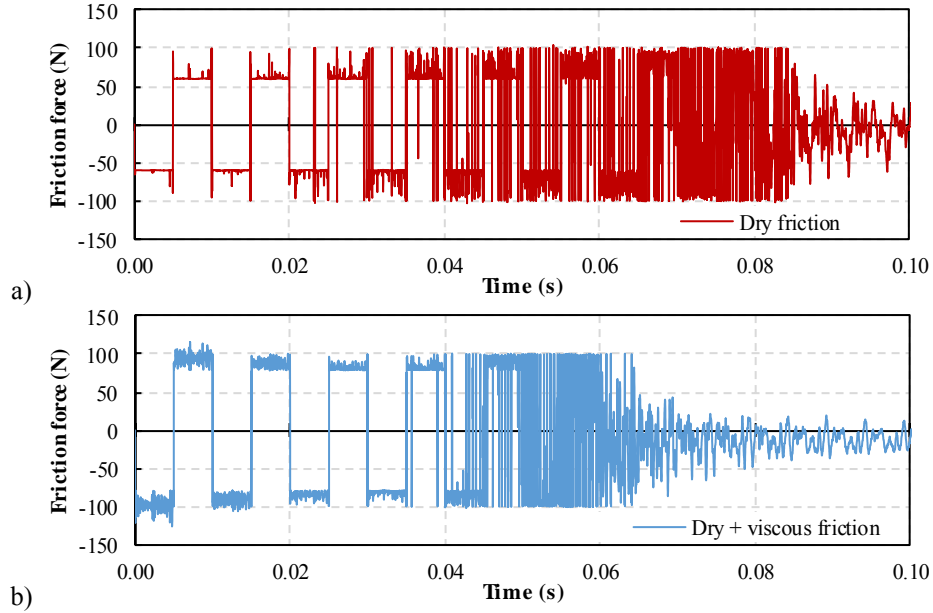


Fig. 15. Friction force for the discretized continuous system: a) dry friction and b) dry and viscous friction

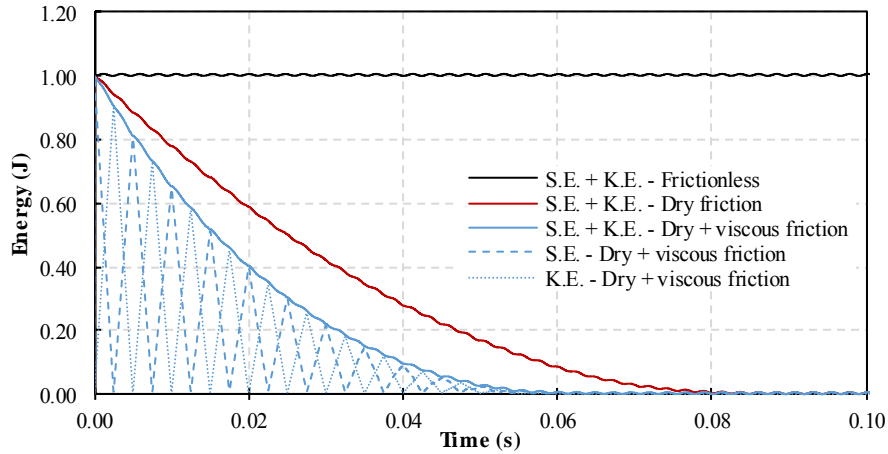


Fig. 16. Energy time history for the discretized continuous system

Driven mechanism with friction

For a more involved application, we propose the mechanism depicted in Fig. 17 subjected to a bending moment pulse in its crank. The moment M increases linearly from zero to 5 kN.m in 1 s and decreases to zero in another 1 s interval. To simulate a rigid crank its cross section is squared with 0.5 m side and its Young modulus is $2 \cdot 10^{13} \text{ Pa}$. The other bars are flexible with

squared cross section of 0.1m side and Young modulus equal $8 \cdot 10^{10}$ Pa . For all bars, the mass density is 8000 kg/m^3 and the shear modulus is half the value of the Young modulus. The adopted time step is 0.01s . Six cubic finite elements were used for the discretization. Friction parameters are: $\mu_s = 0.5$, $\mu_k = 0.3$, $v_\sigma = 0.001 \text{ m/s}$, $\delta_\sigma = 1$ and $v_0 = 1 \cdot 10^{-2} \text{ m/s}$. No viscous friction is considered in the joint.

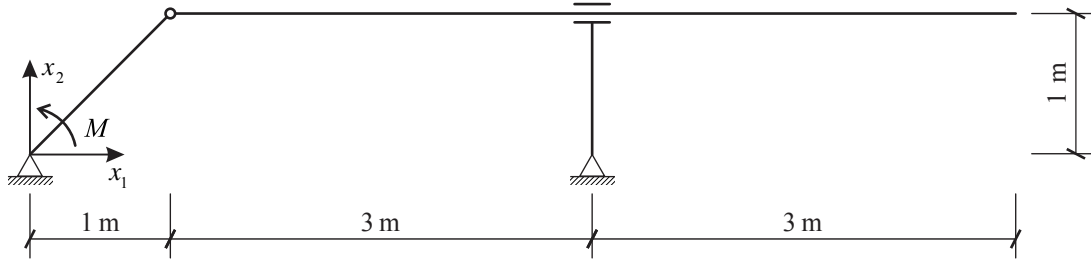


Fig. 17. Mechanism initial configuration

A prismatic joint is employed to connect the arm to a support bar. The sliding connection displacements are shown in Fig. 18 for cases with and without friction. The friction effect is perceived in the joint motion which tends towards rest after the second rotation cycle of the mechanism, while the frictionless case presents free vibrations after the loading phase.

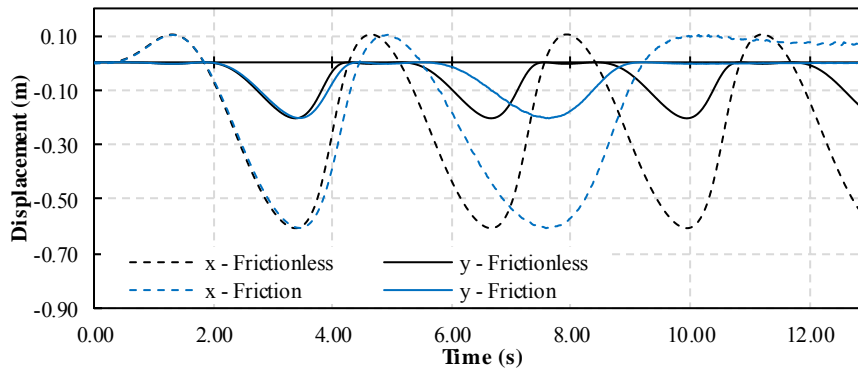


Fig. 18. Sliding connection displacements

The evolution of the curvilinear position (Fig. 19) displays similar results from the joint displacements. In Fig. 19 the position resting value is about 1.5 m from its arbitrary initial value, adopted as zero. Although expected, the curvilinear results are interesting since they can also be noted in the mechanism resting position as illustrated in Fig. 20 (the joint stops about $\frac{1}{4}$ of the arm length). The arm tip displacements (Fig. 21) also display the same behaviour.

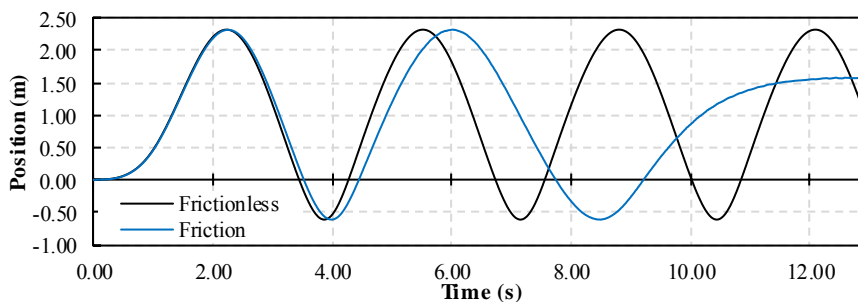


Fig. 19. Curvilinear position evolution

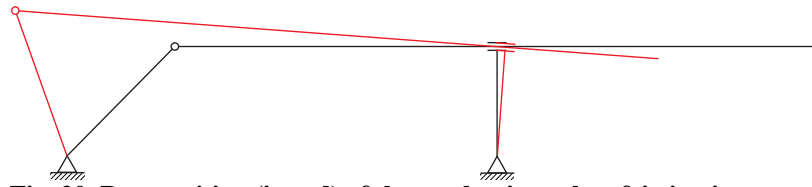


Fig. 20. Rest position (in red) of the mechanism when friction is present

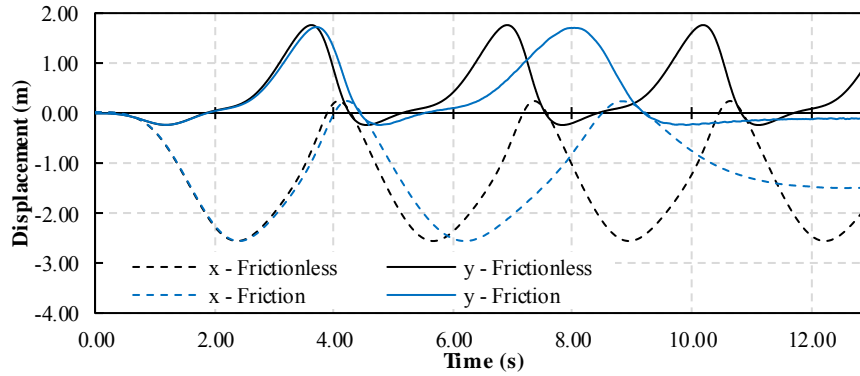


Fig. 21. Arm tip displacements

Conclusions

Friction dissipation was successfully introduced in sliding connections present in structures and mechanisms analysed by a total Lagrangian FEM formulation based on the positional description of the plane frame kinematic. Also, an improvement on the classic Coulomb friction model with Stribeck effect and viscous friction was proposed for a smoother description of the transition between motion and rest states of the joints. The proposed model was able to capture residual displacements of the body since the resultant force could be calculated properly and no instabilities were present in the friction force at null speed. Future studies intent to expand this formulation to 3D applications.

Acknowledgements

The authors would like to thank the São Paulo Research Foundation (FAPESP-2016/00622-0) for the financial support.

References

- [1] Olsson H, Åström KJ, Canudas de Wit C, Gäfvert M, Lischinsky P. Friction Models and Friction Compensation. *European Journal of Control* 1998;4:176–95. doi:10.1016/S0947-3580(98)70113-X.
- [2] Andersson S, Söderberg A, Björklund S. Friction models for sliding dry, boundary and mixed lubricated contacts. *Tribology International* 2007;40:580–7. doi:10.1016/j.triboint.2005.11.014.
- [3] Marques F, Flores P, Pimenta Claro JC, Lankarani HM. A survey and comparison of several friction force models for dynamic analysis of multibody mechanical systems. *Nonlinear Dynamics* 2016;86:1407–43. doi:10.1007/s11071-016-2999-3.
- [4] Threlfall DC. The inclusion of Coulomb friction in mechanisms programs with particular reference to DRAM au programme DRAM. *Mechanism and Machine Theory* 1978;13. doi:10.1016/0094-114X(78)90020-4.
- [5] Karnopp D. Computer simulation of stick-slip friction in mechanical dynamic systems. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME* 1985;107. doi:10.1115/1.3140698.
- [6] Armstrong-Hélouvry B, Dupont P, De Wit CC. A survey of models, analysis tools and compensation methods for the control of machines with friction. *Automatica* 1994;30. doi:10.1016/0005-1098(94)90209-7.

- [7] Ambrósio JAC. Impact of Rigid and Flexible Multibody Systems: Deformation Description and Contact Models. In: Schiehlen W, Valáček M, editors. *Virtual Nonlinear Multibody Systems SE - 4*, vol. 103, Dordrecht, Netherlands: Springer Netherlands; 2003, p. 57–81. doi:10.1007/978-94-010-0203-5_4.
- [8] Wojewoda J, Stefański A, Wiercigroch M, Kapitaniak T. Hysteretic effects of dry friction: modelling and experimental studies. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 2008;366. doi:10.1098/rsta.2007.2125.
- [9] Awrejcewicz J, Grzelczyk D, Pyryev Y. A novel dry friction modeling and its impact on differential equations computation and Lyapunov exponents estimation. *Journal of Vibroengineering* 2008;10.
- [10] Coda HB, Paccola RR. A total-Lagrangian position-based FEM applied to physical and geometrical nonlinear dynamics of plane frames including semi-rigid connections and progressive collapse. *Finite Elements in Analysis and Design* 2014;91:1–15.
- [11] Siqueira TM, Coda HB. Total Lagrangian FEM formulation for nonlinear dynamics of sliding connections in viscoelastic plane structures and mechanisms. *Finite Elements in Analysis and Design* 2017;129:63–77. doi:10.1016/j.finel.2016.12.005.
- [12] Bonet J, Wood RD, Mahaney J, Heywood P. Finite element analysis of air supported membrane structures. *Computer Methods in Applied Mechanics and Engineering* 2000;190:579–95. doi:10.1016/S0045-7825(99)00428-4.
- [13] Coda HB. An exact FEM geometric non-linear analysis of frames based on position description. . IN: 17H INTERNATIONAL CONGRESS OF MECHANICAL ENGINEERING. 2003, São Paulo, São Paulo: ABCM; 2003.
- [14] Ogden RW. *Non-linear elastic deformations*. Chichester: Ellis Horwood; 1984.
- [15] Lanczos C. *The variational principles of mechanics*. New York: Dover Publications ; 1970.
- [16] Gurtin ME, Fried E, Anand L. *The Mechanics and Thermodynamics of Continua*. New York: Cambridge University Press; 2010.
- [17] Lemaitre J, Chaboche J-L. *Mechanics of solid materials*. Cambridge, Massachusetts: Cambridge University Press; 1994.
- [18] G  radin M, Cardona A. *Flexible multibody dynamics: a finite element approach*. Chichester: John Wiley & Sons; 2001.
- [19] Nocedal J, Wright SJ. *Numerical optimization*. New York: Springer; 1999.
- [20] Luenberger DG, Ye Y. *Linear and nonlinear programming*. 3rd ed. New York: Springer; 2008.
- [21] Rao SS. *Engineering optimization: theory and practice*. 4th ed. New Jersey: John Wiley & Sons; 2009.
- [22] Bo LC, Pavelescu D. The friction-speed relation and its influence on the critical velocity of stick-slip motion. *Wear* 1982;82. doi:10.1016/0043-1648(82)90223-X.
- [23] Simo JC, Vu-Quoc L. On the dynamics in space of rods undergoing large motion - a geometrically exact approach. *Computer Methods in Applied Mechanics and Engineering* 1988;66:125–61. doi:10.1016/0045-7825(88)90073-4.
- [24] Simo JC, Vu-Quoc L. On the dynamics of flexible beams under large overall motions — the plane case. II. *Journal of Applied Mechanics-Transactions of the Asme* 1986;53:855–63.
- [25] Simo JC, Vu-Quoc L. On the dynamics of flexible beams under large overall motions — the plane case. I. *Journal of Applied Mechanics-Transactions of the Asme* 1986;53:849–54.
- [26] Rao SS. *Mechanical vibrations*. 5th ed. Upper Saddle River, N.J: Pearson Prentice Hall; 2011.

Effect of soil properties on seismic response of underground station

*Sunbin Liang¹, †Zhiyi Chen¹

¹Department of Geotechnical Engineering, Tongji University, China.

*Presenting author: 1410250@tongji.edu.cn

†Corresponding author: zhiyichen@tongji.edu.cn

Abstract

The soil dynamic properties are the main factors affecting soils' and structural seismic responses. The present paper conducts a numerical analysis to study the effect of uncertainty of elastic modulus at layered soil site using statistic method. A layered soil model is used. The elastic modulus of the concerned soil layer obeys normal distribution with a variant coefficient of 0.2. Eight input ground motions are used to avoid the calculation contingency. The analysis results show that the changing of soil elastic modulus has various influences on structural internal forces with different input ground motions. Whereas the story drift obeys quite similar normal distribution to that of the elastic modulus of the concerned soil layer. And the performance-based seismic design of underground structure should be paid well attention in engineering practice.

Keywords: Soil dynamic properties; Seismic response; Underground station; Statistic method.

1. Introduction

Currently, the soil properties, used for conduct research on geotechnical earthquake engineering, are usually determined through dynamic triaxial or resonant column tests [1-3]. However, the test results are of large-discreteness due to the complex influence factors on dynamic shear modulus and damping ratio. Based on earthquake damage investigation and site seismic response analyses, it is well known that the soil dynamic properties, including dynamic shear modulus and damping ratio, are the main factors affecting soils' seismic responses. Thus, whether the soil dynamic properties are consistent with the actual situation or not has distinct influences on the reliability of analysis results [4-5].

Wen et al. [6] studied on seismic soil-structure interaction and found that the high frequency seismic response spectra is abundant in the hard rock layers in the Middle East of the United States. Dashti et al. [7] conducted centrifuge tests on underground reservoir structures to study the influence of backfill soil type. The test results showed that soft and hard soils have distinct influences on structural dynamic behavior and dynamic earth pressure. Zlatanoviä et al. [8] revealed that the soft clay soil, compared with the sand of medium-compactness, is of smaller shear stiffness and larger damping. And more earthquake energy is dissipated in soft soil, which leads to a weaker amplification effect of soils on seismic wave, and then results in smaller seismic shear force and larger seismic shear strain. Thus, the structure buried in soft soils suffers smaller axial force and larger shear force and bending moment than that in medium-compactness sands. However, the above-mentioned studies, based on the one-layer isotropic soil model, mainly focused on the effect of changing soil types, and the uncertainty of soil dynamic properties have not been considered.

To study the uncertainty of soil dynamic properties, some researchers [9-11] used the one-dimensional equivalent linearization method by adopting a specific variable coefficient to simplify the soil models. Obviously, this method has certain limitations, such as the virtual resonance effect [12], which is an inherent error. At layered sites, considering the uncertainty of soil properties, the calculations are complicated because different kinds of soil properties increase or decrease in different soil layers [13]. An alternative method based on the random vibration theory has been used to simulate dynamic soil properties [14-15]. However, this method is complex and unsuitable for widespread engineering applications in evaluating the seismic performances of underground structures.

The present paper assumes that the elastic modulus of the concerned soil layer obeys the normal distribution with the variation coefficient of 0.2, and this hypothesis is well adopted in previous studies [16-19]. Taking a typical underground station in Shanghai as an engineering reference, the influences of uncertainty of elastic modulus at layered soil site on structural seismic responses are studied by using the statistic method. To avoid the calculation contingency of using only one input ground motion, eight input ground motions are used in the present paper.

2. Finite element modeling

2.1 Numerical model

The finite element code ABAQUS [20] is used to perform the full-time history analyses of underground structure with surrounding soils. Fig. 1 shows the cross-section dimension of the station and its central column. The numerical analyses are performed under plane strain condition. Fig. 2 depicts the analytical model with the area of 500 m × 60.55 m. When considering the boundary effect, as recommended by the Code for Seismic Design of Buildings [21], the width of each side of soil around the structure should be at least triple of the width of the structure. In the present model, the width of each side of soil is 250 m. Further, the infinite element boundary is used as the side boundary. The infinite element (CINPE4) provided by ABAQUS is based on the static analysis [22] and the dynamic response analysis [23], and it can simulate no reflection by setting damping on the boundary. The bottom boundary is placed 45.28 m from the bottom plate of the structure, on which the X and Y displacements are fixed before the ground motion is input.

2.2 Constitutive models and material properties

Quadratic plane strain elements (CPE4R) and beam elements (B21) are used to simulate the soil and the station structure, respectively (Fig. 2). Soil behavior is modelled by Mohr-Coulomb elasto-plastic model, and the specific properties are shown in Table 1. There are three soil layers, which are artificial fill, silty clay and gray clay from top to bottom. And the station structure is cased in the silty clay.

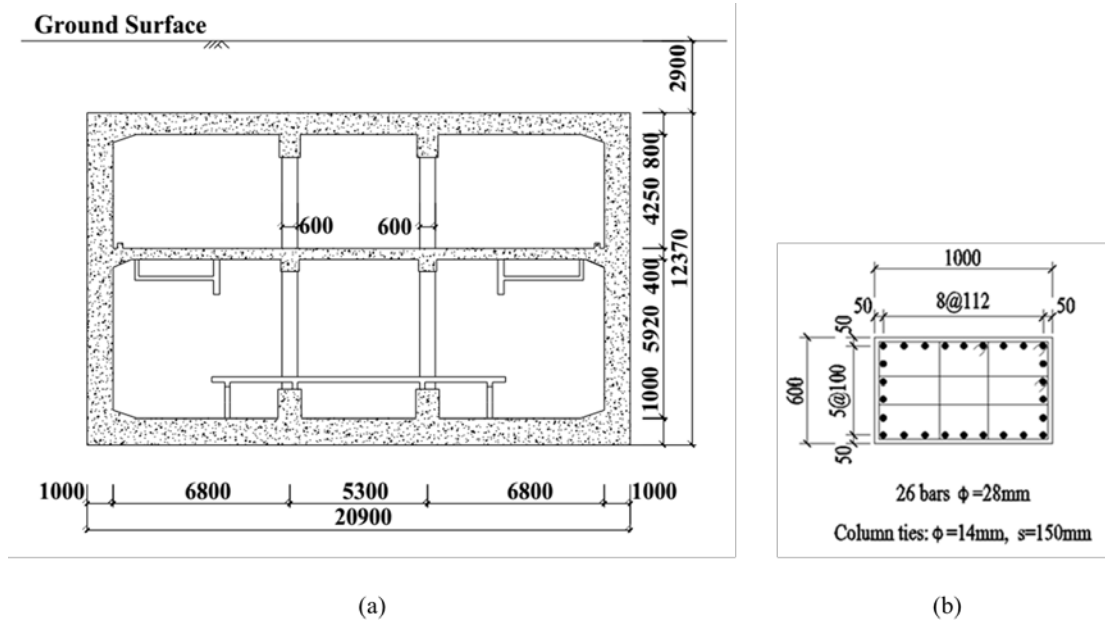


Fig. 1. Cross-section dimension of the station and central column

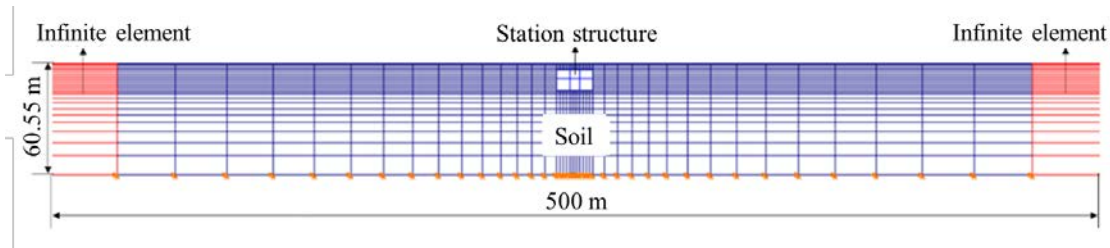


Fig. 2. Numerical analytical model

Table 1. Soil properties of station site

Layer num.	Soil type	Depth (m)	Unit weight (kN/m ³)	Elastic modulus (MPa)	Poisson's ratio ν	Internal friction angle (°)	Cohesion (kPa)
1	Artificial fill	0-1.32	19.00	20.34	0.32	15.0	20.0
2	Silty clay	1.32-19.8	17.72	13.32	0.34	27.7	13.0
3	Gray clay	19.8-60.55	18.10	24.13	0.32	30.3	7.0

The concrete of Grade C45 and Grade C35 [24] are used to build the central columns and the other parts. For concrete C45, its elastic modulus, Poisson's ratio, tensile, and compression strength are 33.5 GPa, 0.2, 2.51 MPa, and 29.6 MPa, respectively. And for concrete C35, its parameters are 31.5 GPa, 0.2, 2.20 MPa, and 23.4 MPa, respectively. To better simulate the

dynamic response of the elastic-plastic stage, the concrete damaged plasticity model is used. And the detailed information of this model and its calculation of damage parameters have been stated in the previous studies [25-26]. Since the longitudinal spacing between columns is 8 m in practical engineering, the reduced stiffness is adopted to consider the spacing [27].

Bilinear isotropic model (idealized elastic-plastic model), following the kinematic hardening rule, is selected to simulate the rebar. Fig. 1(b) shows the reinforcement details of the central column. Rebar HRB400 [24] is used in the present structure with the elastic modulus and yield strength of 200 GPa and 400GPa, respectively.

To simplify the analyses, no-slip condition is assumed for the soil-structure interaction. Although the interface behavior is quite crucial for the dynamic response of underground structures [28-29], this assumption is quite common in engineering practice, as it can be treated as the upper limit for the developed shear stresses around the tunnel [30].

3. Calculation cases and input ground motions

3.1 Calculation cases

Layered homogeneous soil model is adopted in the present paper. The soil properties of Layers 1 and 3 are set as listed in Table 1. There are five calculation cases classified by the elastic modulus of Layer 2, in which the station structure is cased. Table 2 depicts these five calculation cases with details of the elastic modulus of Layer 2. Based on the practical experience and some previous studies [16-19], the elastic modulus of Layer 2 obeys the normal distribution, and the variation coefficient of the mean elastic modulus is set as 0.2. The standard deviation is the product of the mean elastic modulus and the variation coefficient.

Table 2. Calculation cases and corresponding elastic modulus of Layer 2

Case number	Standard deviation (MPa) *	Elastic modulus (MPa)
Case1		7.99 ($\mu-2\sigma$)
Case2		10.66 ($\mu-\sigma$)
Case3	2.67 (σ)	13.32 (μ)
Case4		15.98 ($\mu+\sigma$)
Case5		18.65 ($\mu+2\sigma$)

*Standard deviation (σ) = Variation coefficient \times Mean elastic modulus (μ)

3.2 Input ground motions

There are eight ground motions used as base excitation. The time history data of the input ground motions are from the Pacific Earthquake Engineering Research Center (PEER) in the United States. The detailed information of input ground motions is shown in Table 3. The

predominant frequency of the input ground motions can be classified into three categories as low-, medium-, and high-frequency. Due to space limitation, only two acceleration time histories of Landers and Hector Mine are depicted as in Fig. 3. The peak ground acceleration of the input ground motions is set as 0.1 g. According to the Code for Seismic Design of Buildings [21], a PGA of 0.1 g would correspond to a probability of exceedance of 10% in 50 years for the underground station.

Table 3. Information of input ground motions

Ground motion	Earthquake	Date	Recording stations	Direction	Predominant period (s)	Duration (s)
EQ-1	Landers	28 Jun 92	NO. 11628	90	0.18	134.96
EQ-2	Chichi	20 Sep 99	CHY042	NS	0.22	89.995
EQ-3	Chichi	22 Sep 99	CHY035	EW	0.32	89.995
EQ-4	SMART1	14 Nov 86	SMART1012	EW	0.34	40
EQ-5	Chichi	25 Sep 99	CHY027	EW	0.36	89.995
EQ-6	Northridge	17 Jan 94	NO.14560	90	0.42	60
EQ-7	Hector Mine	16 Oct 99	NO.11628	90	0.42	100
EQ-8	Chichi	20 Sep 99	CHY042	EW	0.8	89.995

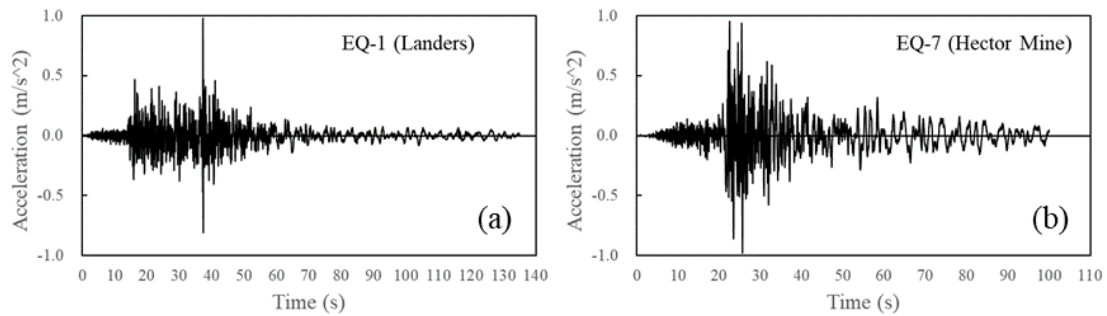


Fig. 3. Time histories of input ground motions, such as (a) Landers; (b) Hector Mine

4. Results and discussion

In engineering practice, the primary indexes, used to conduct seismic performance evaluation on underground stations, include the peak shear force and bending moment of columns and story drift [27, 31]. Tables 4, 5 and 6 show the peak shear force and bending moment of columns and story drift, respectively. And the corresponding numerical characteristics, such as mean value and variation coefficient are also calculated as shown in Tables 4-6. Due to space limitation, only the seismic responses of upper floor are shown herein since the maximum values basically occurs in this floor. And it should be noted that the findings of dynamic behavior of the lower floor are the same as the upper one.

Seen from Tables 4 and 5, it can be easy to find that the changing of soil elastic modulus has various influences on structural internal forces with different input ground motions. For example, in condition EQ-3, the maximum differentials of peak shear force and bending moment of upper column can reach 56% and 57%, respectively. In conditions EQ-4 – EQ-8, the variation coefficient of internal force ranges within 0.01-0.06, which implies the changing of soil elastic modulus barely influences the internal force. And the detailed explanations will be illustrated in the following.

Table 4. Peak shear force of upper column

Ground motion	Peak shear force of upper column (kN)						Variation coefficient*
	Case1	Case2	Case3	Caes4	Case5	Mean value	
EQ-1	567.8	546.6	491.9	651.6	517.8	555.1	0.11
EQ-2	862.9	1060.0	1044.0	848.8	836.2	930.4	0.12
EQ-3	305.2	344.2	374.2	428.8	475.5	385.6	0.17
EQ-4	1136.5	1121.2	1148.2	1146.8	1146.8	1139.9	0.01
EQ-5	1143.0	1128.2	1128.2	1159.0	1103.4	1132.4	0.02
EQ-6	1058.9	1050.2	1078.7	1102.4	1060.3	1070.1	0.02
EQ-7	924.14	1095.3	1025.6	1053.6	1068.9	1033.5	0.06
EQ-8	1085.6	1059.0	1055.9	1078.1	1052.4	1066.9	0.01

*The variation coefficient is a dimensionless value (the same below).

Table 5. Peak bending moment of upper column

Ground motion	Peak bending moment of upper column (kN•m)						Variation coefficient
	Case1	Case2	Case3	Caes4	Case5	Mean value	
EQ-1	1082.7	1046.2	941.5	1226.6	989.9	1057.4	0.10
EQ-2	1648.3	2047.1	1981.8	1629.8	1620.0	1785.4	0.12
EQ-3	577.8	656.6	711.8	821.6	908.6	735.3	0.18
EQ-4	2044.4	2014.5	2086.5	2089.9	2110.2	2069.1	0.02
EQ-5	2141.1	2160.3	2153.3	2106.4	2140.6	2140.4	0.01
EQ-6	2039.4	2026.4	2055.7	2120.2	2047.5	2057.8	0.02
EQ-7	1759.2	2022.4	1954.3	2023.8	2031.2	1958.2	0.06
EQ-8	2021.7	2029.8	2011.8	2062.5	2004.0	2026.0	0.01

Table 6 shows the peak upper story drift and the corresponding numerical characteristics.

Seen from Table 6, it can be concluded that changing the elastic modulus has obvious influence on story drift. And the variation coefficient is basically about 0.2 under all kinds of input excitations, which is the variation coefficient of the elastic modulus of the concerned soil layer. This phenomenon can be explained through Fig. 4. As illustrated in Fig. 4, the deformation of the surrounding soil, rather than the vibration characteristics of the structure, dominates the seismic response of an underground structure embedded in soft soils. Thus, the story drift obeys quite similar normal distribution to that of the elastic modulus of the concerned soil layer, which the structure is cased in.

Table 6. Peak upper story drift

Ground motion	Peak upper story drift						Mean value	Variation coefficient
	(‰)							
	Case1	Case2	Case3	Caes4	Case5			
EQ-1	2.14	1.92	1.42	2.06	1.25	1.76	0.23	
EQ-2	2.54	3.94	3.47	2.72	2.53	3.04	0.21	
EQ-3	1.09	1.31	1.61	1.50	1.20	1.34	0.16	
EQ-4	8.66	7.29	6.24	5.66	4.74	6.52	0.23	
EQ-5	18.47	19.16	14.09	9.05	7.40	13.64	0.39	
EQ-6	5.29	3.52	5.19	3.53	3.16	4.14	0.25	
EQ-7	3.48	4.94	4.40	4.57	2.94	4.07	0.20	
EQ-8	6.34	6.59	5.04	2.89	3.29	4.83	0.35	

Note: The upper limit value recommended by Code is 4‰ [21].

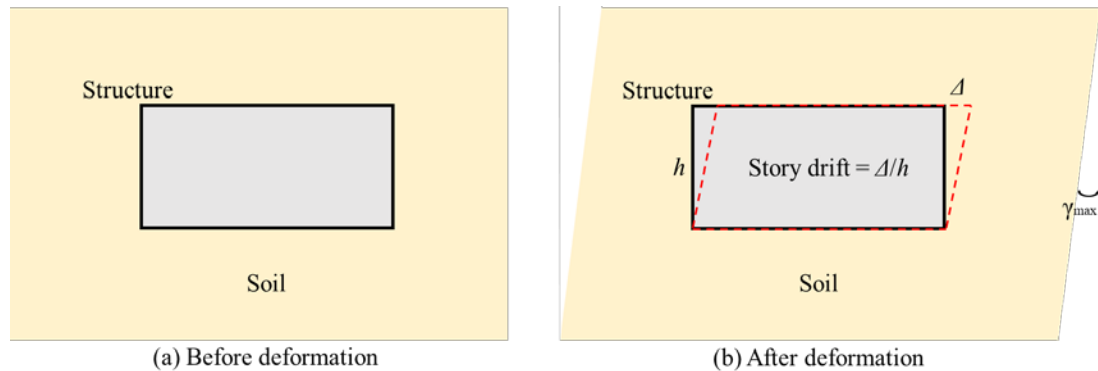


Fig. 4. Racking deformation mode of underground structure

According to the Code [21], the upper limit value recommended is 4‰. As shown in Table 6, the drift story could exceed the upper limit under several cases, when using input ground motions like EQ-4 – EQ-8. And under these excitations, the values of story drift vary significantly (see Table 6), whereas the values of internal forces are very close (see Tables 4 and 5). It implies that the structure is in the plastic state. And this could explain that the

internal forces under these conditions barely change with the changing of soil elastic modulus, whose variation coefficient of internal force ranges within 0.01-0.06. From the above-discussion, it also can be concluded that the index, like internal forces, cannot comprehensively reflect the dynamic behavior state and seismic performance of underground structures when entering severe plastic stage. Thus, the performance-based seismic design of underground structure should be paid well attention in engineering practice.

5. Conclusion

In this paper, a numerical study is conducted to explore the influences of uncertainty of elastic modulus at layered soil site on structural seismic responses through the statistic method. And there are eight input ground motions for avoiding the calculation contingency. Some conclusions can be drawn as follows:

1. The changing of soil elastic modulus has various influences on structural internal forces with different input ground motions. Under EQ-4 – EQ-8 conditions, the changing of soil elastic modulus has little influences on the internal force since the structure enters plastic stage.
2. The story drift obeys quite similar normal distribution to that of the elastic modulus of the concerned soil layer, which the structure is cased in. Because The deformation of the surrounding soil dominates the seismic response of an underground structure embedded in soft soils.
3. Internal forces cannot comprehensively reflect the dynamic behavior state and seismic performance of underground structures when entering severe plastic stage. The performance-based seismic design of underground structure should be paid well attention in engineering practice.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No. 41472246, 51778464), Key laboratory of Transportation Tunnel Engineering (TTE2014-01), and “Shuguang Program” supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission. All supports are gratefully acknowledged.

References

- [1] Chen, G. X., Xie, J. F., Han, W. and Zhang, K. X. (1995) A simplified effective stress method of soil mass earthquake response analysis, *Earthquake Engineering and Engineering Vibration* **15**, 52-61.
- [2] Sun, J. and Yuan, X. M. (2003) A state-of-art of research on dynamic modulus and damping ratio of soils, *World Earthquake Engineering* **1**, 88-95.
- [3] Chen, G. X., Liu, X. Z., Zhu, D. H. and Hu, Q. X. (2006) Experimental studies on dynamic shear modulus ratio and damping ratio of recently deposited soils in Nanjing, *CHN Journal of Geotechnical Engineering* **28**, 1023-1027.
- [4] Seed, H. B., Wong, R. T., Idriss, I. M. and Tokimatsu, K. (1986) Modulus and damping factors for dynamic analyses of cohesionless soils, *Journal of Geotechnical Engineering* **112**, 1016-1032.

- [5] Rollins, K. M., Evans, M. D., Diehl, N. B. and III Daily, W. D. (1998) Shear modulus and damping relationships for gravels, *Journal of Geotechnical and Geoenvironmental Engineering* **124**, 396-405.
- [6] Wen, S. T., Lilhanand, K., Hamasaki, D., Garcia, J. A. and Srinivasan, R. (2014) Seismic soil-structure interaction with consideration of spatial incoherence of seismic ground motions: a case study, *Nuclear Engineering & Design* **269**, 200-206.
- [7] Hushmand, A., Dashti, S., Davis, C., Hushmand, B., McCartney, J., Hu, J. and Lee, Y. (2016) Seismic performance of underground reservoir structures: insight from centrifuge modeling on the influence of backfill soil type and geometry, *Journal of Geotechnical & Geoenvironmental Engineering* **142**, 04016058.
- [8] Zlatanoviä, E., Lukiä, D. Ä., Proloviä, V., Boniä, Z. and Davidoviä, N. (2015) Comparative study on earthquake induced soil tunnel structure interaction effects under good and poor soil conditions, *European Journal of Environmental and Civil Engineering* **19**, 1000-1014.
- [9] Lan, J. Y. (2006) The effect of the soil dynamics parameters on the design response spectrum of site, Institute Engineering Mechanics CNH Earthquake Administrastion, Harbin, China.
- [10] Chen, G. X., Liu, X. Z. and Wang, B. H. (2007) Effect of variability of soil dynamic parameters on ground motion parameters for deep soft sites, *Journal of Disaster Prevention and Mitigation Engineering* **27**, 1-10.
- [11] Liu, X. J. (2007) Effect of soil dynamic feature on ground motion and its testing standard, Institute Engineering Mechanics CNH Earthquake Administrastion, Harbin, China.
- [12] Chen, G. X., Xie, J. F., Han, W. and Zhang, K. X. (1995) A simplified effective stress method of soil mass earthquake response analysis, *Earthquake Engineering and Engineering Vibration* **15**, 52-61.
- [13] Lou, M. L., Yan, G. X., Shen, J. W. and Wen, F. (2004) Effect of variability of dynamic parameters of soft soil in Shanghai region on seismic response of layered soil, *Rock and Soil Mechanic* **25**, 1368-1372.
- [14] Zeng, X. C. and Qin, X. J. (1998) Analysis of random response of soil layer to earthquake, *Earthquake Engineering and Engineering Vibration* **18**, 27-39.
- [15] Chen, Y. and Li, J. (2006) Numerical simulation of coherency function of nonlinear ground motion under consistent excitations, *Journal of Disaster Prevention and Mitigation Engineering* **26**, 369-376.
- [16] Li, T. and Li, J. (1994) Site seismic response analysis with stochastic parameters, *CHN Journal of Geotechnical Engineering* **5**, 79-83.
- [17] Li, J. and Liao, S. T. (2002) The analysis of coherency function of earthquake ground motion considering stochastic effect in site media, *CHN Journal of Geotechnical Engineering* **6**, 685-689.
- [18] Chen, Y. (2006) Seismic response analysis of engineering site with random media, *Engineering Construction* **4**, 5-10.
- [19] Bi K. and Hao H. (2011) Influence of irregular topography and random soil properties on coherency loss of spatial seismic ground motions, *Earthquake Engineering and Structural Dynamics* **40**, 1045-1061.
- [20] ABAQUS Inc. (2014) ABAQUS/Analysis user's manual-version 6.14., Providence, RI, USA.
- [21] GB50011 (2010) Code for Seismic Design of Buildings, China Architecture & Building Press, Beijing, China.
- [22] Zienkiewicz, O. C., Bo, K., Bettess, P., Emson, C. and Chiam, T. C. (1991) Mapped infinite elements for exterior wave problems, *International Journal for Numerical Methods in Engineering* **32**, 207-209.
- [23] Lysmer, J. and Kuhlemeyer, R. L. (1969) Finite dynamic model for infinite media, *Journal of the Engineering Mechanics Division* **95**, 859-878.
- [24] GB50010 (2010) Code for Design of Concrete Structures, China Architecture & Building Press, Beijing, China.
- [25] Chen, Z. Y., Chen, W., Li, Y. Y. and Yuan, Y. (2016) Shaking table test of a multi-story subway station

- under pulse-like ground motions, *Soil Dynamic Earthquake Engineering* **82**, 111-122.
- [26] Chen, Z. Y. and Liu, Z. Q. (2018) Effects of central column aspect ratio on seismic performances of subway station structures, *Advances in Structural Engineering* **24**, 14-29.
- [27] Chen, Z. Y., Chen, W. and Bian, G. Q. (2014) Seismic performance upgrading for underground structures by introducing shear panel dampers, *Advances in Structural Engineering* **17**, 1343-1358.
- [28] Sedarat, H., Kozak, A., Hashash, Y. M. A., Shamsabadi, A. and Krimotat, A. (2009) Contact interface in seismic analysis of circular tunnels, *Tunnelling and Underground Space Technology* **24**, 482-490.
- [29] Tsinidis, G., Pitilakis, K. and Trikalioti, A. D. (2013) Numerical simulation of round robin numerical test on tunnels using a simplified kinematic hardening model, *Acta Geotechnica* **9**, 641-659.
- [30] Pitilakis, K., Tsinidis, G., Leanza, A. and Maugeri, M. (2014) Seismic behaviour of circular tunnels accounting for above ground structures interaction effects, *Soil Dynamic Earthquake Engineering* **67**, 1-15.
- [31] Samata, S., Ohuchi, H. and Matsuda, T. (1997) A study of the damage of subway structures during the 1995 Hanshin-Awaji earthquake, *Cement and Concrete Composites* **19**, 223-239.

Nomograph to calculate amount of reinforcing bar against bending moment in Circular Void slabs

Shigehiro MOROOKA

Department of Architecture and Building Engineering, Tokai University, Japan
morooka@tokai-u.jp

1. Introduction

Void slab has an advantage of reducing weight and increasing stiffness. In 1970s, "structural design of void slab"[1] was published by prof. Gengo Matsui, also well known as a famous structural engineer. In the published book, void slab that has long circular steel pipes etc. is treated as one way structure and is recommended to design like typical beams. For that reason, circular voids were arranged in orthogonal direction against slab boundaries and no stresses are thought to occur at supported beams whose directions are parallel to the void. So the author indicated that the stiffness of slab in orthogonal direction of voids [2] and clarified elastic characteristics occurred in such structures [3]. In these papers, such structures which have cylindrical steel pipes or other materials in one direction can be estimated as an orthotropic plate, and the bending stiffness and the stress concentration, which is clarified to be related to the ratio of the diameter to the thickness of the slab, are presented.

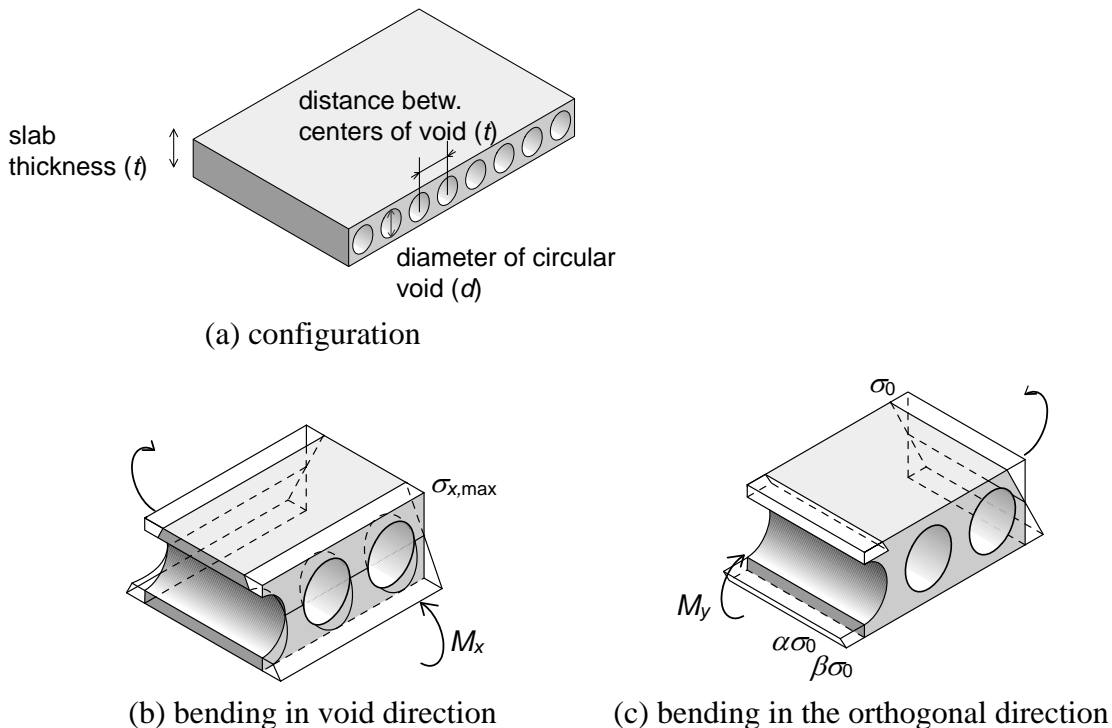


Figure 1. Applicable void slab and stress distribution in each direction

But the distribution of stress in the paper was calculated on the assumption of the linearly elastic isotropic material. As shown in fig. 1 (b), distribution of normal stress in void direction can be

formulated based on Navier's hypothesis, but that in the orthogonal direction is not straight as shown in fig. 1 (c). Figure 2 shows the distribution of normal stress in the orthogonal direction on upper and lower section of circular void calculated by BEM in previous paper. And the author had shown the scaling factor α and β to find the stresses on the surface of slab and the upper edge of the void.

As these results are obtained under the assumption that the slab has no steel bar and concrete has tensile strength, there exists stress distribution in tensile side of the concrete. Consequently, when finding amount of reinforcing bar from these results, resultant stress in the tensile side would be used.

On the other hand, amount of reinforcing bar in typical RC beam is formulated on the assumption that concrete has no tensile strength and Navier's hypothesis [4]. But the hypothesis cannot be established in void slab as shown in fig. 3, so the formulation to find the amount of reinforcing bar in the same method is difficult.

Accordingly, in this paper, amount of reinforcing bar for void slab is numerically calculated by FE analysis.

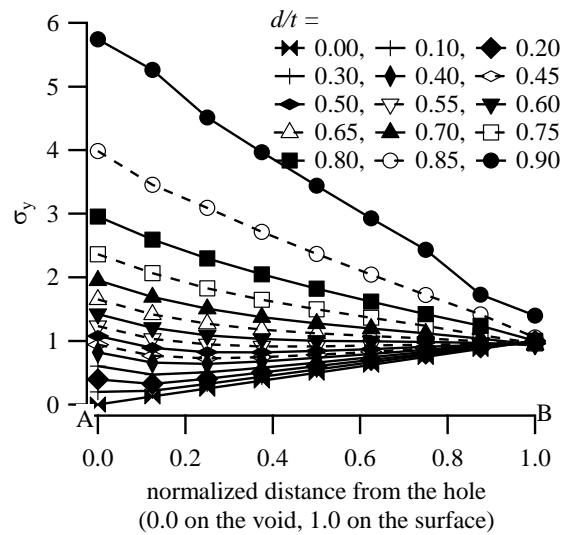


Figure 2. Normal stress distribution on upper and lower section of circular void

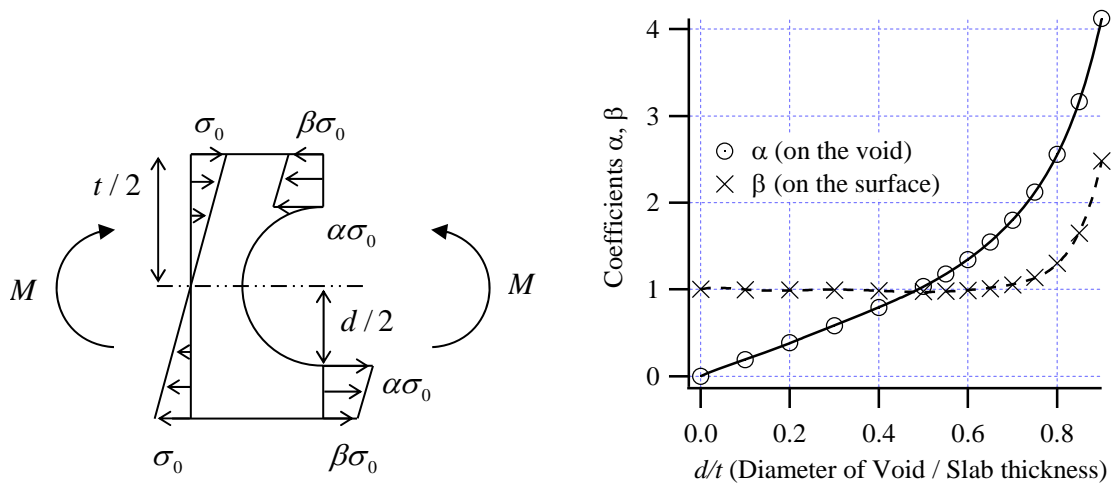


Figure 3. Scaling factor a and b to find the stresses on edge of void and on surface of slab

2. Void Slab configuration and Distribution of normal stress

Circular void slab treated in this paper is shown in fig. 1 (a). Cylindrical steel pipes or other materials are embedded in the center part of cross section in one direction. The distance between steel pipes is same as the thickness of the slab and the diameter of the pipes is equal to the thickness minus 120 mm, which comes from cover depth of reinforcing bars and the pipes.

Figure 4 shows the distribution of normal stress in concrete and tension force in reinforcing bars in typical RC beam. a_{sc} and a_{st} are area of compression reinforcement and that of tension reinforcement, and d_c is distance between extreme compression fiber and center of compression reinforcement. By using these variables in this figure, allowable bending moment for RC beams is formulated in AIJ standard for Structural Calculation of Reinforced Concrete Structures in Japan. Assumption of the neutral axis position x_n gives the maximum stresses σ_c of compressive concrete and σ_{st} of tensile steel bar. And the equations of allowable bending moment are obtained by the maximum stress reaching the allowable strength.

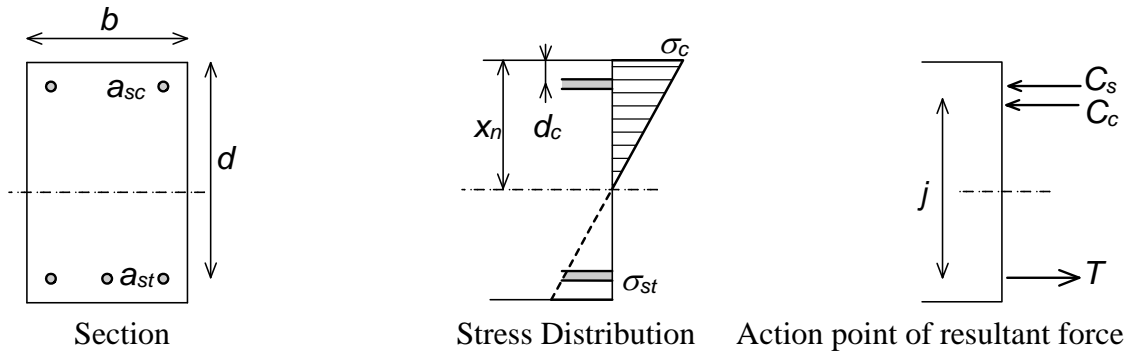


Figure 4. RC beam [4]

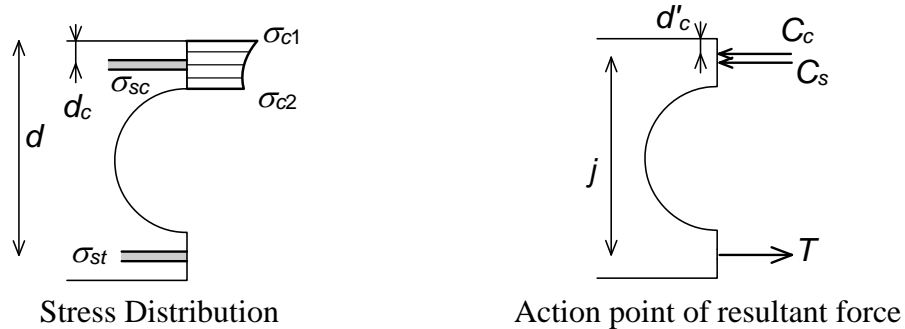


Figure 5. Void slab

By considering in the same manner, the stress distribution and resultant forces around the circular hole would be illustrated as in fig. 5. The maximum compressive stress in the concrete occurs at the surface of the slab or at the top of circular void (σ_{c1} or σ_{c2}).

Similarly in formulation for RC beam, finding resultant force C_c and its acting position from distribution of compressive concrete stress, and the distance j between acting position of composite stress of compressive concrete stress C_c and compressive steel stress C_s and the center of tensile reinforcing bar gives the following two bending moment equations.

$$M_c = C_s(d - d_c) + C_c(d - d'_c)$$

$$M_t = T \cdot j$$

Considering cover depth of concrete, we assume that $d = t - 30$ [mm] and $d_c = 30$ [mm]. Bending moments in which stress reaches its allowable value can be written in

$$M_{ci} = M_c \times \frac{f_c}{\sigma_{ci}} \quad (i=1, 2), \quad M_{c3} = M_t \times \frac{f_t}{\sigma_{st}}$$

where f_c and f_t are allowable compressive stress of concrete and tensile stress of steel bar. Dividing them by bd^2 gives the coefficients $C_1 \sim C_3$ which can be similarly seen in AIJ standard for RC beam.

$$C_1 = \frac{M_{c1}}{bd^2}, \quad C_2 = \frac{M_{c2}}{bd^2}, \quad C_3 = \frac{M_{c3}}{bd^2}$$

As the distribution of stress is dependent on the ratio of the diameter to the thickness of the slab, we need to calculate amount of reinforcing bar for each thickness individually. The thickness of slab is to be set from 200 to 500 mm at 50 mm intervals. Tension reinforcement ratio pt (= section area of tension reinforcement / that of slab) is to be set from 0.2 to 2.0 % at 0.2 % intervals and double reinforcement ratio γ (= section area of compressive reinforcements / that of tension reinforcements) is to be set from 0.0 to 1.0 at 0.2 intervals.

3. FE models

Figure 6 shows analytical model to calculate distribution of normal stress of concrete and tension and compression stresses in reinforcing bars. Simply supported beam which has seven circular holes is loaded with constant load. Firstly, to confirm that the same distributions of normal stress indicated in fig. 2 in a section can be obtained, finely divided 2D model consists only of concrete is used. Numerical models are built by using plane elements. After that, by appending beam elements for reinforcing bars, RC models are formed.

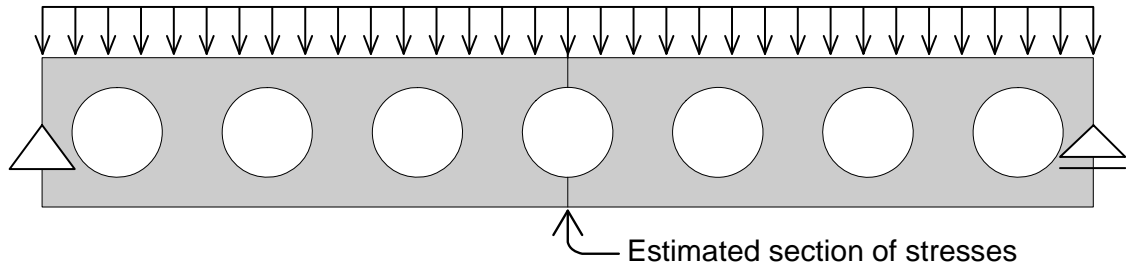


Figure 6. Analytical model and section of estimating stress

In the beginning, analytical models as shown in fig. 7 whose concrete elements around the estimated section are deleted were calculated to confirm the stress distribution. But the distributions are totally different by changing the nodal positions that connect the tension reinforcing elements and the concrete elements, and the models turn to be insufficient. Therefore, analytical models as shown in fig. 8 whose all concrete elements in lower part of

beam are deleted and stiff elements are appended to let tension reinforcing bars be able to be tensioned in adequate distance.

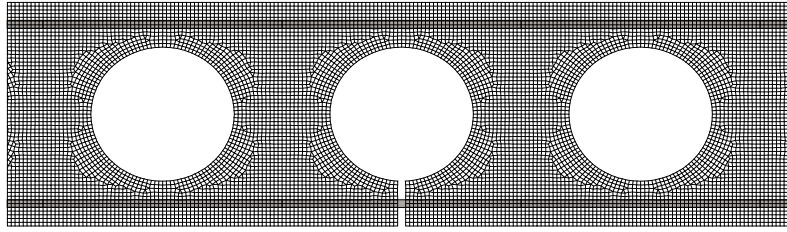


Figure 7. 1st FE model

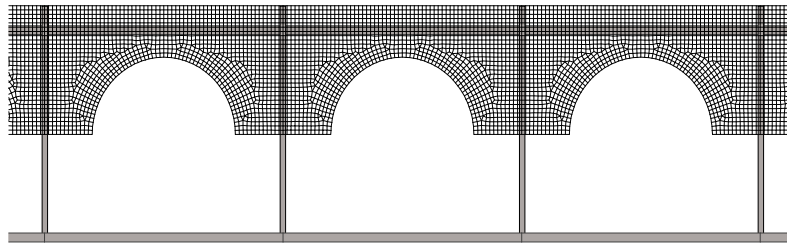


Figure 8. Analytical FE model

4. Results and conclusions

Normal stress distribution and nomograph to find allowable bending moment are illustrated in figures 9 ~ 11 in each thickness of slab. Left side of figs. is normal stress distribution and right one is the nomograph. Plus value of normal stresses means tension stresses and tension stresses are occur near the void in the lower tension reinforcement ratio pt . Allowable bending moment can be calculated as multiplication of minimum value of C1~C3 and bd^2 . C1s illustrated in solid line show the values at which the stress σ_{c1} in fig. 3 reaches the allowable strength and C2s in dashed line show those at which the stress σ_{c2} reaches the allowable strength. And C3 in chain line shows the value at which the stress σ_{st} of reinforcing bar reaches it. As the thickness of slab becomes thicker, stress σ_{c2} at top of circular hole becomes bigger than σ_{c1} at on the surface of slab. And this phenomenon notably appears in higher pt value.

The problem is that tension stresses occur in concrete in some cases. To avoid this fact, we need to use FEM model in which material nonlinearity can be applied, or to use some technics to delete concrete elements that show tension stress.

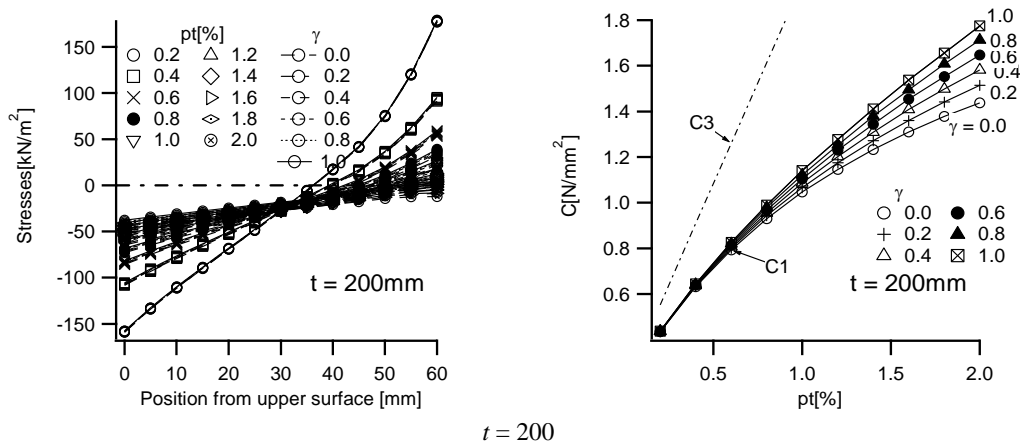


Figure 9. Allowable bending moment for long period load ($F_c = 24[\text{N/mm}^2]$, $n = 9.04$)

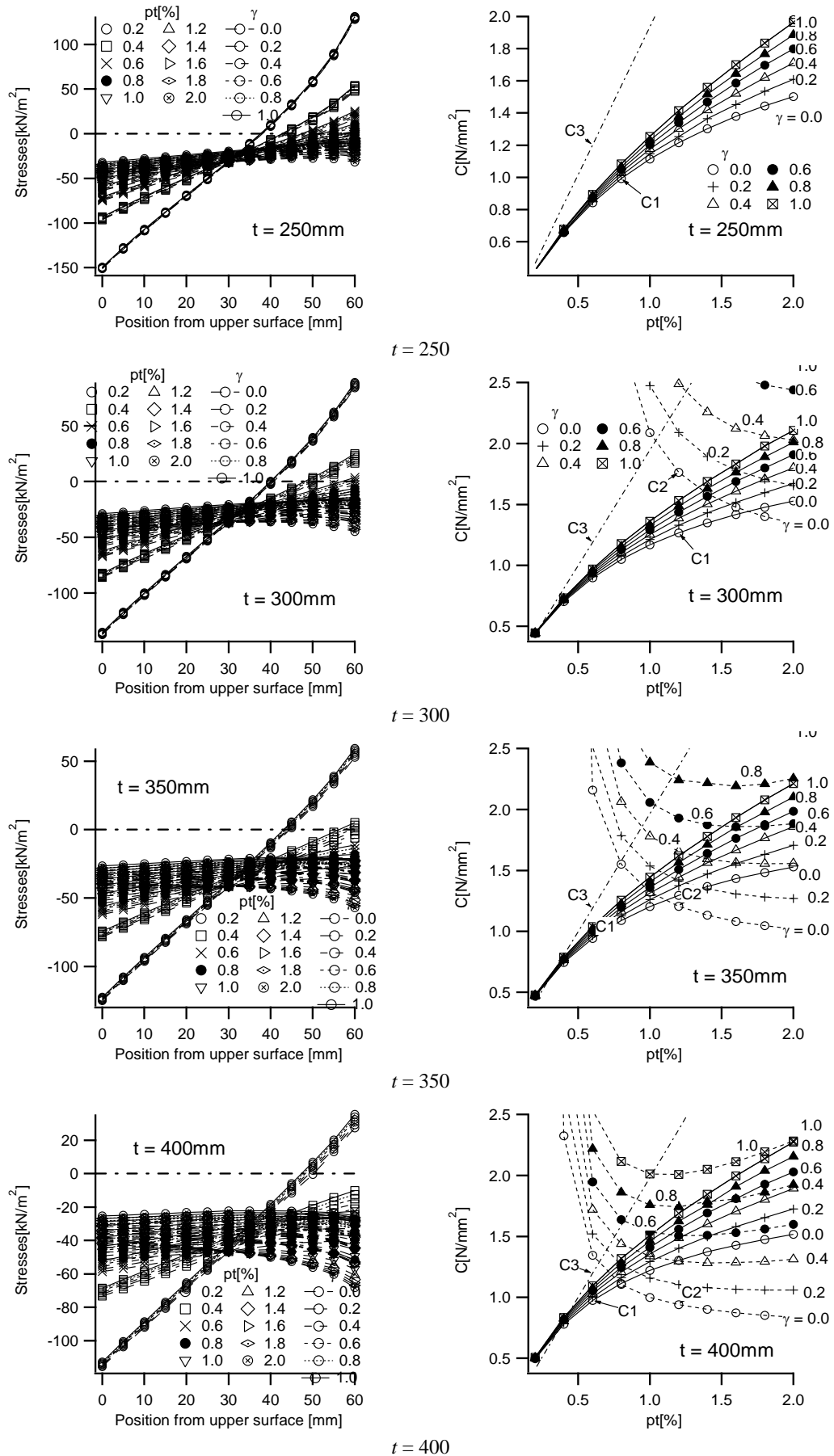


Figure 10. Allowable bending moment for long period load($F_c = 24 \text{ [N/mm}^2\text{]}$, $n = 9.04$)

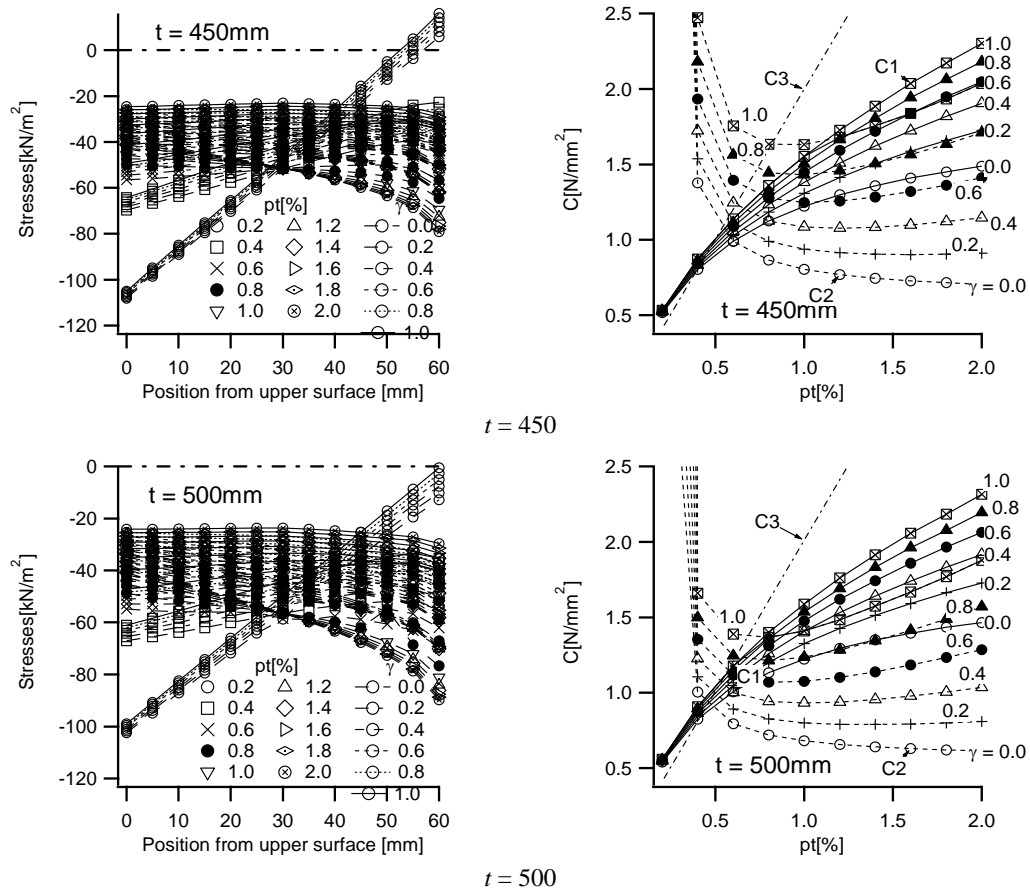


Figure 11. Allowable bending moment for long period load($F_c = 24[\text{N/mm}^2]$, $n = 9.04$)

References

- [1] Gengo Matsui: Design of Void Slab, Shokoku-sha, 1974 (in Japanese)
- [2] Shigehiro MOROOKA, Haruo KUNIEDA and Yukio OGIWARA: ESTIMATION OF BENDING RIGIDITY AND STRESS CONCENTRATION OF CIRCULAR HOLE VOID SLAB, 2002 Volume 67 Issue 558 Pages 133-137, Journal of Structural and Construction Engineering (Transactions of AIJ) (in Japanese)
- [3] Shigehiro MOROOKA: CHARACTERISTICS OF ELASTICITY OF CIRCULAR HOLE VOID SLAB, 2003 Volume 68 Issue 565 Pages 79-85, Journal of Structural and Construction Engineering (Transactions of AIJ) (in Japanese)
- [4] AIJ Standard for Structural Calculation of Reinforced Concrete Structures -Based on Allowable Stress Concept -, revised 1999, Architectural Institute of Japan, 1999 (in Japanese)

Simulation of Smoke Flow in a Longitudinal Ventilated Tunnel in Macau

**Pak Hang Fu^{1,*}, †Hou Kuan Tam¹, Seng Kin Lao¹, Lap Mou Tam^{1,2}, Vai Kuong Sin¹,
Iok Wa Tang¹, Kuan Hou Lam¹**

¹Department of Electromechanical Engineering, Faculty of Science and Technology, University of Macau, Macau SAR, China.

²Institute for the Development and Quality, Macau, Macau SAR, China.

*Presenting author: hang3131@gmail.com

†Corresponding author: hktam@umac.mo

Abstract

In the past decades, tunnel fire safety has been highly concerned since tunnels were widely installed and utilized as transportation links for people and cargos worldwide. In order to achieve a fire safety purpose, smoke extraction system should be well designed and built to extract or control the smoke effectively. The objective of this paper is to investigate the heat and smoke flow behavior based on CFD simulation in a longitudinal ventilated tunnel, Tunnel do Monte da Guia (TMG tunnel) in Macau SAR, China. The dimension and the inclination angle of the existing roadway tunnel are of 285m (L) x 10.5m (W) x 7m (H) and 2.45°, representatively. In this study, a single fire source caused by a vehicle is considered. The smoke flow in the tunnel model is simulated with the CFD program, FDS. In results, the critical velocities are obtained with the tunnel models arranged in the horizontal and inclined positions. Moreover, the smoke temperature is also analyzed in this study.

Keywords: Smoke Flow, Simulation, Longitudinal Ventilation, Tunnel

Introduction

Tunnel ventilation is an important part in a modern tunnel. In the normal traffic condition, the ventilation can keep an acceptable level of contaminants produced by the vehicles. During a fire scenario, the emergency ventilation can remove or control the smoke in order to ensure the life and fire safety in tunnel. Regarding tunnel ventilation, natural ventilation or mechanical ventilation can be selected to use. Natural ventilation is mainly effective for a short and low traffic volume tunnel. For a long and high traffic volume tunnel, mechanical ventilation system is required. Mechanical ventilation is basically classified into three types: full transverse ventilation, semi-transverse ventilation and longitudinal ventilation. Longitudinal ventilation system is the most effective in cost in comparison with the full transverse and semi-transverse systems as the longitudinal system does not require additional space for ducts and requires lower maintenance cost [1]. However, longitudinal ventilation is suitable for a tunnel with unidirectional traffic.

In a longitudinal ventilated tunnel, jet fans are usually directly installed along the longitudinal direction. When the jet fans are activated in the emergency mode, the ventilated air forced the smoke flow from upstream to downstream direction to make a safe route for road passengers to evacuate and fireman to fight the fire. However, if the supply capacity of the ventilation system is not enough or the system is wrongly designed, the smoke will spread to the upstream direction. In such case, people inside tunnel is difficult to escape due to the blockage of evacuation route

and inhalation of excess smoke and toxic gases. The upstream movement of hot smoke and gases is so called the “Back-layering” phenomenon. To prevent the back-layering phenomenon, the minimum ventilation velocity, so-called the critical velocity, provided by the ventilation system is required.

Experiments for the smoke flow in tunnels can be found in the open literature [2-5]. Based on the experimental studies, the critical velocity can be calculated by the empirical equations in terms of the heat release rate and tunnel geometry. Wu and Bakar [2] developed the equations, Eqs. (1-4), which utilized the hydraulic tunnel height \bar{H} to replace the tunnel height as the experimental results showed that the tunnel temperature decreases when the tunnel width increases:

$$V'' = \frac{V}{\sqrt{g\bar{H}}} \quad (1)$$

$$Q'' = \frac{Q}{\rho_0 C_P T_0 \sqrt{g\bar{H}^5}} \quad (2)$$

$$V'' = 0.4[0.20]^{-\frac{1}{3}}[Q'']^{\frac{1}{3}}, \text{ for } Q'' \leq 0.20 \quad (3)$$

$$V'' = 0.4, \text{ for } Q'' > 0.20 \quad (4)$$

where V is the ventilation velocity, ms^{-1} ; Q is the fire convective heat release rate, kW ; V'' is the dimensionless critical velocity based on hydraulic tunnel height; Q'' is dimensionless heat release rate based on hydraulic tunnel height; g is gravitational force, ms^{-2} ; ρ_0 is ambient air density, kgm^{-3} ; C_P is specific heat capacity of air, $\text{kJ kg}^{-1} \text{K}^{-1}$; T_0 is ambient temperature, $^{\circ}\text{C}$;

Lee and Ryou [3] introduced the effect of aspect ratio and tunnel, A_s , into the correlations proposed by Wu and Baker [2]. The new equations are shown as Eqs. (5-7):

$$V'' = \frac{V}{A_s^{0.2} \sqrt{g\bar{H}}} \quad (5)$$

$$Q'' = \frac{Q}{\rho_0 C_P T_0 \sqrt{A_s g \bar{H}^5}} \quad (6)$$

$$V'' = 0.73[Q'']^{\frac{1}{3}}, \text{ for } Q'' \leq 0.20 \quad (7)$$

where A_s is the aspect ratio;

Kennedy and his coworkers [6] proposed the equations, Eqs. (8) and (9) to calculate the critical velocity and average temperature of fire gases, which assumes Froude Number of 4.5. These equations were also quoted in NFPA 502 [7].

$$V_C = K_1 K_g \left(\frac{g \cdot H \cdot Q}{\rho \cdot C_P \cdot A \cdot T_f} \right)^{1/3} \quad (8)$$

$$T_f = \frac{Q}{\rho \cdot C_P \cdot A \cdot V_C} + T \quad (9)$$

where V_C is the critical velocity, m/s ; K_1 is 0.606 which is the Froude number factor ($Fr^{-1/3}$); K_g is the grade factor; g is gravitational force, ms^{-2} ; H is height of tunnel, m ; Q is heat release rate, kW ; ρ is average density of approach (upstream air), kgm^{-3} ; C_P is specific heat capacity of air, $\text{kJ kg}^{-1} \text{K}^{-1}$; A is area perpendicular to the flow, m^2 ; T_f is average temperature of the fire site gases, K ; T is temperature of the approach air, K ;

Nowadays, with the help of the powerful computer and the CFD simulation [2, 8- 9], the smoke flow and heat transfer characteristics in a tunnel fire can be clearly observed and analyzed. Regarding the simulation tool, Fire Dynamic Simulator (FDS) is commonly used for various types of fire engineering problems. Fire Dynamics Simulation (FDS), which is developed by National Institute of Standards and Technology (NIST), has been verified as a reliable tool by a number of tests and experiments in the references [8, 10]. Result of FDS can be displayed in the Smokeview program for analysis of the smoke and heat behavior of tunnel fire.

In this study, FDS is applied to simulate the smoke flow in the tunnel located in Macau. The main objectives of this paper are listed below:

- ✧ Build an accurate tunnel model based on the existing tunnel in Macau - Tunel do Monte da Guia, hereafter called TMG tunnel, with the FDS;
- ✧ Analysis of smoke flow and determine the critical velocity with the tunnel arranged in a horizontal position;
- ✧ Compare the critical velocity obtained from the tunnel models in the horizontal position and in an inclination position;

Description of Tunnel Model

In Macau peninsula, see Figure 1, Tunel do Monte da Guia (TMG tunnel) has been built under the Guia Hill since 1990 and the longitudinal ventilation system has operated for than 28 years. TMG tunnel is built to connect the areas of central residential and outer harbour. Usually, the passenger cars and the medium-sized vans pass through the tunnel. The traffic volume is high daily, especially, during the rush hours. Table 1 lists the geometry of the tunnel.

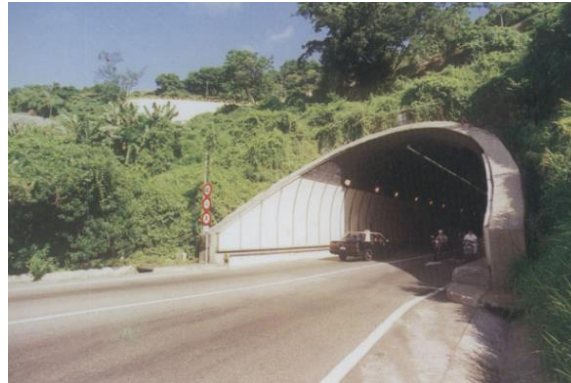


Figure 1. Tunel do Monte da Guia (TMG Tunnel)

Table 1. Parameters of TMG Tunnel

Parameters	Description
No. of lane and direction	Two lanes, bi-directional
Dimension	285m (L) x 10.5m (W) x 7.2m (H)
Height-to-width aspect ratio	0.69
Tilt angle	2.45 °

Figure 2 shows the TMG Tunnel model developed for the FDS simulation. Total length of the TMG Tunnel model is 285 m. In the model, a cubic fire source is arranged at 60 meters after the tunnel inlet and the constant heat release rate (HRR) is set as 14.8MW which is equal to a

van fire [11]. Only the top surface of the cube is set as a ‘fire’ surface. As seen in the figure, seven horizontal thermocouples are set 0.4 m under the ceiling along the downstream direction with a 20 m spacing. To measure the vertical temperature from the ground to the ceiling, eleven vertical thermocouples, with 0.5m spacing vertically, are set at 60 meters downstream from center of fire source (see Figure 3). All the locations of the horizontal and vertical temperature sensors and the fire source of the full scale model is referred to [3].

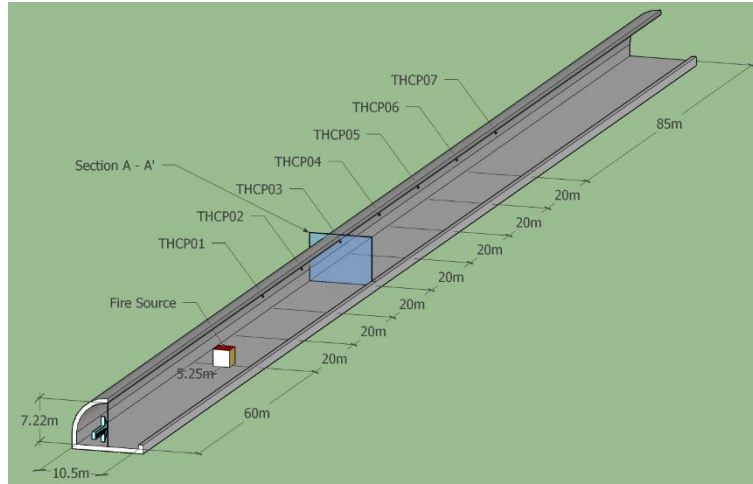


Figure 2. TMG Tunnel Simulation Model

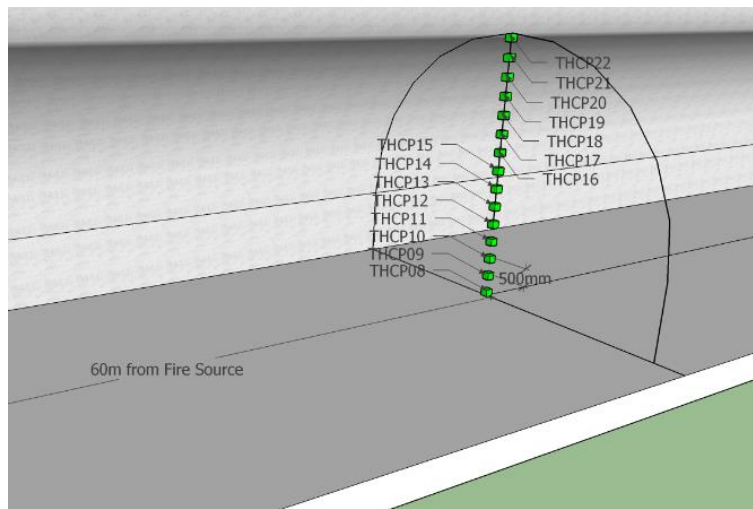


Figure 3. Vertical temperature sensors arranged in the model (Section A-A')

In this study, the electromechanical facilities such as the jet fans are neglected. The ventilation flow is directed obtained from the tunnel inlet. In Table 2, it shows the initial and boundary conditions of the model used in FDS. After a sensitivity analysis of mesh size, the mesh size of 0.4 meters is determined. The total number of meshes of the model is 411,800 and the total simulation time is set as 200 seconds.

Table 2. Initial and Boundary Conditions of FDS Model

Initial Conditions	
Ambient Temperature (T_0)	20 °C
Pressure (P_0)	101.325 kPa (1atm)
Tunnel Wall	Inert wall, 0.1m in thickness

Fire Source	
Heat Release Rate	14.8 MW
Dimension	2.8m (L) x 2.8m (W) x 2.8m (H)
HRRPUA	1888 kW/m ²
Location	Refer to Figure 2
Burning material	Polyurethane
Burning Time	Fire happens at t = 0s; Burning time = Total Simulation time
Boundary Condition	
Inlet	$T_{inlet} = T_0$; P_{inlet} is varied due to ventilation velocity;
Outlet	$T_{outlet} = T_0$; $P_{outlet} = P_0$;
Wall and Floor	$T_W = T_F = T_0$
Ventilation Velocity	$u = u_{critical}$; $v = 0$; $w = 0$;

Results and Discussion

In order to verify the setting of FDS simulation, a reduced scale tunnel model (1040cm (L) x 50cm (W) x 33.3cm (H)) with a height-to-width aspect ratio of 0.67, which is similar to the ratio of 0.69 of the current TMG tunnel simulation is established with the reference of Lee and Ryou [3]. For the reduced scale model, the fire source is reduced from 14.8 MW (for TMG full scale model) to 8.27 kW with the reference to the scaling law [12]. The location of the horizontal and vertical temperature sensors and the fire source of the reduced model is also referred to [3]. After the simulation, as shown in Figure 4, the simulated critical velocity is of 0.59 m/s is almost matched with the experimental critical velocity of 0.58 m/s. Therefore, the same FDS settings for the reduced scale model is applied to the full scale TMG tunnel.

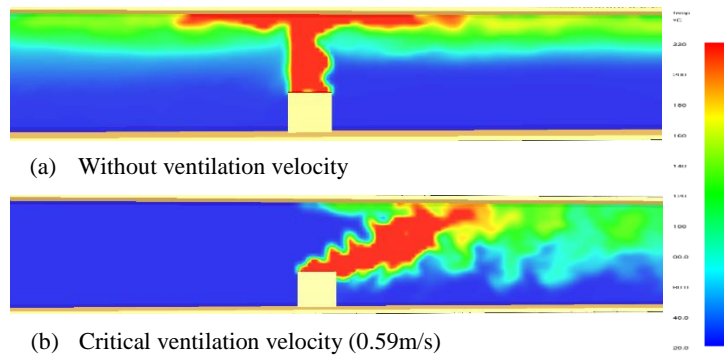


Figure 4. Simulation of the reduced scale tunnel

For the full scale TMG tunnel, a horizontal case is executed first because it is used as a base case to analyze the effect of slope on the smoke flow and critical velocity in the tunnel. Figure 5 shows the temperature contour plot for the TMG tunnel arranged in horizontal position. From the figure (a) – (d), it can be observed that the hot gases or smoke rise from the fire source to the ceiling. After that, the smoke layer becomes thicker and spread to the both ends of the tunnel. To avoid the back-layering phenomena, as shown in the figure (e), the critical ventilation velocity of 3 m/s can force all the gases to the downstream direction and solve the back-layering problem.

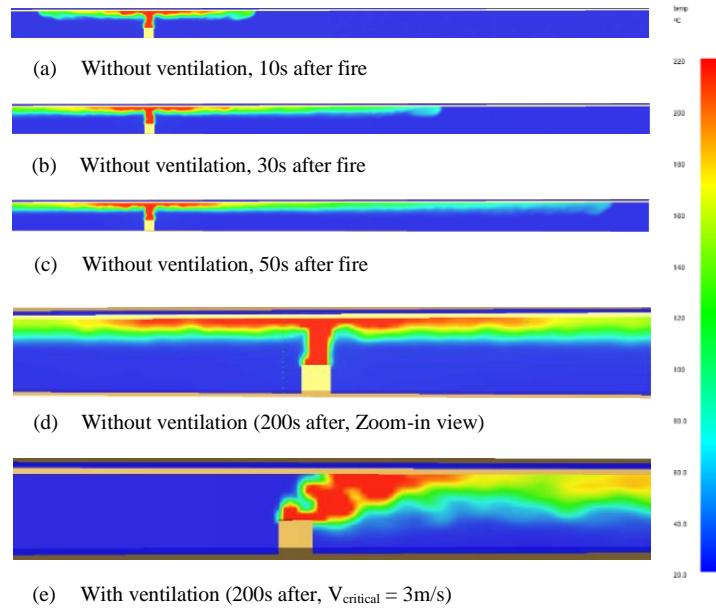


Figure 5. Temperature contour plot for the horizontal TMG Tunnel

Figure 6 shows the temperature contour plot for the TMG tunnel with a upward tilt angle of 2.45° . From the figure (a) – (d), it can be observed that the hot gases or smoke rise from the fire source to the ceiling. After that, the smoke spread to the both ends of the tunnel. However, it seems smoke is faster to spread to the downstream due to a buoyant effect. To avoid the back-layering phenomena, as shown in the figure (e), the critical ventilation velocity of 2.9 m/s can force all the gases to the downstream direction. When compared with a horizontal tunnel, less critical ventilation velocity for inclined tunnel is sufficient to solve the back-layering problem.

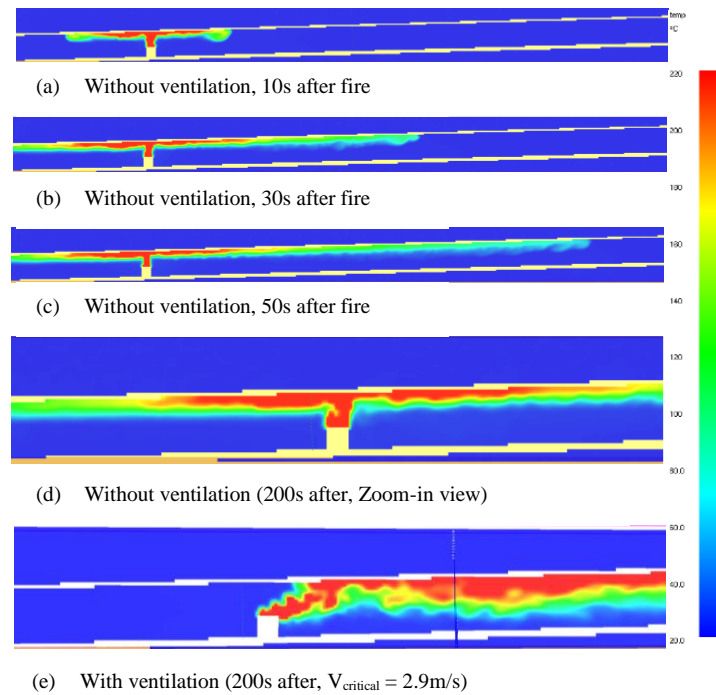


Figure 6. Temperature Contour Plot of the TMG Tunnel (with a tilt angle of 2.45°)

For both of the horizontal and inclined TMG tunnel models, the critical velocities calculated by the FDS simulation are also compared with those predicted by the correlations, Eqs. (1-9) in Table 3. From the table, it is observed that the critical velocities calculated by Wu and Bakar [2] and Lee and Ryou [3] are equal to or larger than 2.5 m/s. Those values are more comparable to the values obtained from the current simulation. As the curved ceiling is used in the current simulation, it is reasonable that there is a discrepancy between the simulation critical velocity values and the values predicted by those correlations. However, the critical velocity is likely under estimated by NFPA 502 [7]. That is only 2 m/s required for the TMG tunnel. As a safety concern, the higher ventilation velocities calculated by the simulation or the correlations [2-3] are recommended for the practical smoke management design.

Table 3: Comparison of Critical Velocity of the Simulation and Calculated by Correlations

Model	FDS Critical Velocity (m/s)	Critical Velocity Calculated by the Correlations (m/s)		
		Wu and Bakar [2]	Lee and Ryou [3]	NFPA 502 [7]
Horizontal TMG	3	2.5 (19% deviation)	2.7 (13% deviation)	2.0 (50% deviation)
Inclined TMG	2.9	2.5 (15% deviation)	2.7 (9% deviation)	2.0 (45% deviation)

Figure compares the temperature measured from the eleven vertical thermocouples placed at 60 meters downstream from center of fire source for the horizontal tunnel and the inclined tunnel without ventilation velocity. From the figure, the hot gas layer is maintained to the height of around 3 meters away from the ground regardless of the horizontal or inclined tunnel is observed. It is also observed for both cases that the temperature trends are basically the same until the height over 6 meters. Higher temperature is observed in the inclined tunnel. Perhaps, more hot gases or smoke moves to the downstream due to the buoyant effect caused by the tunnel inclination to the fire source.

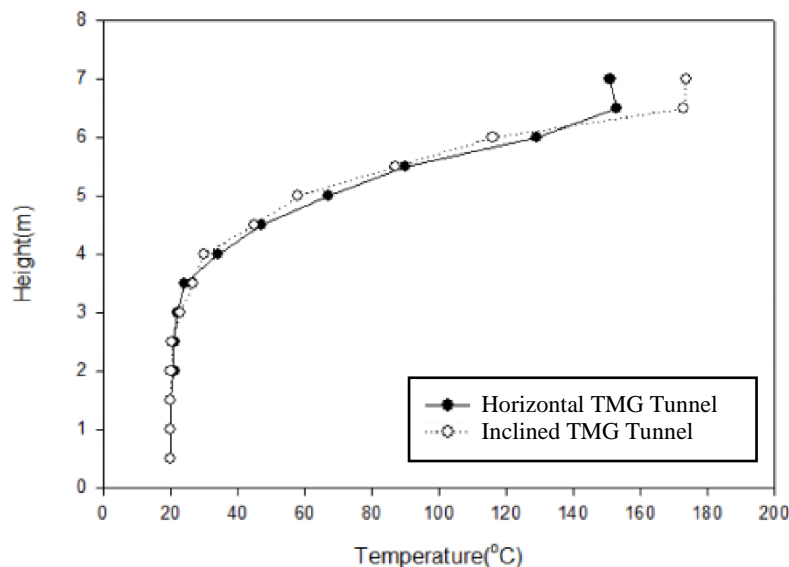


Figure 7. Vertical Temperature profiles of horizontal and inclined TMG Tunnel (without ventilation velocity)

Figure compares the temperature measured from the seven horizontal thermocouples placed under the ceiling for the horizontal tunnel and the inclined tunnel. From the figure, it is observed for both cases that the first thermocouple is much higher because the thermocouple is closer to the fire source. Regardless of the horizontal or sloped tunnel, the temperature curve trend goes downward because the smoke temperature gradually decreases when the smoke moves downstream along the tunnel. For the inclined tunnel, the temperature is higher than that of the horizontal tunnel. It can be explained that more hot smoke is accumulated at the ceiling of the inclined tunnel than the horizontal one.

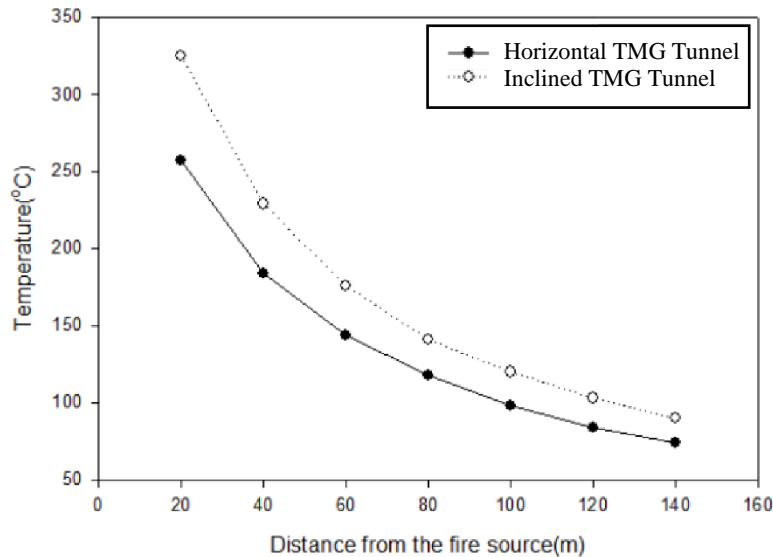


Figure 8. Horizontal Temperature profiles of horizontal and inclined TMG Tunnel Models

Conclusions

In this study, the horizontal and inclined TMG tunnel models simulated by FDS software. From the results, the critical velocity of 2.9 m/s is obtained for the practical TMG tunnel with a tilt angle of 2.45°. That critical velocity is basically matched with the published correlations. From the temperature results, the hot gas layer temperature in the inclined tunnel is higher than that of the horizontal tunnel.

The critical velocity obtained from this study can be used as a reference to verify whether the existing ventilation system has sufficient velocity to ensure the life and fire safety in the practical TMG tunnel in Macau.

Acknowledgement

This research is supported by the Institute for the Development and Quality, Macau.

References

- [1] Kim, J.H., Kim J. K., Yoon, J. Y., Choi, Y. S. and Yang, S. H.. (2015). "Application of multi-objective optimization techniques to improve the aerodynamic performance of a tunnel ventilation jet fan" , *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 229: 91-105.
- [2] Wu, Y., and Bakar, M.Z.A. (2000). "Control of smoke flow in tunnel fires using longitudinal ventilation systems—a study of the critical velocity", *Fire Safety Journal*, 35: 363-90.
- [3] Lee, S.R., and Ryou, H. S.. (2005). "An experimental study of the effect of the aspect ratio on the critical velocity in longitudinal ventilation tunnel fires", *Journal of Fire Sciences*, 23: 119-38.
- [4] Oka, Y. and Atkinson, G. T. (1995) "Control of Smoke Flow in Tunnel Fires", *Fire Safety Journal*, 25 : 305-322.
- [5] Tso, C.Y. and Chow, W. K. (2011), "Obersvation of smoke movement pattern in a scale tunnel model with longitudinal ventilation," *International Journal on Engineering Performance-based Fire Codes*, 10(2) : 36-39.
- [6] Kennedy, W. D, Gonzalez, J. A., Sanchez, J. G. (1996). "Derivation and application of the SES critical velocity equations", *ASHRAE Transactions-American Society of Heating Refrigerating Airconditioning Engin*, 102: 40-44.
- [7] NFPA 502, *Standard for road tunnels, bridges, and other limited access highways*, 2011 Edition.
- [8] Lee, S.R., and Ryou, H. S.. (2006). "A numerical study on smoke movement in longitudinal ventilation tunnel fires for different aspect ratio", *Building and Environment*, 41: 719–725.
- [9] Jojo, S.M. Li, and Chow, W. K. (2003), "Numerical studies on performance evaluation of tunnel ventilation safety systems," *Tunnelling and Underground Space Technology*, 18 : 435-452.
- [10] Guo, X.P., and Zhang, Q. (2014). "Analytical solution, experimental data and CFD simulation for longitudinal tunnel fire ventilation", *Tunnelling and Underground Space Technology*, 42: 307-13.
- [11] Cheong, MK, Spearpoint, MJ, and Fleischmann, CM. (2008) "Design fires for vehicles in road tunnels." *Proc7th International Conference on Performance-Based Codes and Fire Safety Design Methods*, Auckland, New Zealand, pp.229-240.
- [12] Quintiere, J.G. (1989). Scaling Applications in Fire Research, *Fire Safety Journal*, 15: 3–29.

A finite element based procedure for accurate determination of mode I SIF of orthotropic materials based on two parameter strain series

†*Debabrata Chakraborty¹, Debaleena Chakraborty², and K. S R. K. Murthy³

^{1,3}Professor, Department of Mechanical Engineering, I.I.T. Guwahati, Assam, India

²Ex-Research Scholar, Department of Mechanical Engineering, I.I.T. Guwahati, Assam, India

*Presenting author: chakra@iitg.ernet.in

†Corresponding author: chakra@iitg.ernet.in

Abstract

In strain gage based determination of stress intensity factor (SIF), the degree of accuracy of strain readings is affected by the radial position of the gage ahead of the crack tip. In this paper, a theoretical frame work based on two parameter strain series has been developed for accurate determination of mode I SIF (K_I). Based on the two parameter strain series a finite element analysis (FEA) based methodology is presented by which the limit of the radial location (r_{\max}) up to which the two parameter strain series is valid could be determined. Therefore knowing the r_{\max} for a specimen and placing the strain gages within r_{\max} will ensure accurate determination of K_I . Even though two parameter strain series for determination of K_I in orthotropic laminates was used by earlier researchers, radial location of the strain gage was decided by trial and error and hence accuracy could not be ensured. Using the proposed methodology, optimal radial location for the strain gage has been determined for $[0_2/90]_{2s}$ glass-epoxy edge-cracked laminates and numerical simulations were performed for determination of K_I . Results from numerical simulations show that the present technique using two parameter strain series not only allows the strain gage to be placed at a radial distance sufficiently away from the crack tip compared to that reported by earlier researchers, but also ensures accurate determination of K_I in orthotropic composites.

Keywords: Stress intensity factor, Orthotropic composites, Strain gage, Finite element

Introduction

The use of strain gages for the determination of SIF was first proposed by Irwin many years back [1]. However, complications related to a strain gage based application like high strain gradients, 3D effects etc. restricted the usage of strain gages to its full potential. Dally and Sanford [2] then proposed a single strain gage technique for the determination of K_I in isotropic materials. They employed a truncated three parameter strain series representation for the strain field around the crack tip and after careful mathematical simplification they devised that a single strain gage placed sufficiently away from the crack tip with certain orientations and location (decided by θ and ϕ) (Fig. 1) can be used to measure K_I for isotropic materials. In the case of orthotropic composites, the development of similar technique was little more involved due direction dependent properties. Shukla and co-workers [3] were the first to make a similar attempt for a single strain gage technique for the

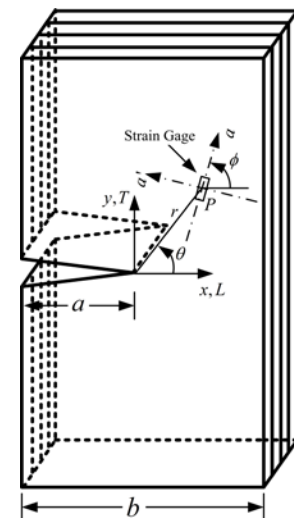


Figure 1. Strain gage location ahead of the crack tip

determination of K_I of orthotropic materials. However, instead of a three parameter series as originally proposed by Dally and Sanford [2], they used a two parameter series representation of the strain field around the crack tip to make the analysis simpler. They selected the radial distance for mounting the strain gage as 5 mm and 9 mm based on previous experience and compared the results and did not present any further explanation for the selection of gage locations. It is important that for accurate determination of SIF, the gage needs to be placed at an optimum location which is neither very near to (where strain gradients and 3D effects are prevalent) nor far away from the crack tip (where other terms in addition to the selected two terms of the infinite series become significant). The lower limit of the gage location was established experimentally [3-4]. However, the upper limit of the strain gage though so important was purely based on intuition and was not addressed. Only recently Sarangi et al. [5-6] came up with an approach to determine the upper limit of the radial distance (r_{\max}) of placing the strain gage in the Dally and Sanford technique for determination of K_I for isotropic materials. Even though, recently, Chakraborty et al [7-8] extended Dally and Sanford technique to orthotropic composites, no such attempts have been made to determine r_{\max} corresponding to the relatively simpler single strain gage method proposed by Shukla et al. [3] for determination of K_I for orthotropic materials. The present work attempts at establishing a theoretical basis for developing a procedure for determination of r_{\max} using FEA corresponding to the two parameter method. Numerical simulations are also presented supporting the present theoretical formulation.

Theoretical Development and Numerical Simulation

The two parameter strain field representation of the normal strain component ε_{aa} at an angle ϕ with the crack axis (CCW direction with crack axis is positive) at a point P located by r and θ (Fig. 1) obtained using strain transformation laws as [3]

$$\varepsilon_{aa} = A_0 \left\{ \left[\frac{1}{\sqrt{r_1}} \left(\cos \frac{\theta_1}{2} \frac{\alpha - \beta}{2\alpha} \left(\cos^2 \phi (-a_{11}(\alpha + \beta)^2 + a_{12}) \right) + \sin^2 \phi (-a_{12}(\alpha + \beta)^2 + a_{22}) \right) - \left(\sin \frac{\theta_1}{2} a_{66} \sin \phi \cos \phi \left(\frac{\alpha^2 - \beta^2}{2\alpha} \right) \right) \right] + \left[\frac{1}{\sqrt{r_2}} \left(\cos \frac{\theta_2}{2} \frac{\alpha + \beta}{2\alpha} \left(\cos^2 \phi (-a_{11}(\alpha - \beta)^2 + a_{12}) \right) + \sin^2 \phi (-a_{12}(\alpha - \beta)^2 + a_{22}) \right) + \left(\sin \frac{\theta_2}{2} a_{66} \sin \phi \cos \phi \left(\frac{\alpha^2 - \beta^2}{2\alpha} \right) \right) \right] \right\} + B_0 \left\{ \frac{\beta}{2\alpha} [(\alpha + \beta)^2 - (\beta - \alpha)^2] (a_{11} \cos^2 \phi + a_{12} \sin^2 \phi) \right\} \quad (1)$$

Inspection of Eq. (1) suggests that the coefficient of B_0 term can be eliminated by selecting an angle ϕ such that

$$\tan^2 \phi = -a_{11}/a_{12} = 1/\nu_{12} \quad (2)$$

Thus, with ϕ determined from Eq. (2), the strain ε_{aa} can be written in terms of r , θ and unknown term A_0 as

$$\varepsilon_{aa} = \frac{1}{\sqrt{r}} \times A_0 \left\{ \begin{aligned} & \left[\frac{1}{E_2} \left(\frac{1 - \nu_{LT}\nu_{TL}}{1 + \nu_{LT}} \right) \frac{1}{2\alpha} \left[\frac{\cos\left(\frac{1}{2}\tan^{-1}\left((\beta + \alpha)\tan\theta\right)\right)}{\sqrt[4]{(\cos^2\theta + (\beta + \alpha)\sin^2\theta)}} (\alpha - \beta) + \frac{\cos\left(\frac{1}{2}\tan^{-1}\left((\beta - \alpha)\tan\theta\right)\right)}{\sqrt[4]{(\cos^2\theta + (\beta - \alpha)\sin^2\theta)}} (\alpha + \beta) \right] \right. \\ & \left. + \frac{1}{G_{12}} \left[\frac{\nu_{LT}}{(1 + \nu_{TL})\sqrt{\nu_{TL}}} \right] \frac{1}{2\alpha} \left[\frac{\sin\left(\frac{1}{2}\tan^{-1}\left((\beta + \alpha)\tan\theta\right)\right)}{\sqrt[4]{(\cos^2\theta + (\beta + \alpha)\sin^2\theta)}} - \frac{\sin\left(\frac{1}{2}\tan^{-1}\left((\beta - \alpha)\tan\theta\right)\right)}{\sqrt[4]{(\cos^2\theta + (\beta - \alpha)\sin^2\theta)}} \right] \right] \end{aligned} \right\} \quad (3)$$

Therefore, by placing a single strain gage as shown in Fig. 1 at a radial distance r from the crack tip along the line at angle θ and oriented at an angle ϕ , the measured strain ε_{aa} can be equated to Eq. (3) to obtain the value of unknown coefficient A_0 . The mode I SIF can then be determined using $K_I = \sqrt{2\pi}A_0$. For a given cracked configuration, applied load and material properties and after a careful selection of the angle θ , Eq. (3) can be written as

$$\varepsilon_{aa} = \frac{C}{\sqrt{r}} \quad (4)$$

where C is a constant. Taking logarithm on both sides of Eq. (4) results as

$$\ln(\varepsilon_{aa}) = -0.5\ln(r) + \ln(C) \quad (5)$$

A plot of Eq. (5) on log-log axes depicts a straight line of slope equals to -0.5 , with an intercept of $\ln(C)$. If r_{\max} is the extent of valid two parameter zone theoretically, the straight line property will break beyond $r > r_{\max}$ as more than two parameters would be needed in Eq. (3) to estimate the ε_{aa} . Thus, Eq. (5) is valid along the gage line determined by the angle, θ for $r \leq r_{\max}$.

Numerical Simulation, Results and Discussions

In the present study, a $[0_2/90]_{2s}$ glass-epoxy edge-cracked configuration with $a/b = 0.4$ as used by Shukla et al. [3] for their experimentations has been considered. Table 1 lists the material properties, loading and geometric parameters of the cracked orthotropic panel considered in this section. The parameters $\alpha = 0.9684$, $\beta = 1.4496$ and orientation $\phi = 68.01^\circ$ are determined from the material properties. The exact model used by Shukla and co-workers [3] has been duplicated so that the results can be easily compared and the value of θ as suggested after taking care of various factors like strain gradients, averaging error due to finite gage size and the like is found to be $\theta = 38^\circ$.

Table 1. Geometry and material parameters for edge-cracked $[0_2/90]_{2s}$ glass-epoxy

b (mm)	h/b	ν_{LT}	E_L (GPa)	E_T (GPa)	G_{LT} (GPa)	σ (MPa)
50	3	0.163	33.3	24.6	5.2	100

Finite element analysis for the present example has been carried out in ANSYS employing eight noded isoparametric quadrilateral elements. A typical finite element mesh used in the present analysis considered after proper convergence study is shown in Fig. 2(a). Following Eq. (5), the plots of $\ln(\varepsilon_{aa})$ versus $\ln(r)$ obtained from the FEA is shown in Fig. 2(b). It may be observed from Fig. 2(b) that plot consists of well demarcated zones defining the linear and non-linear portions as predicted by the theory. The extent of the linear portion which gives the value of r_{\max} for that configuration is found to be 16.74 mm. The numerically determined SIF values computed at different locations are compared with the analytical results. The analytical expression for mode I SIF of this configuration is given by [3]

$$K_I = Y_I (a/b) \sigma \sqrt{a} \quad (6)$$

where σ is the applied stress, a is the crack length and Y_I is the specimen geometric factor given by [3]

$$Y_I = 1.99 - 0.41(a/b) + 18.7(a/b)^2 - 38.48(a/b)^3 + 53.85(a/b)^4 \quad (7)$$

The computed strains at all the nodes on the gage line are considered as the measured strains using a single strain gage oriented at angle of $\phi = 68.01^\circ$ with the crack axis at the corresponding radial distances. Following the procedure explained, the r_{\max} value of this configuration is found to be 16.74 mm and the thickness of the plate is set to 1 mm. Therefore, according to the present approach, any radial distance of the strain gage from the crack tip that satisfies

$$1 \text{ mm} \leq r \leq 16.74 \text{ mm} \quad (8)$$

is an optimal or valid gage location for accurate determination of mode I SIF for the problem considered. The gage locations for which $r \geq r_{\max}$ are invalid or non-optimal locations. Accordingly, the strain ε_{aa} is sampled at two optimal gage locations (for which, $r < 16.74 \text{ mm}$) and two non-optimal gage locations (for which, $r \geq 16.74 \text{ mm}$) as shown in Table 2. For this configuration at $\sigma = 100 \text{ MPa}$ the reference value of mode I SIF determined using Eq. (7) is $K_I = 52.80 \text{ MPa}\sqrt{\text{m}}$. The measured mode I SIF using a single strain gage located at those optimal and non-optimal radii are determined using the simulated finite element strain values ε_{aa} at those radii (Table 2). The percent relative error in measured is computed as

$$\% \text{ Rel. error} = \frac{K_{\text{Analytical}} - K_{\text{measured or simulated}}}{K_{\text{Analytical}}} \times 100 \quad (9)$$

Table 2 shows the comparison of numerically simulated SIFs with the analytical value and the corresponding percent error at different radial locations. It may be observed from the results of numerical simulations that for the strain readings at radial locations within r_{\max} , the values of SIF are highly accurate and those estimated at non-optimal radial locations (beyond r_{\max}) are erroneous. This also shows that the radial location ($r = 9 \text{ mm}$) used by Shukla et al. [3] was very well within the estimated r_{\max} of 16.74 mm and hence resulted in accurate K_I . However, using the proposed method, it was observed that highly accurate value of K_I could

actually be obtained even by placing the strain gage well beyond 9 mm thereby substantiating the importance and usefulness of the method in fracture mechanics in accurate determination of SIFs.

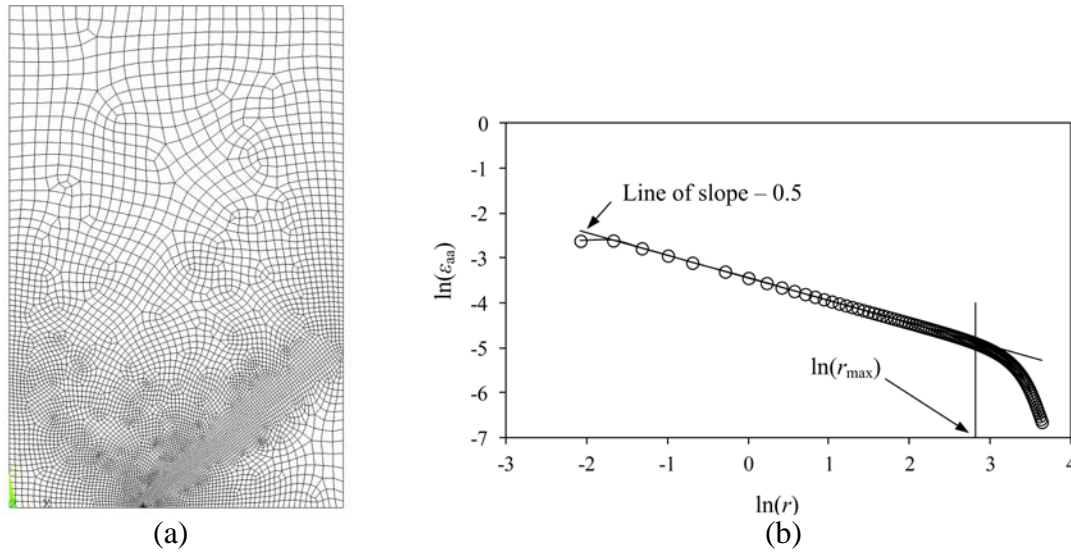


Figure 2. (a) FE mesh (b) Variation of $\ln(\epsilon_{aa})$ vs. $\ln(r)$

Table 2. Simulated K_I at the optimal and non-optimal locations for edge cracked $[0_2/90]_{2s}$ glass-epoxy laminate

r (mm)	ϵ_{aa}	Analytical K_I ($\text{MPa}\sqrt{\text{m}}$)	Measured K_I ($\text{MPa}\sqrt{\text{m}}$)	% Relative Error
10.98	9.57E-03	52.80	51.96	1.59
14.29	8.26E-03		51.16	3.11
18.56	6.92E-03		48.85	8.1
22.01	5.92E-03		45.5	13.83

Conclusions

In the present work, a theoretical frame work has been developed for determination of maximum radial location (r_{max}) up to which the two parameter strain series is valid ahead of the crack tip in an orthotropic laminate. Based on the theoretical formulation, a finite element based methodology has been proposed by which the r_{max} for a particular configuration of edge cracked orthotropic composites could be determined. Numerical simulations have been performed and the results show that it is possible to accurately determine mode I SIF in orthotropic composites using a single strain gage considering two parameter strain series only when the gage is placed within r_{max} . In the absence of prior knowledge of r_{max} , and placing gages arbitrarily might lead to erroneous values of SIF. Therefore the development of this FE based procedure will be immensely useful for experimentalists using the procedure put forward by Shukla et al [3] for determination of K_I in edge cracked orthotropic composites.

References

- [1] Irwin G. R. (1957) Analysis of stresses and strains near the end of a crack traversing a plate, *Journal of Applied Mechanics*, **24**, 361–364.
- [2] Dally, J. W. and Sanford, R. J. (1987) Strain gage methods for measuring the opening mode stress intensity factor, *Experimental Mechanics*, **27**, 381–388.
- [3] Shukla, A., Agarwal, B. D. and Bhusan, B. (1989) Determination of stress intensity factor in orthotropic composite materials using strain gages, *Engineering Fracture Mechanics*, **32**, 469 – 477.
- [4] Rosakis, J and Ravi-Chandar K (1986) On crack-tip stress state: An experimental evaluation of three dimensional effects, *International Journal of Solids and Structures*, **22**:121–134.
- [5] Sarangi H, Murthy KSRK, Chakraborty D. (2010) Radial locations of strain gages for accurate measurement of mode I stress intensity factor, *Materials & Design*, **31**, 2840–2850.
- [6] Sarangi H, Murthy KSRK, Chakraborty D. (2010) Optimum strain gage location for evaluating stress intensity factors in single and double ended cracked configurations, *Engineering Fracture Mechanics*, **77**, 3190–3203.
- [7] Chakraborty, D, Murthy, K. S. R. K. and Chakraborty, D. (2014) A new single strain gage technique for the accurate determination of mode I stress intensity factor in orthotropic composite materials, *Engineering Fracture Mechanics*, **124–125**, 142-154.
- [8] Chakraborty, D, Murthy, K. S. R. K. and Chakraborty, D. (2016) Determination of K_I in orthotropic laminates with double ended cracks using a single strain gage technique, *Theoretical and Applied Fracture Mechanics*, **82**, 96-106.

KBL: A knowledge based learning method for extracting formulas of aerodynamic heating*

†**Changtong Luo¹, Zonglin Jiang^{1,2}**

¹ State Key Laboratory of High Temperature Gas Dynamics, Institute of Mechanics,
Chinese Academy of Sciences, China.

² School of Engineering Sciences, University of Chinese Academy of Sciences, China

†Corresponding and presenting author: luo@imech.ac.cn

Abstract

Aerodynamic heating of hypersonic flights originates from the action of strong shock and viscous boundary layer. Accurate prediction of the heating flux is a classic problem in aerodynamics. Great efforts including theoretical analysis, numerical simulation and wind-tunnel/flight test, have been devoted to it since the middle of last century, and many empirical formulas for some typical shapes have been obtained, which have supported the development of current aerospace engineering. However, formulas of quantitative relationships with more universal meaning are still lacking, and conventional analytical method is difficult to solve this problem. In this work, an intelligent optimization method, knowledge based learning (KBL), is proposed to detect the underlying laws of aerodynamic heating that could unify the data distribution of different models from different wind tunnels. In KBL, the wind-tunnel data to be learned are preconditioned and enriched with expert knowledge and numerical simulation of high fidelity, and then a special genetic programming algorithm is developed to detect the underlying universal laws of aerodynamic heating. The proposed method has been applied to two groups of aerodynamic heating data collected from different wind tunnels. One is of sphere heat transfer, and the other is of a double ellipsoid configuration. The study indicates that the KBL can discovered concise formulas and is promising for detecting more universal laws of aerodynamic heating.

Keywords: Artificial intelligence, Data correlation, Aerodynamic heating, Wind tunnel, Computational fluid dynamics

Introduction

Aerodynamic heating of hypersonic flights originates from the action of strong shock and viscous boundary layer. It is a critical issue to consider when developing a new hypersonic vehicle [1]. Accurate prediction of the heating flux is a classic problem in aerodynamics. Great efforts including theoretical analysis, numerical simulation and wind-tunnel/flight test [5,6,9-11,18-22], have been devoted to it since the middle of last century, and many empirical formulas for some typical shapes have been obtained.

Early research was mainly based on boundary layer analysis and experimental data fitting. Empirical formulas are available for many typical configurations. The classical Fay-Riddle formula can be used in the stagnation area. The surface friction formula of Blasius and the modified Reynolds analogy can be used in the large-area of laminar flow, and the reference enthalpy method can be used to consider the effect of high-speed laminar compressibility.

* This work has been supported by the National Natural Science Foundation of China (Grant No. 11532014).

For the turbulent area, the surface friction formula of Schultz-Grunow, reference enthalpy methods and Reynolds analogy, or the surface friction formula of Spalding-chi and Reynolds analogy could be applied. To distinguish the laminar and turbulent flow, Van Driest criterion, Batt criterion, Dirling criterion, and Bishop criterion could be used to estimate the position of transition.

With the development of computer and CFD, numerical simulation plays an increasingly important role in the prediction of aerodynamic heating. However, direct CFD method is Computationally intensive. Therefore, researchers began to combine theoretical empirical formulas with simplified CFD (e.g., inviscid CFD) simulations to sketch surface streamlines and predict the aerodynamic heating. The hybrid method needs less calculation time, and does work for more complex shapes. For example, the NASA Langley Research Center developed AEROHEAT into the AA3DBL program based on the theory of three-dimensional axisymmetric boundary layer. This program can not only predict the heat flow distribution of the centerline, but also predict the lateral heat flow distribution off the centerline, which can be used for calculations such as space shuttles, etc. It can be used to calculate the heat flux distribution of three-dimensional aircrafts like the Space Shuttle. Later, a generalized body-fitted coordinate system was used to allow boundary layer calculations to be used in conjunction with inviscid flow field solution methods. Thus, it can be used to solve the heat flux distribution of any shape of aircrafts.

It is worth noting that people's pursuit of universal laws of aerodynamic heating will never stop even if the CFD and computational capabilities become so advanced that heat flux distribution can be accurately predicted. In fact, CFD is a gray box process. The inputs are flow parameters (i.e., P , T , V , etc.), mesh grids, turbulence models, boundary conditions, etc. The output is the heat flux at each point of the grid-nodes or grid-cells, and it cannot directly depict the intrinsic quantification relationship of heat flux and flow parameters (law of aerodynamic heating).

Comparing classical mechanics with aerodynamic heating, universal laws that can quantitatively describe the relationship of different parameters such as Newton's second law ($F=ma$) and the law of universal gravitation ($F=GMm/r^2$) are still lacking.

Of course, it is not easy to obtain such universal laws. For example, Kepler's three laws of planetary motion (the Law of Ellipses, the Law of Equal Areas, the Law of Harmonies) have cost Kepler more than 8 years of hard work, based on the set of astronomical observation data collected by Danish astronomer Tycho Brahe. The discovery of Kepler's three laws has undergone two phases, knowledge accumulation (relevant theories, observation/experiment data) and formula extracting.

Comparing the present and the past, the phase of knowledge accumulation has matured. In fact, many countries have been made great efforts to develop new types of hypersonic vehicles, and have conducted a large number of wind tunnel experiments and flight tests, and accumulated a large number of aerothermal data with high-precision [16]. So it is time for the stage of formula extracting.

Note that extracting universal formulas of aerodynamic heating would be much more difficult. In fact, the aerodynamic heat flux is a gradient quantity influenced by global parameters, and it has is strong nonlinearity. Compared to Kepler's law, the formula that can depict the intrinsic quantification relationship of heat flux and flow parameters must have higher

complexity. For example: the heat flux formula of Fay-Riddell [6], at the stagnation point is relatively complex:

$$q = 0.763(Pr)^{-0.6} \left(\frac{\rho_w \mu_w}{\rho_s \mu_s} \right)^{0.1} \sqrt{\rho_s \mu_s \left(\frac{du_e}{dx} \right)_s} \left[1 + (Le^\alpha - 1) \frac{h_D}{h_s} \right] (h_s - h_w)$$

where $\left(\frac{du_e}{dx} \right)_s$ is the velocity gradient at the edge of boundary layer, and equals to

$$\frac{1}{2R_0} \sqrt{\frac{2(P_s - P_\infty)}{\rho_s}}, \quad \text{if } Ma_\infty > 1.12$$

$$\frac{V_\infty}{R_0} (2.0 - 0.872Ma_\infty^2 - 0.328Ma_\infty^4), \quad \text{if } Ma_\infty \leq 1.12$$

As above shown, conventional analytical method of manual deduction is difficult to extracting laws of aerodynamic heating. Fortunately, artificial intelligence (AI), including genetic programming [8], big data algorithm, deep learning, has made considerable progress in recent years, and has been widely concerned and applied [2,3,4,7,13,14,19]. In near future, AI as an extension of human brain will be an inevitable trend. However, general AI are not suitable for learning the wind tunnel / flight test data. In fact, compared with the big data obtained through the Internet (large amount, timeliness and weak correlation), the data of wind tunnel / flight test can be regarded as a small data with strong correlation characteristics [17]. On the one hand, it is difficult to carry out a large number of experiments to obtain massive data for each model. On the other hand, the effects of the parameters on the heat flow are almost deterministic. Furthermore, even if a general AI method (e.g., deep learning) could give an accurate model, the neural network is still a gray box, which involves a large number of coefficients. It cannot give a concise, intuitive, easy to understand law of aerodynamic heating.

In this work, a special AI method, knowledge based learning (KBL), is proposed. In KBL, the wind-tunnel data to be learned are preconditioned and enriched with expert knowledge and numerical simulation of high fidelity, and then a special genetic programming algorithm is developed to detect the underlying universal laws of aerodynamic heating. The proposed method has been applied to two sets of aerodynamic heating data collected from different wind tunnels. One is of sphere heat transfer, and the other is of a double ellipsoid configuration. Study shows that the KBL can discovered concise formulas and is promising for detecting more universal laws of aerodynamic heating.

Knowledge based learning

In order to detect universal laws of aerodynamic heating using the wind tunnel/flight test data, both aspects, regarding to data and algorithm, respectively, must be customized. On one hand, the wind tunnel/flight test data are parameterized, localized, and standardized by CFD, and enriched and transformed using expert knowledge. On the other hand, general AI algorithm is specialized to enhance the ability of nonlinear function evolution, dimension analysis, and interval analysis. Then, the specialized AI algorithm will be used to detect universal laws of aerodynamic heating using the wind tunnel/flight test data as its training, test, and validation sets. The sketch could be illustrated as Figure 1.

Data preconditioning

Usually, the information of the heat flux data from wind tunnel tests at different measurement points is very brief. The test condition parameters involve only the total temperature, the total pressure and the Mach number. In this way, not only the heat flux data of different models cannot be compared, but also the data of the same model under different wind tunnels cannot

be directly compared and correlated. For example, it is usual that although the Mach number of test flows in two wind tunnels is the same, the unit Reynolds number may be quite different. In this work, CFD simulation is applied to identify the local parameter near the measurement points (i.e., the parameters on the edge of boundary layer). And then the raw data is extended with the knowledge of high temperature gas dynamics. For example, the original data (q_w , P_0 , T_0 , M_∞ , α , ...) could be extended to (q_w , C_q , St , St_e ; P , T , v , C_p , Entropy, γ , h , k , μ , ν , ρ , Schmidt, c , thermalDiff, M , Re , Re_x).

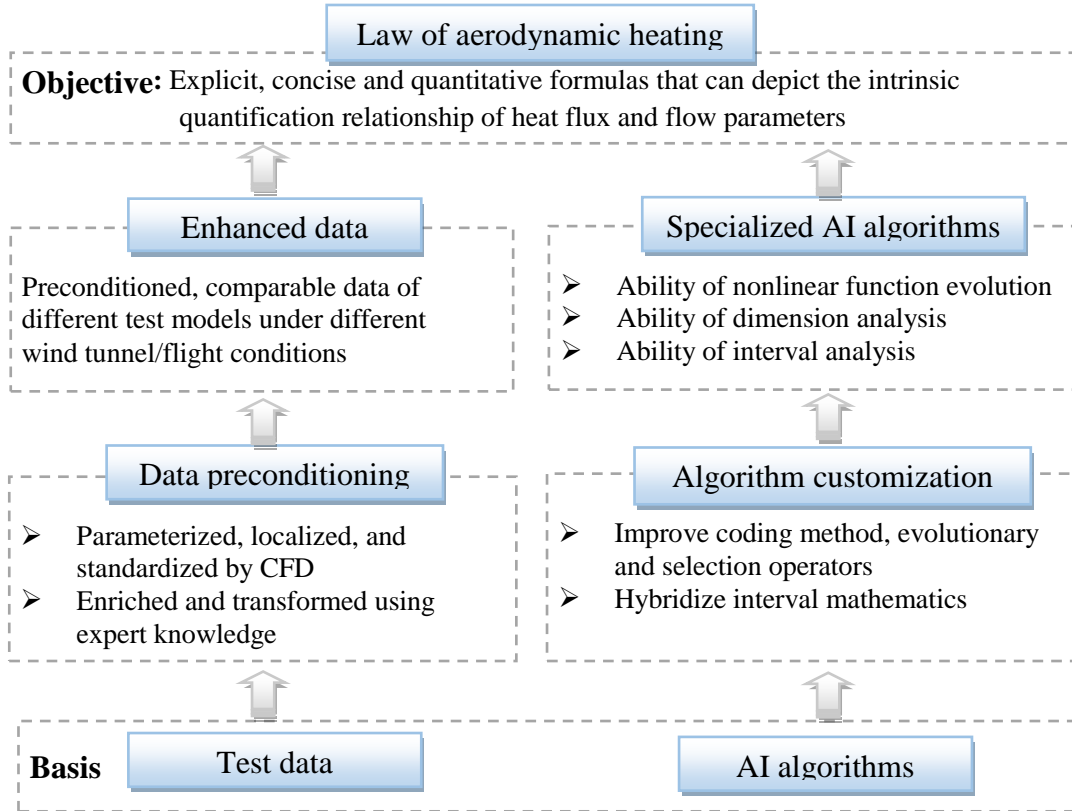


Figure 1. Sketch of knowledge based learning (KBL)

Note that the aim of CFD simulation here is to obtain the macroscopic structure of flow field and the local parameters of boundary layer outer edge, where the CFD model could be laminar, or turbulent, the real gas effect might be considered or ignored, according to different cases. Small errors are supposed to be “commonplace”, since the AI algorithm to be applied is expect to be capable of handling noise data within some tolerance error. The purpose is different from both direct CFD simulation of aerodynamic heating and the CFD simulation in engineering methods. In fact, the CFD simulation in engineering methods should be laminar to save the computational cost, and the direct CFD simulation of aerodynamic heating is not only computationally intensive but also has high requirements for Y^+ and orthogonality on mesh grid. Determining the macroscopic structure of flow field and the local parameters of boundary layer outer edge needs much less CPU time and requires much less on mesh grid. In fact, during the CFD iteration process, after the macroscopic structure of flow field become steady, the flow field inside the boundary layer still needs a lot of computation time to get the right heat flux.

Algorithm customization

According to the data characteristics of aerodynamic heating, it is necessary to improve the existing artificial intelligence (AI) algorithms in order to carry out intelligent learning. In this

work, general AI algorithm is specialized to enhance the ability of nonlinear function evolution, dimension analysis, and interval analysis. In fact, the aerodynamic heat flux is a gradient quality and has strong nonlinearity, which requires the AI algorithm used to have a good nonlinear function evolution ability. Meanwhile, the aerodynamic heating data are strongly correlated. The expression of the quantitative relation that reflects the aerodynamic heating law should be dimensional compatible, so the algorithm needs a certain ability of dimensional analysis. Finally, the interval analysis ability of the algorithm is also obligatory. This can not only help the stability of the algorithm, but also give the confidence interval of the result. The intelligent learning algorithm, which has the capability of nonlinear function evolution, dimension analysis and interval analysis, is called specialized AI algorithm.

Preliminary applications

Two sets of aerodynamic heating data collected from different wind tunnels are selected to demonstrate the usage of KBL method. One is of sphere heat transfer, and the other is of a double ellipsoid configuration.

Formulas learned for sphere heat transfer

The data set consists 530 measurement points distributed on the windward side of a sphere of radius 0.1 meters. The data are collected under the test conditions of two wind tunnels of different freestream flows, JF-10 for real-gas flow and JF-12 for ideal-gas flow. The original data set $\{(q_w^{(i)}; T_0^{(i)}, P_0^{(i)}, M^{(i)})\}_{i=1}^{530}$ is parameterized, localized, and standardized using inviscid CFD simulations, and extended using expert knowledge of high temperature gas dynamics. The enhanced data set is as follows.

$$\left\{ (qw, Cq, St, St_e; P, T, v, Cp, Entropy, gamma, h, k, Levis, mu, nu, pr, rou, Schmidt, c, thermalDiff, M, Re, Re_x)^{(i)} \right\}_{i=1}^{530}$$

The specialized AI algorithm is then applied to automatically search the best formula to fit the data, and balance the fitting error, formula complexity and stability. Mean Absolute Error (MAE), Pearson's correlation coefficient r and the coefficient $1-R^2$ are applied to measure the goodness of a fitting formula, where $1-R^2 = \frac{SSE}{SST}$. SSE is the Error Sum of Squares of observed and predicted values $(\sum_i (q_i^o - q_i^p)^2)$, and SST is the Total Sum of Squares of observed and average values $(\sum_i (q_i^o - \bar{q})^2)$.

Table 1. Aerodynamic Heating Formulas of sphere with boundary layer parameters

Complexity	formula	MAE	r	1-R ²
3	$q_w = k \cdot P_e$	1.4E5	0.99	0.04
9	$q_w = k \cdot T_e \cdot \sqrt{P_e}$	5.3E4	0.997	0.006
17	$q_w = k \cdot (T_e)^{\frac{5}{4}} \cdot \sqrt{P_e}$	3.9E4	0.997	0.006

From Table 1, we can see that the pressure at the edge of boundary layer P_e is a key parameters for aerodynamic heating, and the heat flux is almost linear to the square root of the pressure, i.e., $q_w : \sqrt{P_e}$. The results indicate that the thickness of boundary layer is also linear to the square root of the pressure, i.e., $\delta : \sqrt{P_e}$. This result might be helpful to study the development of boundary layer.

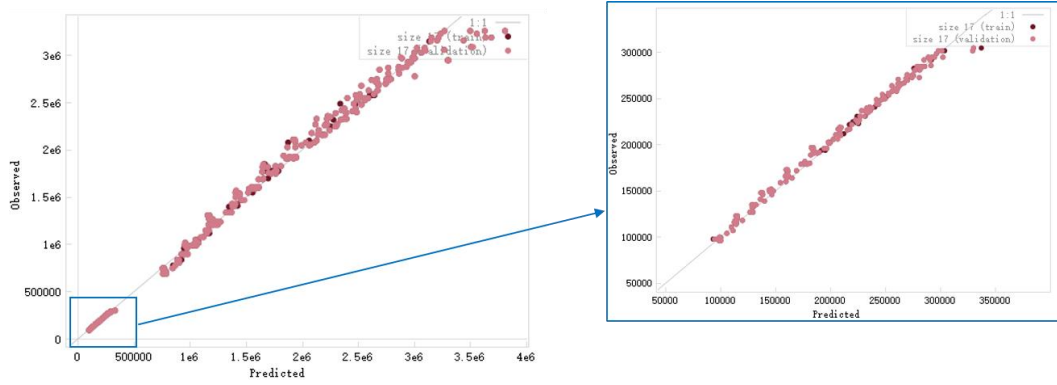
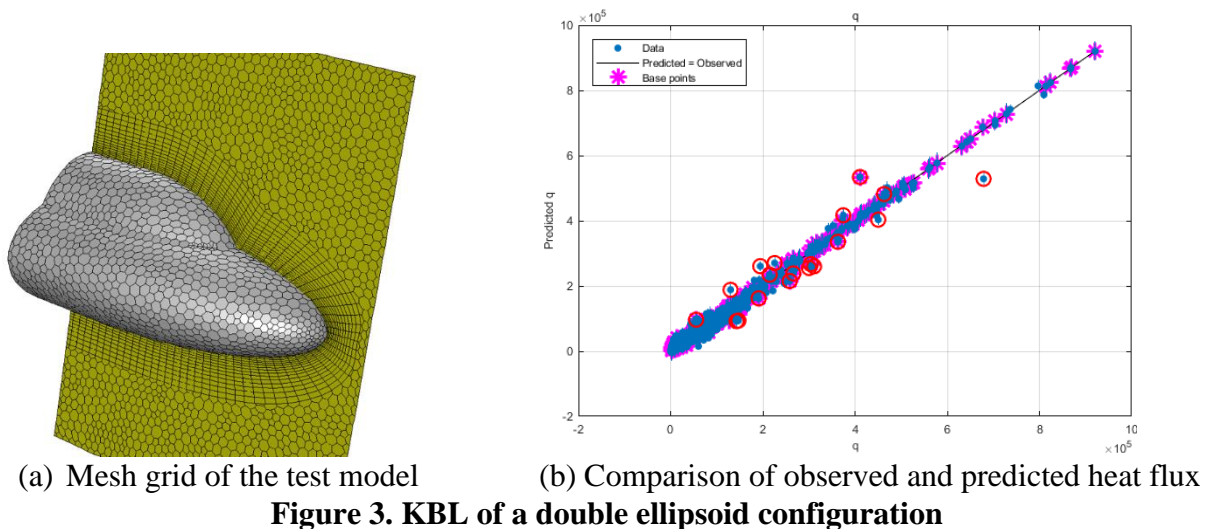


Figure 2. Comparison of observed and predicted heat flux of the sphere under the conditions of two wind tunnels (with $q_w = k \cdot (T_e)^{\frac{5}{4}} \cdot \sqrt{P_e}$), JF-10 and JF-12. The left figure shows the overall results, and the right figure is the local zoom of the left, results of JF-12 with smaller heat flux

Preliminary results of a double ellipsoid configuration

The data set consists 1044 measurement points distributed on both windward and leeward side of a double ellipsoid model of length 0.215 meters (see Fig. 3(a)). The data are collected from two wind tunnels, FD-14A and FD-20, of several tests [12]. For model and experimental parameter details, please refer to reference [12]. Similarly, the original data set is parameterized, localized, and standardized using CFD simulations (Fig. 3(a)), and extended using expert knowledge of high temperature gas dynamics. An enhanced data set is obtained. Then the specialized AI algorithm is applied to search the best formula automatically to fit the data, and balance the fitting error, formula complexity and stability. The study shows that the heat flux could be quantitatively determined by the air's temperature, pressure, speed at the outer edge of boundary layers, i.e., $q_w = f(T_e, P_e, v_e)$. The learning results are shown in Fig. 3(b). The Pearson's correlation coefficient r is 0.98 and the coefficient $1 - R^2 = 0.03$. The explicit expression is omitted here since it is very complicated at this stage due to the limitations of current AI algorithm. The improvement of AI algorithm and the simplification of the expression are left for future research.



(a) Mesh grid of the test model

(b) Comparison of observed and predicted heat flux

Figure 3. KBL of a double ellipsoid configuration

Conclusions

An intelligent optimization method, knowledge based learning (KBL), has been presented to extract formulas of aerodynamic heating. In KBL, the wind-tunnel data to be learned are regarded as a small data with strong correlation characteristics. They are preconditioned and enriched with expert knowledge and numerical simulations. Then the enhanced data are used to train and validate the candidate formulas in the guide of a special artificial intelligence (AI) algorithm, which is adapted from general AI algorithm to fit the characteristics of aerodynamic heating data. The proposed KBL method has been applied to two data sets of aerodynamic heating from different wind tunnels of different models. The study indicate that the KBL method is promising for detecting more universal laws of aerodynamic heating.

The basic idea of KBL is using specialized AI algorithm to replace conventional analytical method of manual deduction. This does make sense, and will be an inevitable trend in the near future. However, current AI algorithm still needs improving, and the data sets are not merged together. Furthermore, to get a formula that really matters, the training data set needs more data from more wind tunnels of more test models. These tasks are left for future studies.

References

- [1] Anderson, J. Jr. (2006), *Hypersonic and High-Temperature Gas Dynamics*, 2nd edn, American Institute of Aeronautics and Astronautics, Inc., Reston, Virginia.
- [2] Brunton, S. L., Proctor, J. L. and Kutz, J. N. (2016) Discovering governing equations from data by sparse identification of nonlinear dynamical systems, *Proceedings of the National Academy of Sciences of the United States of America*, 113(15), 3932-3937.
- [3] Chen, C., Luo*, C., Jiang, Z. (2018) Block building programming for symbolic regression, *Neurocomputing* 275 (31), 1973-1980.
- [4] De Jesus, and Hagan, M. T. (2007), Back propagation algorithms for a broad class of dynamic networks, *IEEE Transactions on Neural Networks* 18(1), 14-27.
- [5] Eckert, R. G. (1955), Engineering Relations for Friction and Heat Transfer to Surface in High Velocity Flow, *Journal of the Aeronautical Sciences* 22(8), 585-587.
- [6] Fay, J. A. and Riddell, F. R. (1958), Theory of Stagnation Point Heat Transfer in Dissociated Air, *Journal of the Aeronautical Sciences* 25(2), 78-85.
- [7] Horn, J. M., De Jesus, O. and Hagan, M. T. (2009), Spurious Valleys in the Error Surface of Recurrent Networks - Analysis and Avoidance, *IEEE Transactions on Neural Networks* 20(4), 686-700.
- [8] Koza, J. R. (1992) *Genetic programming: on the programming of computers by means of natural selection*, MIT Press, Cambridge, Massachusetts.
- [9] Lee, J.-H., Kim, E. T., Chang, B.-H., Hwang, I.-H., and Lee, D.-S. (2009) The accuracy of the flight derivative estimates derived from flight data, *International Journal of Aerospace and Mechanical Engineering* 3(10), 1317-1323.
- [10] Lees, I. (1956) Laminar Heat Transfer over Blunt Nosed Bodies at Hypersonic Flight Speeds, *Jet Propulsion* 26 (4), 259-269.
- [11] Lesin, A. B. and Lunev, V. V. (1994). Heat transfer peaks on a blunt-nosed triangular plate in hypersonic flow, *Fluid Dynamics* 29(2), 258-262.
- [12] Li, S. X. (2007) *Hypersonic flow characteristics around typical configurations*, National Defence Industry Press, Beijing. (in Chinese)
- [13] Luo, C., Zhang, S.-L. (2012) Parse-matrix evolution for symbolic regression, *Engineering Applications of Artificial Intelligence* 25, 1182-1193.
- [14] Luo, C., Hu, Z., Zhang, S.-L., Jiang, Z. (2015) Adaptive space transformation: An invariant based method

- for predicting aerodynamic coefficients of hypersonic vehicles, *Engineering Applications of Artificial Intelligence* 46, 93-103.
- [15] Mason, E. A., and Saxena, S. C. (1958). Approximate Formula for the Thermal Conductivity of Gas Mixtures, *Physics of Fluids* 1(5), 361-369.
 - [16] Morelli, E. A., DeLoach, R. (2003) *Wind tunnel database development using modern experiment design and multivariate orthogonal functions*, AIAA paper, 2003-0653.
 - [17] Nicoli, A., Imperatore, B., Marini, M., Catalano, P., Pizzicaroli, A., and Perigo, D. (2006) *Ground-to-flight extrapolation of the aerodynamic coefficients of the VEGA launcher*, AIAA paper, 2006-3829.
 - [18] Pulliam, T. H. (1986) Artificial dissipation models for the Euler equations, *AIAA Journal* 24(12), 1931-1940.
 - [19] Schmidt, M., and Lipson, H. (2009). Distilling free-form natural laws from experimental data. *Science* 324, 81-85.
 - [20] Vaglio L. R. (1960) Turbulent Heat Transfer on Blunt Nosed Bodies in Two-dimensional and General Three-dimensional Hypersonic Flow, *Journal of the Aerospace Sciences* 27(1): 27-38.
 - [21] Vlasov, V. I., and Gorshkov, A. B. (2001) Comparison of the Calculated Results for Hypersonic Flow Past Blunt Bodies with the OREX Flight Test Data, *Fluid Dynamics* 36(5), 812-819.
 - [22] Vlasov, V. I., Gorshkov, A. B., Kovalev, R. V. , and Lunev, V.V. (2009) Thin triangular blunt-nosed plate in a viscous hypersonic flow, *Fluid Dynamics* 44(4), 596-605.

A multi-level Method of Fundamental Solutions using quadtree-generated sources

†Csaba Gáspár¹

¹Department of Mathematics and Computational Sciences, Széchenyi István University, Győr, Hungary

†Corresponding author: csgaspar@sze.hu

Abstract

A new strategy to define source points for the Method of Fundamental Solutions is presented based on a quadtree-generated cell system controlled by the boundary of the domain in which the partial differential equation is defined. The quadtree (in 3D, octtree) algorithm results in a cell system, the spatial density of which decreases rapidly when moving away from the boundary. The sources are defined to be the cell centers of the external cells of the quadtree subdivision. This makes it possible to build up a multi-level method, where the ‘coarse’ sources generate the ‘coarse’ approximation, while the ‘fine’ (i.e. the near-boundary) sources provide the ‘fine’ approximation. On each level, the problem is discretized by using the sources belonging to the actual level only. Thus, the computational cost can be kept under an acceptable limit. Moreover, the problem of severely ill-conditioned linear systems is completely avoided.

Keywords: **Method of Fundamental Solutions, multi-level method, quadtrees**

Introduction

The Method of Fundamental Solutions (MFS, see e.g. [8]) is now a popular computational method for solving elliptic partial differential equations due to its simplicity and meshfree character and also to the fact that it is a boundary-only technique i.e. no discretization is needed inside the domain.

In its original form, the approximate solution is defined as a linear combination of the fundamental solution shifted to some external points (*source points*). Thus, the approximate solution exactly satisfies the partial differential equation to be solved. The a priori unknown coefficients of the linear combination are calculated by enforcing the boundary conditions at some *boundary collocation points*.

For instance, consider the example of the simplest 2D Laplace equation:

$$\Delta u = 0 \tag{1}$$

defined in a bounded 2D domain Ω . Suppose that Equation (1) is equipped with pure Dirichlet boundary condition:

$$u|_{\Gamma} = u_0, \tag{2}$$

where $\Gamma := \partial\Omega$, the boundary of the domain Ω . The approximate solution defined by the MFS has the form:

$$u(x) \sim \sum_{j=1}^N \alpha_j \Phi(x - \tilde{x}_j), \quad (3)$$

where Φ denotes the fundamental solution of the Laplacian (apart from a multiplicative constant):

$$\Phi(x) = \log ||x|| \quad (4)$$

and $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N$ are predefined external source points. Here $||\cdot||$ denotes the 2D Euclidean norm. The coefficients $\alpha_1, \alpha_2, \dots, \alpha_N$ can be calculated by requiring the boundary conditions. In case of Dirichlet boundary condition, this results in the following linear system of equations:

$$\sum_{j=1}^N \alpha_j \Phi(x_k - \tilde{x}_j) = u_k := u_0(x_k) \quad (k = 1, 2, \dots, M), \quad (5)$$

where x_1, x_2, \dots, x_M are predefined boundary collocation points.

The numbers of sources and the boundary collocation points need not be equal. If $N \neq M$, Equation (5) should be solved in a generalized sense using e. g. the least squares approach. For the sake of simplicity, however, in a lot of practical cases, the numbers N and M are defined to be equal, that is, Equation (5) has a square matrix. Unfortunately, though the MFS has excellent accuracy in general (see [10]), in a number of cases, the discretized linear system is severely ill-conditioned, especially when the sources are located far from the boundary. On the other hand, if they are too close to the boundary, numerical singularities appear in the approximate solution.

Another problem of the Method of Fundamental Solutions is the proper definition of the locations of sources (preferably in an automated way). In [1], [4], the sources are located along a sufficiently large circle; however, this leads to extremely ill-conditioned linear systems. In [12], the initial set of points is thinned by several strategies. See also [3], where the original boundary is transformed to the boundary of a somewhat larger domain along which the source points are located.

A popular technique is to allow the source and boundary collocation points to coincide. Thus, the problem of the proper definition of sources is automatically circumvented. However, in this approach, some singular terms generally appear, and the main problem is how to evaluate these singular terms properly or how to avoid the singularity. To treat this difficulty, a lot of special methods have been developed. The boundary knot method [2] utilizes general nonsingular solutions instead of the traditional fundamental solutions: thus, the problem of singularity is avoided, but the problem of severely ill-conditioned character of the discretized system remains the case. The situation is similar, when fundamental solutions concentrated to straight lines instead of points are used, see [6]. Using the traditional fundamental solutions, the evaluation of singular terms can be performed by special tools (regularization and desingularization techniques, see e.g. [7], [9], [11], [13]).

In this paper, we return to the traditional form of the MFS. However, the sources are generated in a completely automatic way using the well-known quadtree/octtree subdivision technique (see e.g. [5]). This algorithm produces a cell system; the individual cells belong to different levels of subdivision. The cell system exhibits automatic local refinements in the vicinity of the boundary. Taking the centers of the outer cells as source points, we obtain a point set, the spatial density of which decreases rapidly when moving away from the

boundary. This makes it possible to build up a multi-level discretization in a natural way. The method avoids also the problem of solving severely ill-conditioned systems of equations and has a relative low computational complexity as well.

A two-level technique

As a model problem, consider the 2D Dirichlet problem (1) - (2). Suppose that the boundary collocation points x_1, x_2, \dots, x_M are given. Let $\tilde{x}_1^F, \tilde{x}_2^F, \dots, \tilde{x}_N^F$ be external source points at a distance δ from the boundary (more or less equally spaced), they will be considered ‘fine level’ sources. Moreover, let $\tilde{x}_1^C, \tilde{x}_2^C, \dots, \tilde{x}_{N/2}^C$ be additional (‘coarse level’) sources at a distance 2δ from the boundary (N is supposed to be an even number). Define the approximate solution of (1) - (2) as follows:

$$u(x) \sim \sum_{j=1}^N \alpha_j^F \Phi(x - \tilde{x}_j^F) + \sum_{j=1}^N \alpha_j^C \Phi(x - \tilde{x}_j^C) \quad (6)$$

Enforcing the boundary condition in the boundary collocation points, we have:

$$\sum_{j=1}^N \alpha_j^F \Phi(x_k - \tilde{x}_j^F) + \sum_{j=1}^N \alpha_j^C \Phi(x_k - \tilde{x}_j^C) = u_k \quad (k = 1, 2, \dots, M) \quad (7)$$

In a more compact form:

$$A^F \alpha^F + A^C \alpha^C = \mathbf{u} \quad (8)$$

where A^F is an M -by- N and A^C is an M -by- $N/2$ matrix with entries:

$$A_{kj}^F = \Phi(x_k - \tilde{x}_j^F), \quad A_{kj}^C = \Phi(x_k - \tilde{x}_j^C) \quad (9)$$

The direct solution of Equation (8) is not recommended, since Equation (8) is even more ill-conditioned than the single-level equation

$$A^F \alpha^F = \mathbf{u}$$

Instead, it can (and should) be solved in an iterative way (in the sense of least squares) by splitting it into a coarse-level and a fine-level subproblem:

$$A^C \alpha^C = \mathbf{u} - A^F \alpha^F \quad (10)$$

$$A^F \alpha^F = \mathbf{u} - A^C \alpha^C \quad (11)$$

The above equations are to be solved in the sense of least squares, i.e. by solving the corresponding Gaussian normal equations:

$$(A^C)^* A^C \alpha^C = (A^C)^* (\mathbf{u} - A^F \alpha^F) \quad (12)$$

$$(A^F)^* A^F \alpha^F = (A^F)^* (\mathbf{u} - A^C \alpha^C) \quad (13)$$

The main idea of the method is that if the coarse subproblem is already solved, then, in order to solve the fine level subproblem, it is sufficient to apply some steps of the familiar (conjugate) gradient method, which significantly reduces the computational complexity.

Remark: Without going into deep details, the idea behind the method is as follows. The solution of the coarse level subproblem (nearly) eliminates the low-frequency error components from the approximate solution. Thus, the fine level operator maps the subspace of the high-frequency components into itself. It can be shown that the fine level operator

restricted to the high-frequency subspace is uniformly well-conditioned (independently of the fineness of the discretization). Thus, though the (conjugate) gradient method converges slowly, if the corresponding operator is not well-conditioned, the high-frequency error components are damped much more efficiently.

By introducing additional sources on even coarser levels, the method can be extended to a multi-level technique in a straightforward way. At the coarsest level, the corresponding subproblem should be solved exactly. In practice, it is often sufficient to apply several (conjugate) gradient iterations at the coarsest level as well.

Automatic generation of source locations using quadtrees

To build up a multi-level method outlined above, several *groups* of sources are needed. The greater the distance from the boundary, the lower the spatial density of the sources is. The quadtree algorithm produces point sets with exactly the same property. Recall that the quadtree subdivision is a systematic, recursively defined subdivision of an initial square controlled by a finite set of points (controlling points). A subsquare is divided into four congruent subsquares (cells), if the number of controlling points contained in the actual subsquare exceeds a predefined minimal value, provided that the level of subdivision remains under a predefined maximal level. This results in automatic local refinements in the vicinity of the controlling points. By additional subdivisions, it can be assured that the ratio of the neighboring cell sizes is at most 2, i.e. no abrupt changes in cell sizes occur. Note that in 3D, the procedure is similar: here an initial cube is divided recursively into eight congruent subcubes (octtree algorithm). Note also that the obtained cell system is suitable for defining simple finite volume schemes as well (see e.g. [5]), however, here it is used to define source point locations only.

In the presented multi-level technique, the quadtree subdivision is controlled by the boundary of the domain of the original partial differential equation, more precisely, by the predefined boundary collocation points. Having created the quadtree cell system, the source points are defined to be the centers of the external cells. The cell centers belonging to low levels of subdivision are considered ‘coarse level’ sources, while the (near-boundary) cell centers belonging to high levels of subdivision are regarded as “fine level” sources.

Numerical examples

The above outline method is demonstrated through two simple examples.

Example 1. Let Ω be a circle centered at the midpoint of the unit square with radius 0.3. Consider the test solution of the Laplace equation

$$u(x, y) = e^{4\pi x} \cdot \sin 4\pi y, \quad (14)$$

where the more familiar notations x, y are used for the space variables. The Laplace equation (1) is supplied with Dirichlet boundary condition consistent with the above test solution. Figure 1 shows the quadtree cell system controlled by the boundary Γ of the domain and the source point locations as well. The maximal subdivision level was 8, i.e. the smallest cell size was $1/256$. Table 1 shows the relative L_2 -errors of the above outlined two-level method for different numbers of sources calculated on the boundary of the domain. Here L_{coarse} and L_{fine} are the quadtree subdivision levels of the coarse and fine sources, respectively, while

N_{coarse} and N_{fine} denote the numbers of sources at the coarse (resp. fine) level. The number of boundary collocation points was always = 476 .

The results demonstrate that the accuracy is acceptable. Note, however, that the numerical complexity is much less than that of the traditional direct method.

Table 1. Two-level MFS, relative boundary L_2 -errors. Domain: circle

L_{coarse}/L_{fine}	3/4	4/5	5/6	6/7
N_{coarse}/N_{fine}	12/88	88/104	104/216	216/376
Relative L_2 -error (%)	0.16591	0.04687	0.01604	0.02571

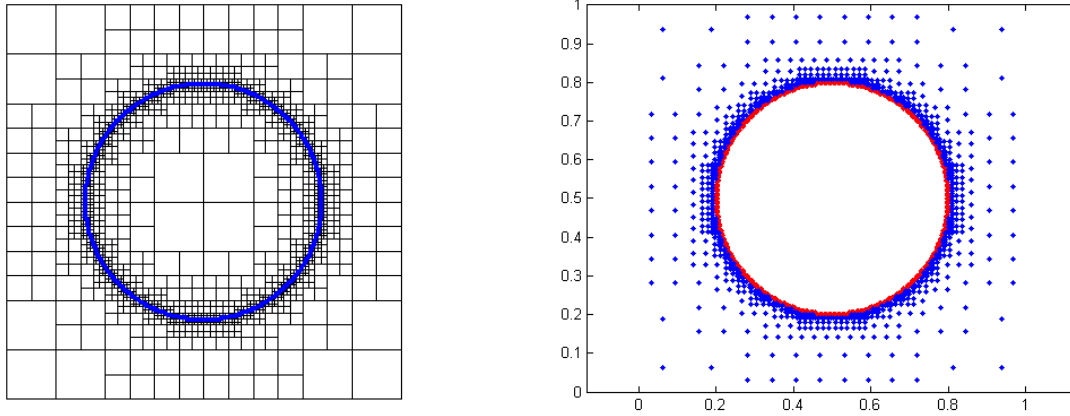
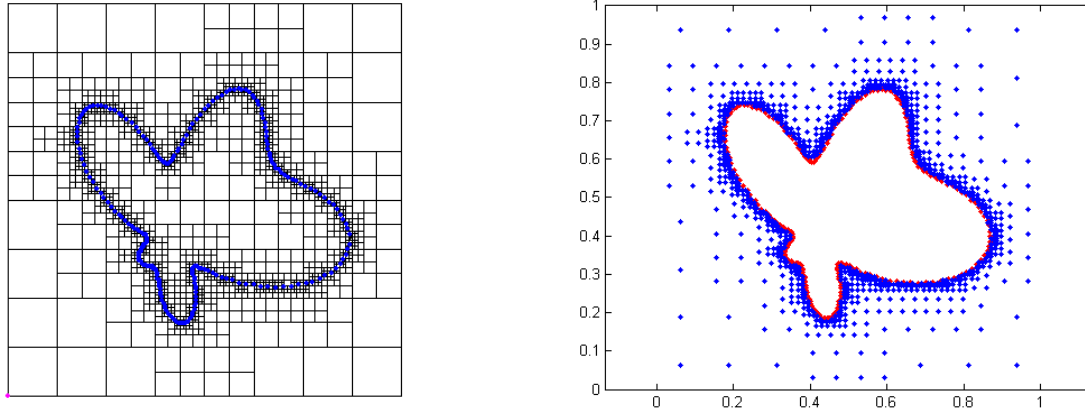


Figure 1. A quadtree cell system generated by a circle and the external source points

Example 2. Let Ω be an amoeba-shaped domain contained in the unit square. Figure 2 shows the quadtree cell system controlled by the boundary Γ of the domain and the source point locations as well. The maximal subdivision level was again 8. The test solution (14) was the same as in Example 1. The number of boundary collocation points was always $M = 236$. Table 2 shows the relative L_2 -errors of the above outlined two-level method calculated on the boundary of the domain. L_{coarse} and L_{fine} are the quadtree subdivision levels of the coarse and fine sources, respectively. N_{coarse} and N_{fine} denote the numbers of sources at the coarse (resp. fine) level. Due to the more complicated geometry, the accuracy is now somewhat less than in Example 1, but it is still acceptable.

Table 2. Two-level MFS, relative boundary L_2 -errors. Amoeba-shaped domain

L_{coarse}/L_{fine}	3/4	4/5	5/6	6/7
N_{coarse}/N_{fine}	19/70	70/122	122/227	227/480
Relative L_2 -error (%)	1.3801	0.12707	0.10469	0.06078

**Figure 2. A quadtree cell system generated by an amoeba-like curve and the external source points**

Summary and conclusions

The traditional Method of Fundamental Solutions has been revisited. The sources are defined in a completely automatic way using the quadtree/octtree subdivision algorithm. This algorithm generates sources, the spatial density of which is greater in the vicinity of the boundary and becomes low far away from the boundary. These groups of sources result in multi-level MFS-based approximations. As a smoothing procedure the classical (conjugate) gradient method was used. The number of boundary collocation points was always greater than that of the sources at any level, so that the MFS-equations were solved in the sense of least squares, i.e. the Gaussian normal equations were taken into account. The accuracy of the method has been proved acceptable. At the same time, the computational complexity of the method is much less than that of a traditional direct solver. Moreover, the problem of the severely ill-conditioned algebraic system is also avoided.

Acknowledgement: The research was partly supported by the European Union in the framework of the project EFOP-3.6.2-16-2017-00015.

References

- [1] Alves, C. J. S., Chen, C. S. and Šarler, B., The method of fundamental solutions for solving Poisson problems. *Int. Series on Advances in Boundary Elements. (Proceedings of the 24th International Conference*

- on the Boundary Element Method incorporating Meshless Solution Seminar. Sintra, Portugal, 2002, C.A.Brebbia A.Tadeu, V.Popov, Eds. Vol. 13. WitPress, Southampton, Boston, 67–76.
- [2] Chen, W., Shen, L. J., Shen, Z. J. and Yuan, G. W. (2005) Boundary Knot Method for Poisson Equations. *Engineering Analysis with Boundary Elements* **29**, 756–760.
- [3] Chen, C. S., Karageorghis, A. and Li, Y. (2016) On choosing the location of the sources in the MFS. *Numerical Algorithms* **72**, 107–130.
- [4] Fam, G. S. A. and Rashed, Y. F., A Study on the Source Points Locations in the Method of Fundamental Solutions. *Int. Series on Advances in Boundary Elements. (Proceedings of the 24th International Conference on the Boundary Element Method incorporating Meshless Solution Seminar. Sintra, Portugal, 2002, C.A.Brebbia A.Tadeu, V.Popov, Eds. Vol. 13. WitPress, Southampton, Boston, 297–312.*
- [5] Gáspár, C. (2004) A Meshless Polyharmonic-type Boundary Interpolation Method for Solving Boundary Integral Equations. *Eng. Anal. with Boundary Elements*, **28**, 1207–1216.
- [6] Gáspár, C. (2013) Some Variants of the Method of Fundamental Solutions: Regularization Using Radial and Nearly Radial Basis Functions. *Central European Journal of Mathematics* 11/8, 1429–1440.
- [7] Gáspár, C. (2015) A Regularized Multi-level Technique for Solving Potential Problems by the Method of Fundamental Solutions. *Engineering Analysis with Boundary Elements* **57**, 66–71.
- [8] Golberg, M. A. (1995) The Method of Fundamental Solutions for Poisson’s Equation. *Engineering Analysis with Boundary Elements* **16**, 205–213.
- [9] Gu Y., Chen W. and Zhang J. (2012) Investigation on Near-boundary Solutions by Singular Boundary Method. *Engineering Analysis with Boundary Elements* **36**, 1173–1182.
- [10] Li, X. (2005) On convergence of the method of fundamental solutions for solving the Dirichlet problem of Poisson's equation. *Advances in Computational Mathematics* **23**, 265–277.
- [11] Šarler, B., A Modified Method of Fundamental Solutions for Potential Flow Problems. *The Method of Fundamental Solutions - A Meshless Method. C.S.Chen, A.Karageorghis, Y.S.Smyrlis Eds. Dynamic Publishers, Inc., Atlanta, USA, 2008, 299–321.*
- [12] Schaback, R., Adaptive Numerical Solution of MFS Systems. *The Method of Fundamental Solutions - A Meshless Method. C.S.Chen, A.Karageorghis, Y.S.Smyrlis Eds. Dynamic Publishers, Inc., Atlanta, USA, 2008, 1–27.*
- [13] Young, D. L., Chen, K. H. and Lee, C.W. (2005) Novel Meshless Method for Solving the Potential Problems with Arbitrary Domain. *Journal of Computational Physics* **209**, 290–321.

FEM non-linear modelling of cob using ANSYS

*†Jimenez Alejandro¹, O'Dwyer Dermot¹

¹Department of Civil Structural and Environmental Engineering, Trinity College Dublin, Ireland.

*†Presenting & Corresponding author: jimnezra@tcd.ie

Abstract

Cob is an earthen construction technique used to build monolithic load-bearing walls. Mainly of vernacular nature, remaining cob buildings can be found throughout Europe as well as in other specific locations around the world.

The aim of this paper is to evaluate the available material models' suitability for the modelling of cob's structural behavior in one of the most commonly used FEM software in the market. Previous stress-strain graphs obtained and failure mechanisms observed after a simple compression and diagonal compression experimental campaign in cob wallettes were replicated using three different material constitutive models, namely, MISO, CONCR, and DMGE/DMGI. Furthermore, a mesh size sensitivity analysis was performed following a mesh refinement approach.

MISO could reproduce the pre-peak behavior of cob and principal stresses could be used as an indication of the opening of cracks. On the other hand, it did not capture the softening post-peak behavior of the material. CONCR provided quite accurate pre-peak behavior results and peak strength values. Moreover, it was possible to plot the opening of cracks and those plots agreed with the experimental results. Nevertheless, as CONCR is suitable to reproduce brittle failures, it did not capture the long deformations characteristic of cob. Finally, the DMGE/DMGI proved to be inaccurate to reproduce both pre-peak and post-peak behavior. Although parameters could be calibrated to obtain the appropriate peak strength, neither stresses nor did strains corresponded to what was observed in the experimental campaign.

Keywords: **Non-linear, FEM, Cob, ANSYS**

Introduction

Cob is an earthen construction technique used to build monolithic load-bearing walls. The loam, mixture of soil, water and straw, is placed wet in a horizontal layer which after being left to dry for some time is shaped with sharp instruments into its final form. New layer is placed in top of the previous one and the process is repeated until the desired wall height is reached.

Monolithic techniques have the advantage of not presenting weakness planes such as it is in the case of modular constructions (namely adobe or rammed earth). Cob's cohesion is provided mainly by the clay cementing properties and the added organic fibers such as straw or heather [1]. Despite the low compressive strength of cob, this material presents a relatively good performance regarding the shear strength. Moreover, cob appears to be more flexible in comparison with the other earthen construction techniques, since it presents a relatively ductile post-peak behavior due to the fibers added to the mixture [1].

Cob remaining buildings are of vernacular nature and can be found in eighteen European countries [2]. According to Hamard et al. only within Germany, France and the UK there are at least 200 000 cob buildings [3]. Furthermore, the bulk of remaining earthen buildings, cob included, is located within the dry climate regions of the world [4] and in total earthen buildings house at least 30 % of the world's population [5].

Despite its importance, cob has not received as much attention by researches as other earthen techniques such as adobe, rammed earth or compressed earth blocks (CEB). Therefore, cob's structural behavior is not yet fully understood, nor material constitutive models have been specifically developed to simulate cob's structural response. Moreover, neither standards are available for the design of new cob buildings nor for the conservation of existent ones. After an extensive research that involved 55 documents related with the normalization and standardization of earthen construction techniques around the world, Cid et al. [6] identified that none of them was specialized on providing guidance for cob.

The purpose of a non-linear model is to identify the peak strength of a structure and reproduce more accurately its pre-peak and post-peak behavior. By doing so, a better safety evaluation of existent buildings can be performed thus avoiding the implementation of over conservative intervention measures that may cause the loss of their authenticity. This paper aims at assessing the suitability of the material constitutive models available in ANSYS [7], to replicate the non-linear response of cob wallettes. The simulations are based on the experimental campaign and the numerical simulations performed previously by Miccoli et al. [1] which represent one of the most complete and detailed studies of cob at the moment.

Methodology

Miccoli et al. [8] determined experimentally the stress-strain curves of cob (as well as for adobe and rammed earth) both under simple compression and under diagonal compression. They also reported the failure mechanisms and the crack patterns of the tested wallettes (see Figure 1). Pull-off test were also carried out to determine cob's tensile strength. All cob's mechanical properties determined by Miccoli et al. are summarized and presented in Table 1.

Their numerical simulations to describe cob's structural behavior consisted in 2D plane stress models. They implemented a macro-modelling approach with a TSRCM (Total Strain Rotating Crack Model) constitutive model and a multilinear definition of the stress-strain relationship for the compressive behavior with an initial linear segment of $0.3 f_c$ and a post peak segment with negative slope. An exponential relationship was employed for the tensile behavior [8]. The software used was Diana [9].

Miccoli's et al. tests were simulated in this paper using ANSYS [7]. The values for the geometry and material properties were adopted as those reported by them [8] to replicate as accurately as possible the non-linear response of cob observed during their experimental campaign. Three different material constitutive models were employed, MISO, CONCR, and DMGE/DMGI.

MISO, which stands for multilinear isotropic hardening, is a rate-independent plasticity model characterized by a Von Mises yield criterion [10], an associative flow rule, and an isotropic hardening in which the yield surface remains centered about its initial centerline and expands in size as the plastic strains develop [11]. It is supported by plane and solid finite elements. The stress strain multilinear behavior and the initial and subsequent yield surfaces for isotropic hardening plasticity are shown in Figure 2. MISO does not support the definition of negative

slopes for the stress-strain relationship. A post-peak horizontal curve was defined with constant stress equivalent to the compressive strength of cob reported in Table 1.

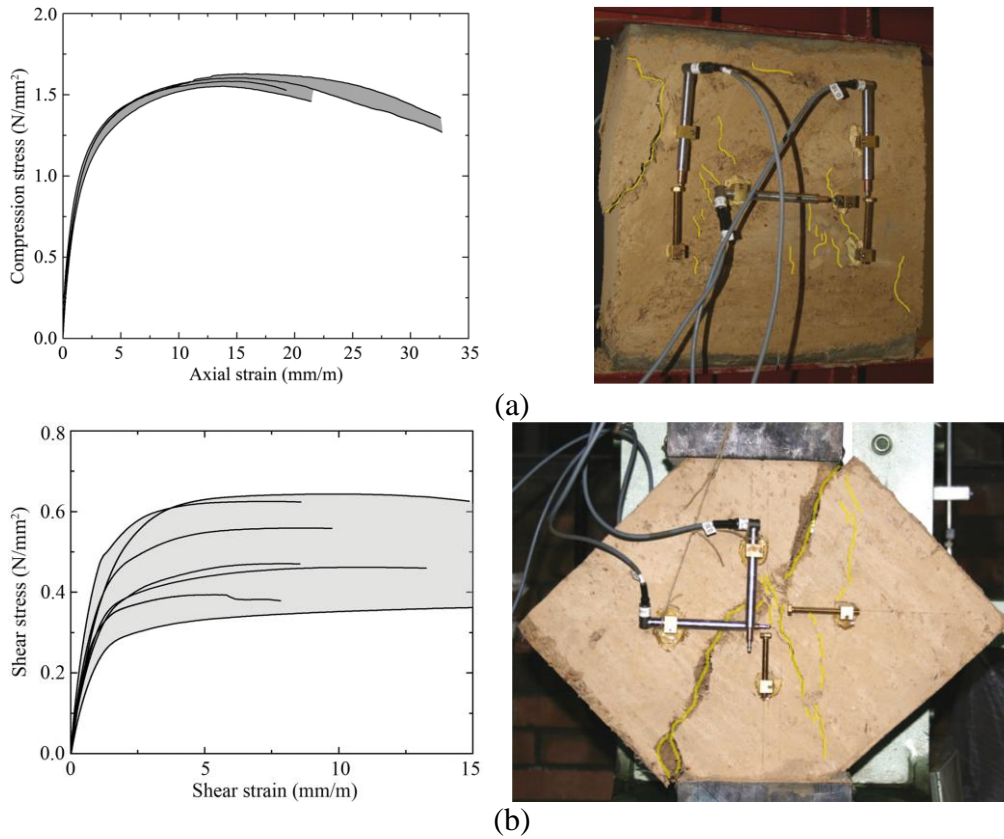


Figure 1. Stress-strain curves and crack patterns of cob wallettes under: (a) simple compression & (b) diagonal compression [8].

Table 1. Cob's mechanical properties [8].

Property	Value
Compressive strength f_c (MPa)	1.59
Tensile strength f_t (MPa)	$(0.10-0.16) f_c$
Tensile fracture energy (N/mm)	$(0.3-0.8) f_t$
Shear strength (MPa)	0.5 [1]
Shear modulus (MPa)	420 [1]
Modulus of elasticity (MPa)	1021
Poisson's ratio (-)	0.14
Density (kg/m^3)	1475

CONCR is a material constitutive model that can only be applied in combination with the legacy element called SOLID65 [10]. SOLID65 is an eight-node 3D element capable of cracking in tension and crushing in compression. It is suitable to model geological materials and reinforced composites [12]. The CONCR material model predicts the failure of brittle materials. The criterion for failure due to a multiaxial stress state can be expressed as follows:

$$\frac{F}{f_c} - S \geq 0 \quad (1)$$

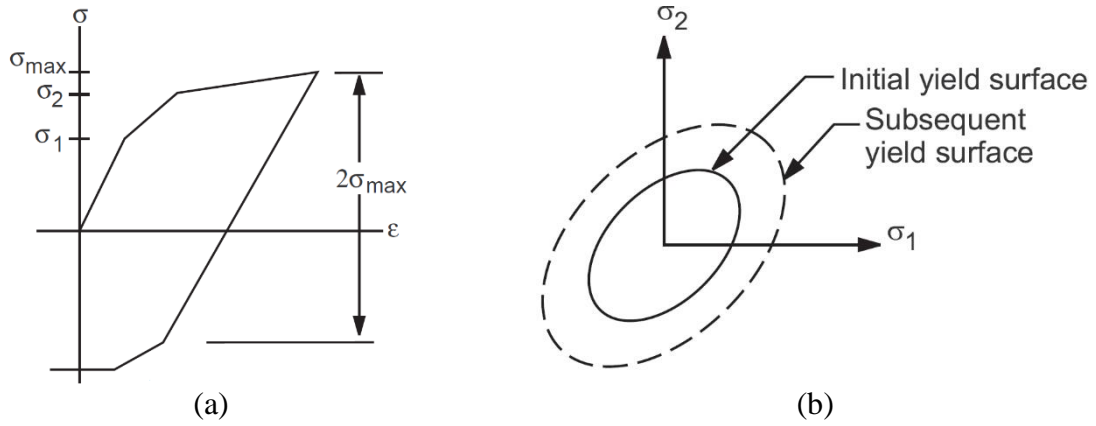


Figure 2. (a) Stress-strain multilinear isotropic behavior. (b) Initial and subsequent yield surfaces for isotropic hardening plasticity [11].

Where:

- F = Function of principal stress state ($\sigma_{xp}, \sigma_{yp}, \sigma_{zp}$).
- S = Failure surface expressed in terms of principal stresses and five input parameters:
 - f_c = Ultimate uniaxial compressive strength (taken from Table 1).
 - f_t = Ultimate uniaxial tensile strength (taken from Table 1).
 - f_{cb} = Ultimate biaxial compressive strength ($f_{cb} = 1.2f_c$).
 - f_1 = Ultimate compressive strength for a state of biaxial compression superimposed on hydrostatic stress state ($f_1 = 1.45f_c$).
 - f_2 = Ultimate compressive strength for a state of uniaxial compression superimposed on hydrostatic stress state ($f_2 = 1.725f_c$).

The presence of a crack at an integration point is represented through modification of the stress-strain relations by introducing a plane of weakness in a direction normal to the crack face. Also, a shear transfer coefficient β_t is introduced which represents a shear strength reduction factor for those subsequent loads which induce sliding (shear) across the crack face. If the crack closes, then all compressive stresses normal to the crack plane are transmitted across the crack and only a shear transfer coefficient β_c for a closed crack is introduced [11]. This condition can be seen in Figure 3. β_t , β_c and T_c , multiplier for amount of tensile stress relaxation, were calibrated to obtain a response from the wallettes as similar as possible to the one reported in the referenced experimental campaign (the values adopted were 0.25, 0.9 and 0.8 respectively).

If the material at an integration point fails in uniaxial, biaxial, or triaxial compression, the material is assumed to crush at that point. In SOLID65, crushing is defined as the complete deterioration of the structural integrity of the material [11]. To obtain a more detailed description of CONCR see [13].

DMGI, which stands for damage initiation, determines the onset of material damage under loading. It needs to be used in combination with DMGE, damage evolution, which defines the way in which the material degrades once the damage has started [10]. A Hashin criteria was adopted to determine the DMGI material model with a continuum damage mechanics method (not supported by any 3D finite element). This physical failure criteria accounts for four damage modes, namely, fiber tension (rupture), fiber compression (kinking), matrix tension (cracking), and matrix compression (crushing) [11].

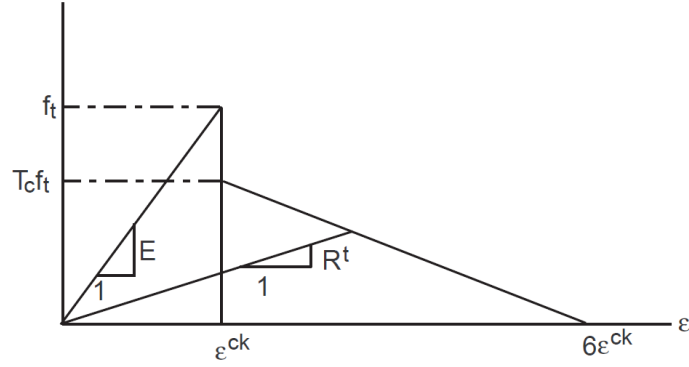


Figure 3. CONCR model for cracked condition [11].

For the continuum damage mechanics method damage variables increase gradually based on the energy amounts dissipated for the various damage modes. To achieve an objective response, the dissipated energy for each damage mode is regularized as follows [11]:

$$g_v = \frac{G_c}{L_e} \quad (2)$$

Where:

- g_v = Energy dissipated per unit volume.
- L_e = Characteristic length of the element calculated from the element area, A, as equal to:

$$L_e = \begin{cases} 1.12\sqrt{A}, & \text{for a square element} \\ 1.52\sqrt{A}, & \text{for a triangular element} \end{cases} \quad (3)$$

- G_c = Energy dissipated per unit area which for an specific damage mode is given by:

$$G_c = \int_0^{U_e^f} \sigma_e dU_e \quad (4)$$

Where:

- σ_e = Equivalent stress.
- U_e = Equivalent displacement.
- U_e^f = Ultimate equivalent displacement, where total material stiffness is lost for the specific mode.

Viscous damping coefficients η are also specified respectively for all four damage modes. For a specific damage mode, the damage evolution is regularized as follows [11]:

$$d'_{t+\Delta t} = \frac{\eta}{\eta + \Delta t} d'_t + \frac{\Delta t}{\eta + \Delta t} d_{t+\Delta t} \quad (5)$$

Where:

- $d'_{t+\Delta t}$ = Regularized damage variable at current time.
- d'_t = Regularized damage variable at the end of the last sub step.
- $d_{t+\Delta t}$ = Unregularized current damage variable.

Hashin's fiber and matrix failure criterion are described according to EQ. (6) and to EQ. (7) respectively.

$$\xi_4 = \begin{cases} \left(\frac{\sigma_x}{\sigma_{xt}^f} \right)^2 + \frac{\sigma_{xy}^2 + \sigma_{xz}^2}{(\sigma_{xy}^f)^2}, & \text{if } \sigma_x > 0 \\ \left(\frac{\sigma_x}{\sigma_{xc}^f} \right)^2, & \text{if } \sigma_x \leq 0 \end{cases} \quad (6)$$

$$\xi_5 = \begin{cases} \left(\frac{\sigma_y + \sigma_z}{\sigma_{yt}^f} \right)^2 + \frac{\sigma_{yz}^2 - \sigma_y \sigma_z}{(\sigma_{yz}^f)^2} + \frac{\sigma_{xy}^2 + \sigma_{xz}^2}{(\sigma_{xy}^f)^2} & \text{if } \sigma_y + \sigma_z > 0 \\ \frac{1}{\sigma_{yc}^f} \left(\left(\frac{\sigma_{yc}^f}{2\sigma_{yz}^f} \right)^2 - 1 \right) (\sigma_y + \sigma_z) + \left(\frac{\sigma_y + \sigma_z}{2\sigma_{yz}^f} \right)^2 + \frac{\sigma_{yz}^2 - \sigma_y \sigma_z}{(\sigma_{yz}^f)^2} + \frac{\sigma_{xy}^2 + \sigma_{xz}^2}{(\sigma_{xy}^f)^2} & \text{if } \sigma_y + \sigma_z \leq 0 \end{cases} \quad (7)$$

Orthotropic elasticity parameters, Young's modulus, Poisson's ratio and Shear modulus, were defined according to the values presented in Table 1 assuming same value for all directions (x, y and z). Similarly, orthotropic stress limits, tensile, compressive and shear, were assumed to have same value in all directions. Finally, the dissipated energy and viscous damping coefficients values used for the simulations are those presented in Table 2.

Table 2. Dissipated energy and viscous coefficient values for the DMGE/DMGI models.

Material constant	Meaning	Value
C1	Energy dissipated per unit area from tensile fiber damage (N/m).	1.0×10^{10}
C2	Viscous damping coefficient for tensile fiber damage.	0.001
C3	Energy dissipated per unit area from compressive fiber damage (N/m).	1.0×10^{10}
C4	Viscous damping coefficient for compressive fiber damage.	0.001
C5	Energy dissipated per unit area from tensile matrix damage (N/m).	1.035×10^5
C6	Viscous damping coefficient for tensile matrix damage.	0.37
C7	Energy dissipated per unit area from compressive matrix damage (N/m).	7.950×10^5
C8	Viscous damping coefficient for compressive matrix damage.	0.37

The values of the non-participating modes (C1 and C3) were set at a relatively high value to avoid their interference in the study of the cob typical damage modes, namely, cracking and crushing (C5 and C7), as advised in [14]. The values of the correspondent viscous damping coefficient for tensile and compressive fiber damage (C2 and C4) were randomly assigned as they are negligible. Whereas that those for tensile and compressive matrix damage (C6 and C8) were calibrated to obtain the reported peak strength in the simple compressive test.

The dimensions of the FEM models are shown in Figure 4. To simulate the compressive tests a displacement control approach was used. The steel plates in the bottom were fixed whereas that vertical displacements were applied to the plates on top of the wallettes. Self-weight was neglected. For the MISO model, large simulations were taken into account. On the other hand,

no large deformations were used for CONCR as advised in [12] as they would cause convergence problems. A summary of FEM simulations is presented in Table 3.

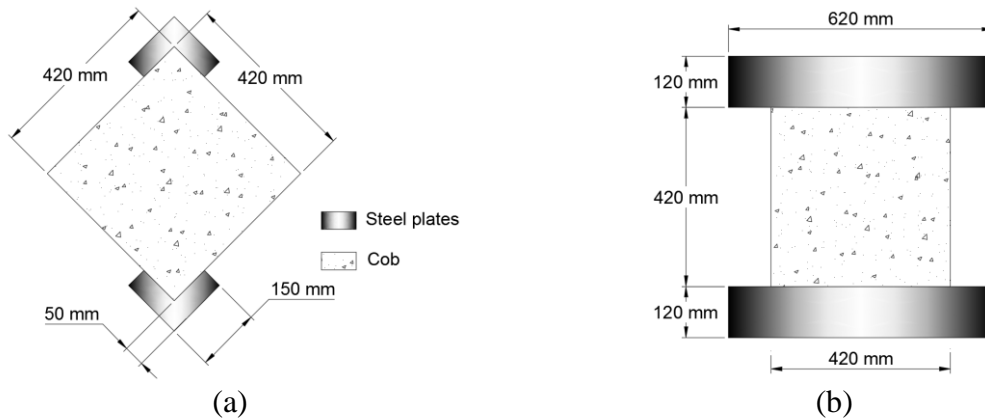


Figure 4. Dimensions used for: (a) diagonal compression and (b) simple compression simulations.

For the mesh sensitivity analysis, a mesh refinement method was selected. The maximum element sizes studied were 50, 30 and 10 mm. The element shape was verified with the mesh quality metrics provided by the software. Thanks to the relatively simple geometry of the models, the values for the average mesh quality were located between 0.86 and 0.99 for all simulations. Finally, the element type for each one of the simulations performed is indicated in Table 3.

Table 3. Finite element analyses set-ups.

Material	Geometry	Finite element (# of nodes)	Large deformations
MISO	2D	PLANE183 (8)	On
	3D	SOLID186 (20)	On
CONCR	3D	SOLID65 (8)	Off
DMGE/DMGI	2D	PLANE183 (8)	Off

Results and discussion

The results obtained from the simulations performed using ANSYS are presented in the form of stress-strain graphs and maximum principal stress plots. The upper and lower limits of the stress-strain graphs represent the experimental envelope reported by Miccoli et al. [8]. Besides, crack and crushing plots are shown for the CONCR models. Finally, a table is presented to show the mesh sensitivity using as reference the values obtained for the peak strengths of each simulation.

Plane stress and 3D simple compression stress-strain curves using MISO are shown in Figure 5 and Figure 6 respectively. The plane stress model reproduced quite accurately the pre-peak behavior of cob. Unfortunately, as this plasticity model is not capable to reproduce post-peak softening, it was not able to follow the loss of strength of the material after the maximum stress was attained. Regarding the mesh sensitivity, same behavior path was reproduced for the three element sizes. However, the finer the mesh the larger were the values obtained for the strains. The discrepancy values computed between the peak strengths obtained and the reference value are smaller than 10 % for the three mesh sizes implemented as can be seen in Table 4.

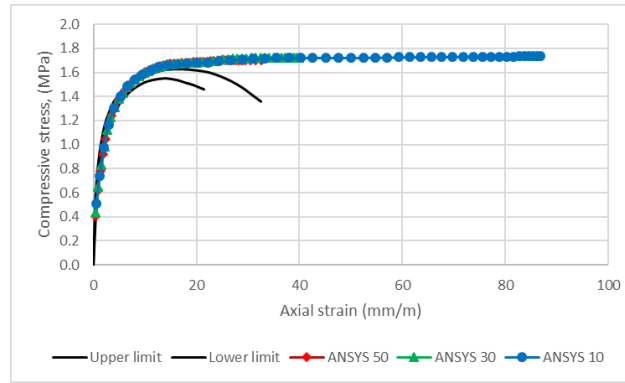


Figure 5. Plane stress simple compression stress-strain curves using MISO.

The 3D MISO model results were not as accurate as the ones obtained from the plane stress one. It presented a stiffer behavior and the plastic strains started to develop at a higher stress. The plots produced do not fit with the experimental ones. As 3D models include the Poisson's ratio effect in the orthogonal directions to the plane of the wall, the stress-strain multilinear curve used, which was calibrated with a plane stress simulation that neglects such effect, turns out to be inadequate. Table 4 presents discrepancy values for the peak strengths obtained for the 3D MISO model above 20 and 30 % with respect to the reference value for the different mesh sizes implemented.

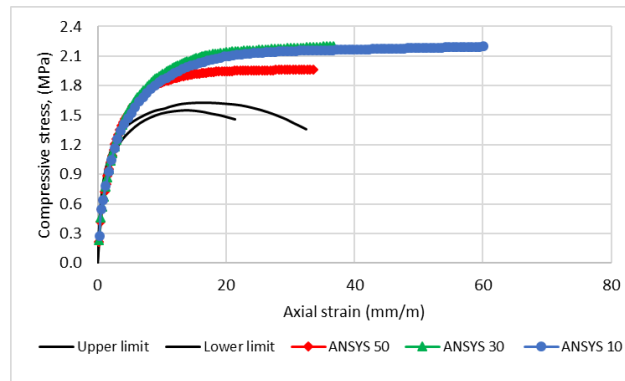


Figure 6. 3D simple compression stress-strain curves using MISO.

MISO models do not provide directly the appearance of cracks nor crushing within the material. As an indication for such failure modes the maximum principal strains, which are shown in Figure 7, can be interpreted (only the plots obtained for the 10 mm models are presented for the sake of brevity but similar patterns were found for the 30 and 50 mm models). As it can be seen, both plane stress and 3D models show a symmetric x pattern of the strains which is typical on simple compression tests of brittle materials and is considered as a satisfactory failure pattern [15].

The obtained stress-strain curves from the diagonal compression simulations are shown in Figure 8 and Figure 9 for the plane stress and 3D models respectively. Wider scattering was reported by Miccoli et al. for the shear response of the cob wallets as can be seen from the upper and lower limit curves. All plane stress MISO models fit within such limits. Unfortunately, the discrepancy values for the peak strengths are slightly high, between 13 and 20 %, as can be seen in Table 4.

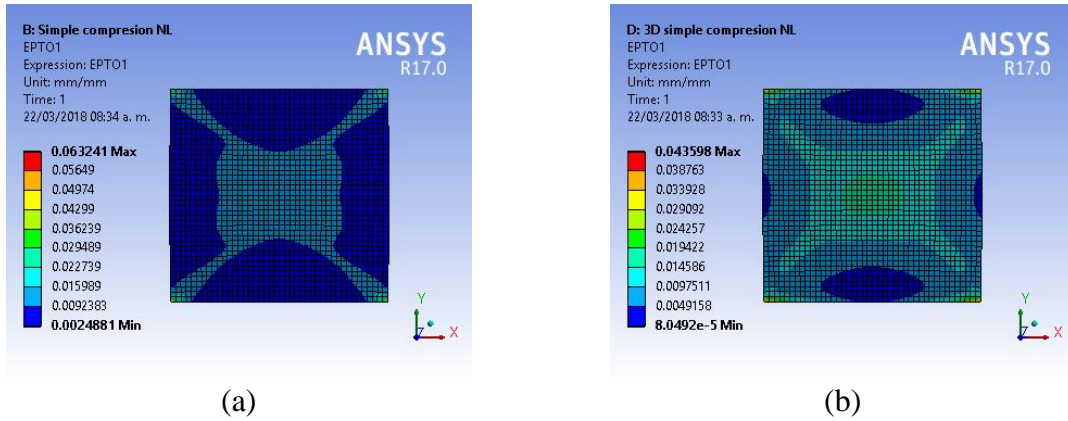


Figure 7. Maximum principal strains for the simple compression MISO models (a) plane stress & (b) 3D.

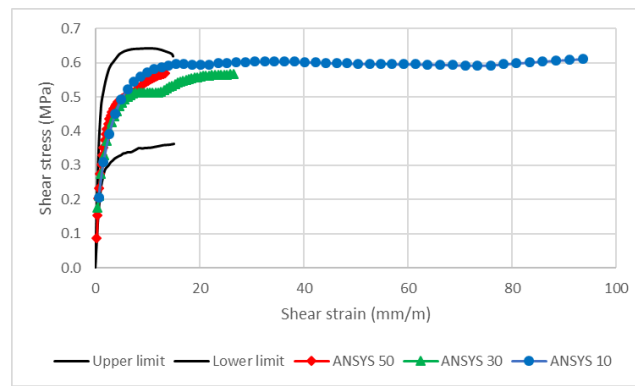


Figure 8. Plane stress diagonal compression stress-strain curves using MISO.

On the other hand, 3D MISO stress-strain curves do not fit within the reference range. These results are similar as those reported for the 3D MISO simple compression curves. A stiffer behavior can be observed which may be explained by the Poisson's ratio effect in the orthogonal directions of the wall plane for the 3D models. Moreover, an important difference between the finer mesh (10 mm) and the coarser ones (30 & 50 mm) can be seen. Discrepancy values for the peak strengths are within 39 and 41 % for the 30 and 50 mm models whereas that for the 10 mm model this value increases up to 76 % (see Table 4).

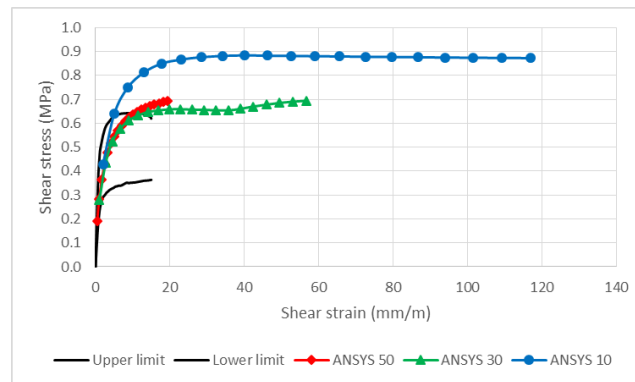


Figure 9. 3D diagonal compression stress-strain curves using MISO.

Figure 10 shows the maximum principal strains from the diagonal compression simulations. Due to the geometry of the models, singularity points appear at the contact between the sharpened

edges of the steel plates with the cob wallettes. These figures do not match with the failure pattern observed during the experimental campaigns and place serious doubts regarding the validity of the results obtained. In future work, those singularities must be removed for example by rounding the edges of the plates or by applying the displacements directly on the edge nodes of the cob wallettes.

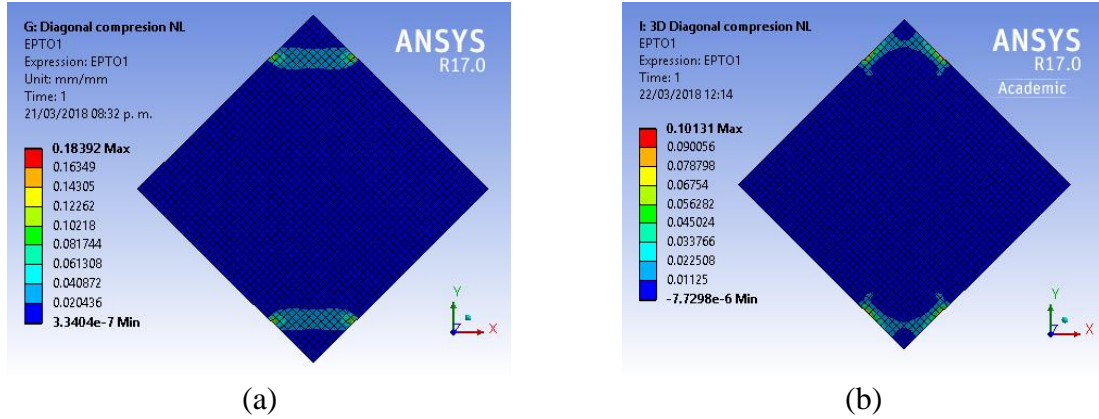


Figure 10. Maximum principal strains for the diagonal compression MISO models (a) plane stress & (b) 3D.

The stress-strain curves obtained with the CONCR model for the simple compression simulation are shown in Figure 11. They display a slightly stiffer pre-peak behavior in comparison with the experimental response of cob. Nonetheless, the overall behavior turned out to be quite accurate. Regarding the mesh sensitivity results, discrepancy values for the peak strengths were computed within 10 % for the three mesh sizes implemented.

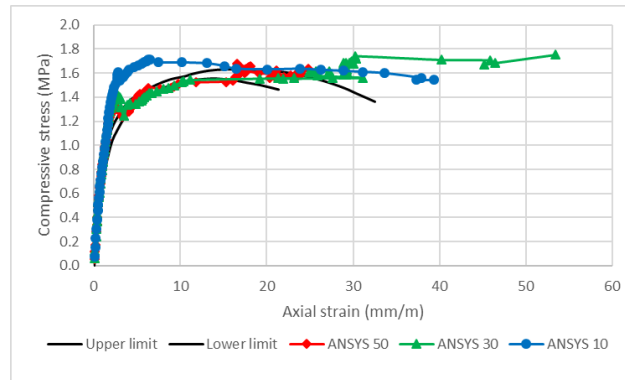


Figure 11: 3D simple compression stress-strain curves using CONCR.

The CONCR material model, in combination with the finite element SOLID65, has the capability to represent cracks and/or crushing explicitly in a graphical way. Figure 12 shows those plots at three different stages of the simulation, namely, at substeps 40, 50 and 100 (last substep). It can be appreciated how the failure starts at the corners of the wallette (a), propagates to the center forming the typical cone shape of compression tests (b) and finally reaches the total damage of the material (c).

The simulation of the diagonal compression test with CONCR was more mesh sensitive as can be seen in Figure 13. None of the three models was capable to reproduce the post-peak ductile behavior of cob reported after the experimental campaign. Moreover, the models with 30 and 50 mm mesh size gave quite big discrepancy values regarding the peak strengths of around 47 %. On the other hand, the 10 mm model gave a very accurate peak strength value. Even though

it presented a relatively brittle post peak behavior, the values for the final strains are similar to those reported by Miccoli et al. This model represents the more accurate way to reproduce the non-linear behavior of cob from the three different material models studied in this paper.

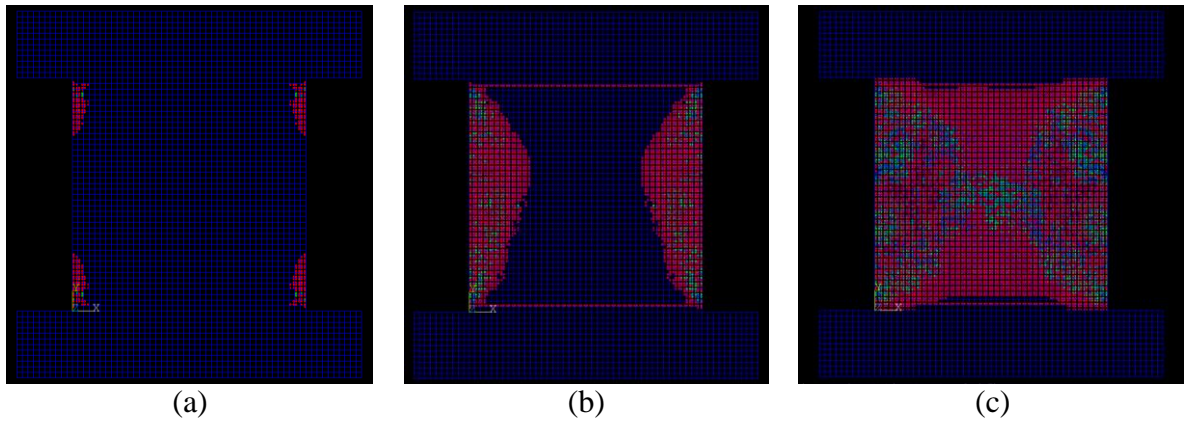


Figure 12. Crack and crushing development for a simple compression test using CONCR; (a) at substep 40, (b) at substep 50, (c) at substep 100.

Regarding the failure pattern of the CONCR model, it can be seen from Figure 14 that cracks initially appear at the center of the wallette (a), then propagate diagonally in both directions (b) until they reach the faces of the wallette (c) forming the typical x pattern expected from a compression test.

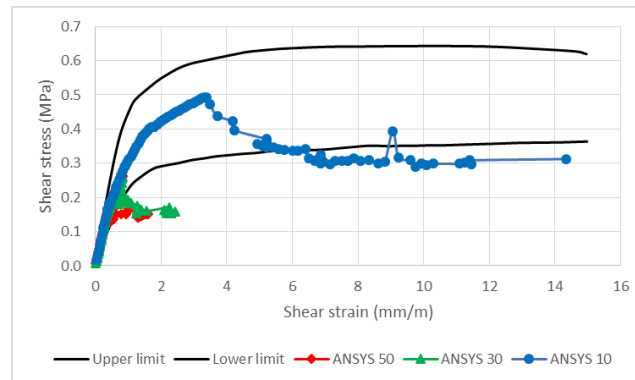


Figure 13. 3D diagonal compression stress-strain curves using CONCR.

Finally, the stress-strain curves and the maximum principal strain plots for the DMGE/DMGI models are presented in Figure 15 and Figure 16 respectively. Even though input parameters were calibrated to provide an accurate strength value for the simple compression test (see Figure 15 (a) and the mesh sensitivity values for the discrepancy of the peak strengths presented in Table 4), neither the pre-peak nor the post peak behavior of cob was captured properly. DMGE/DMGI material shows a stiffer pre-peak behavior and, after reaching the peak strength, a sudden loss of strength. Thus, depicting a fully brittle material behavior rather than the progressive loss of capacity and ductile post peak behavior of cob. Furthermore, the shear responses from the simulations were far from replicating the cob performance reported by Miccoli et al. (see Figure 15 (b)).

As can be seen in Figure 16, neither the simple compression nor the diagonal compression maximum principal strains correspond with the expected x pattern considered as satisfactory. For the simple compression plot, strains accumulate at the interface between the steel plates and the top and bottom of the cob wallette. Whereas that for the diagonal compression plot,

strains accumulate at the singularity points between the sharpened edges of the steel plates in contact with the faces of the cob wallettes.

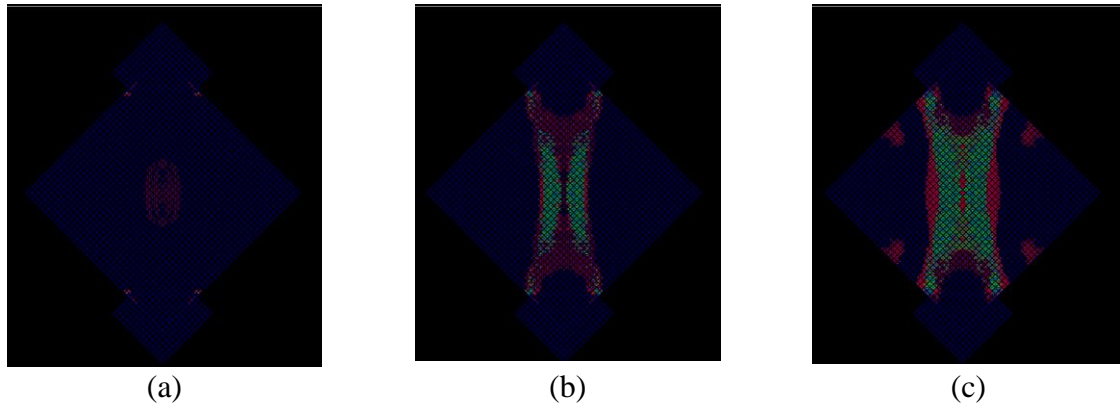


Figure 14. Crack and crushing development for a diagonal compression test using CONCR; (a) at sub step 50, (b) at sub step 70, (c) at sub step 100.

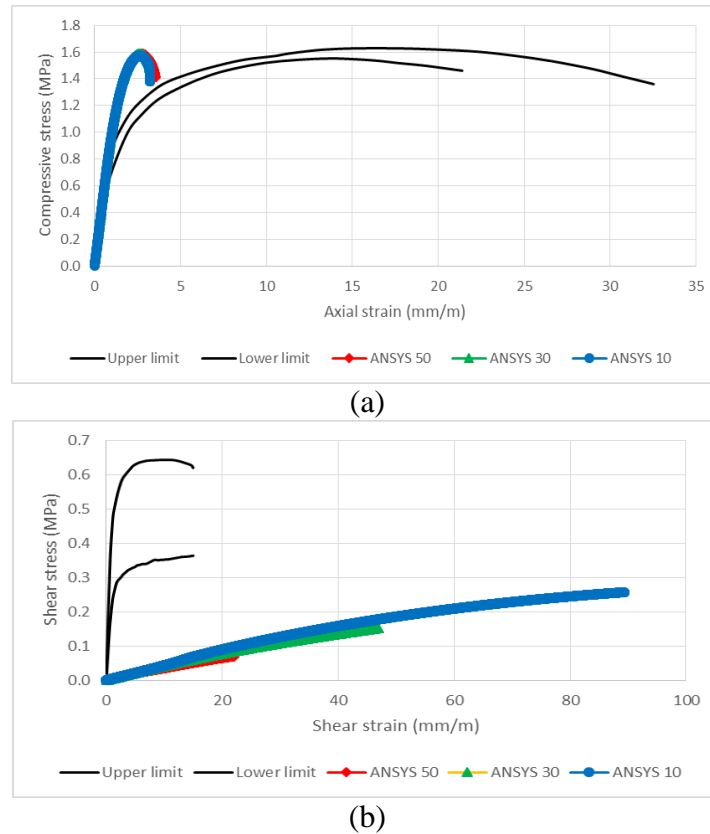


Figure 15. Plane stress stress-strain curves using DMGE/DMGI for (a) simple compression & (b) diagonal compression simulations.

As said before, the input parameters of the DMGE/DMGI models were calibrated to obtain the same peak strength as the reference for the simple compression test. Therefore, the discrepancy values are quite small (within 1 %). On the other hand, the discrepancy values for the peak strengths of the diagonal test are the higher ones among all set of simulations (going from 48 up until the 86 %). As neither pre-peak/post-peak behavior nor the distribution of the maximum principal strains correspond to the results obtained from the experimental campaign, the

DMGE/DMGI material models are the less suitable to reproduce cob's structural response among the set of three different material models implemented.

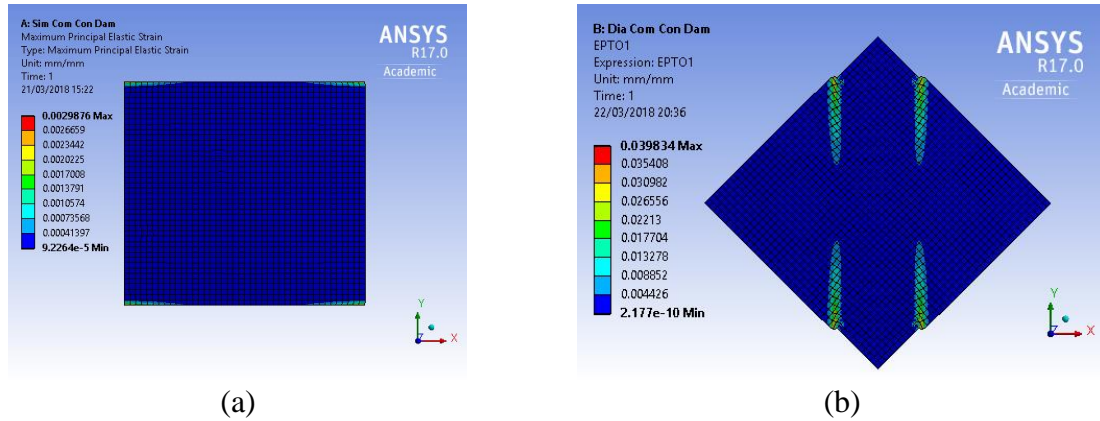


Figure 16. Maximum principal strains using DMGE/DMGI for (a) simple compression model & (b) diagonal compression model.

Table 4. Mesh sensitivity in terms of peak strengths.

Test	Model	Mesh size (mm)	Peak strength (MPa)	Reference value (MPa)	Discrepancy (%)
Simple compression	2D MISO	50	1.704	1.59	7.16
		30	1.722	1.59	8.28
		10	1.736	1.59	9.15
	3D MISO	50	1.965	1.59	23.60
		30	2.201	1.59	38.44
		10	2.196	1.59	38.08
	CONCR	50	1.680	1.59	5.67
		30	1.752	1.59	10.21
		10	1.714	1.59	7.81
	DMGE/DMGI	50	1.587	1.59	0.18
		30	1.587	1.59	0.21
		10	1.572	1.59	1.11
Diagonal compression	2D MISO	50	0.569	0.50	13.75
		30	0.566	0.50	13.27
		10	0.611	0.50	22.20
	3D MISO	50	0.706	0.50	41.26
		30	0.696	0.50	39.11
		10	0.883	0.50	76.55
	CONCR	50	0.263	0.50	47.36
		30	0.262	0.50	47.57
		10	0.492	0.50	1.70
	DMGE/DMGI	50	0.070	0.50	86.07
		30	0.152	0.50	69.64
		10	0.257	0.50	48.69

Conclusions

The suitability of three constitutive material models (MISO, CONCR & DMGE/DMGI) available in ANSYS to replicate the non-linear response of cob wallettes was assessed. The simulations were based on results obtained from previous experimental campaigns.

MISO could reproduce the pre-peak behavior of cob and principal stresses could be used as an indication of the opening of cracks. On the other hand, it did not capture the softening post-peak behavior of the material. CONCR provided quite accurate pre-peak behavior results and peak strength values. Moreover, it was possible to plot the opening of cracks and those plots agreed with the experimental results. Nevertheless, as CONCR is suitable to reproduce brittle failures, it did not capture the long deformations characteristic of cob. Finally, the DMGE/DMGI proved to be inaccurate to reproduce both pre-peak and post-peak behavior. Although parameters could be calibrated to obtain the appropriate peak strength, neither stresses nor do strains correspond to what was observed in the experimental campaign.

The model that better represented the behavior of cob was the 3D CONCR with a 10 mm mesh size. The discrepancy values for both simple and diagonal compression peak strengths are smaller than 10 %. Moreover, the cracks obtained in the model are in good agreement with the typical failure pattern presented in these types of experimental tests.

References

- [1] Miccoli, L., U. Müller, and P. Fontana, *Mechanical behaviour of earthen materials: a comparison between earth block masonry, rammed earth and cob*. Construction and building materials, 2014. **6**: p. 327 - 339.
- [2] Terra Incognita, *Terra Europae, Earthen Architecture in the European Union* ed. E. ETS. 2011.
- [3] Hamard, E., et al., *Cob, a vernacular earth construction process in the context of modern sustainable building*. Building and environment, 2016. **106**: p. 103-119.
- [4] Piesik, S., *Habitat: Vernacular Architecture for a Changing Planet*. 2017, UK: Thames & Hudson Ltd.
- [5] Keefe, L., *Earth building : methods and materials, repair and conservation*. 2005, London: Taylor & Francis. xii, 196 p.
- [6] Cid, J., F.R. Mazarron, and I. Canas, *The earth building normative documents in the world*. Informes de la CConstruccion, 2011. **63**(523): p. 159-169.
- [7] ANSYS®, *ANSYS® Academic Research Mechanical*. 2017.
- [8] Miccoli, L., et al., *In-plane behaviour of earthen materials: a numerical comparison between adobe masonry, rammed earth and cob*, in *6th ECCOMAS Thematic Conference on Computational Methods in Structural Dynamics and Earthquake Engineering* M. Papadrakakis and M. Fragiadakis, Editors. 2017: Rhodes Island, Greece.
- [9] DIANA, *TNO DIANA*. 2014, TNO DIANA BV: Delft, The Netherlands.
- [10] ANSYS Inc., *ANSYS Mechanical APDL Material Reference*. 2013.
- [11] ANSYS Inc., *ANSYS Mechanical APDL Theory Reference*. 2013.
- [12] ANSYS Inc., *ANSYS Mechanical APDL Element Reference*. 2013.
- [13] William, K. and E. Warnke, *Constitutive model for the triaxial behaviour of concrete*, in *Seminar on concrete structures subjected to triaxial stresses*. 1974: Bergamo, Italy.
- [14] Barbero, E.J. and M. SHahbazi, *Determination of material properties for ANSYS progressive damage analysis of laminated composites*. Composite structures, 2017. **176**: p. 768-779.
- [15] British Standard, *BS EN 12390-3:2009 Testing of hardening concrete, Part 3: Compressive strength of test specimens*. 2009.

Parametric Study on RPV Integrity Assessment under Pressurized Thermal Shock

***Jongwook Kim¹, Hanbum Surh¹, and Shinbeom Choi¹**

¹Korea Atomic Energy Research Institute, Republic of Korea

*Presenting & Corresponding author: kjwook1231@gmail.com

Abstract

The objective of this study is to evaluate the structural integrity of a reactor pressure vessel during a pressurized thermal shock event by applying deterministic fracture mechanics. The deterministic fracture mechanics analysis was performed using three-dimensional finite element models. The impacts of the input parameters, such as the crack location, the aspect ratio, and the cladding properties were reviewed through the corresponding sensitivity analyses. In addition, the maximum allowable transition temperatures were estimated.

Keywords: Reactor Pressure vessel (RPV), Pressurized Thermal Shock, Deterministic Fracture Mechanics

Introduction

The reactor pressure vessel integrity is ensured by a proper margin between its loads bearing capacity given by the vessel design and material properties, and the acting loads, which could occur during a plant operation. Thus, it is designed and manufactured according to the strict code requirements to ensure its structural integrity.

Before the late 1970s, it was postulated that the most severe thermal shock that a pressurized water reactor (PWR) vessel must withstand is a large break loss of coolant accident (LOCA). In this type of overcooling transient, low-temperature emergency core coolant will rapidly enter the reactor pressure vessel (RPV) and cool the vessel wall. The resulting temperature gradient in the vessel wall will cause a significant thermal stress, within the inner surface of the wall. However, the stresses due to the system pressure along with the thermal stresses were not considered, since it was expected that during a large break LOCA, the system will depressurize fast and remain at a low pressure. In 1978, the occurrence of a pressurized thermal shock (PTS) at the Rancho Seco nuclear power plant in California showed that some overcooling transients can be accompanied by a re-pressurization of the primary system, which will compound the effects of the thermal stresses. When a system pressure remains high or slowly decreases during thermal shock events, an additional stress from the system pressure greatly increases the possibility of a crack initiation and propagation. In particular, the surface cracks and underclad cracks located in the RPV inner wall can pose concerns under a PTS event. To assure the integrity of RPVs under a PTS event, the PTS rule requires that the RT_{NDT} of RPV beltline materials should be lower than the PTS screening criteria. However, a PTS analysis for the integrity of a RPV is a complex task, which places significant requirements on the experts performing it. These requirements include knowledge of the dominant physical phenomena and associated computer codes, knowledge of the plant being analyzed, and knowledge of the relevant codes and standards for a RPV integrity assessment. In addition, several different procedures and approaches are presently used for the integrity assessments of RPVs. This implies that the results from the assessments are not comparable between individual RPVs as different procedures and approaches are used.

Therefore, it is necessary to undertake a parameter study for the fracture mechanics evaluation of a RPV under a PTS event.

The objective of this study is to evaluate the structural integrity of a RPV under PTS conditions by applying deterministic fracture mechanics. The deterministic fracture mechanics analysis was performed using three-dimensional finite element models. The crack configurations, crack aspect ratio, and cladding properties were considered in the parametric study. In addition, the maximum allowable transition temperatures were investigated.

Problem Definition

Geometric

The RPV considered in the analysis is a typical 3-loop PWR, which is made of ASTM A508 Class 3 with an inner radius of 1994 mm, a base metal thickness of 200 mm, and a cladding thickness of 7.5 mm. The postulated defect as a base case (Case 1) is a through-clad surface-breaking semi-elliptical crack of 19.5 mm in depth by 117 mm in length for $a/c = 1/3$, as shown in Fig. 1. The configuration of an elliptical underclad crack (Case 6) is shown in Fig. 2.

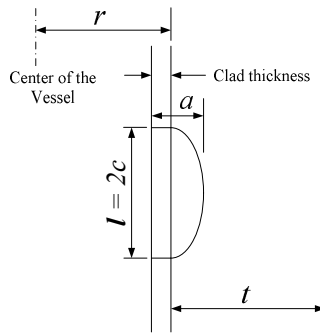


Figure 1. Schematic Illustration of a postulated crack (Case 1)

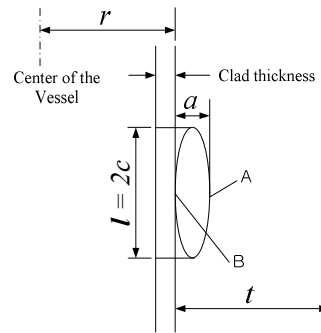


Figure 2. Schematic Illustration of a postulated crack (Case 6)

Transient Conditions

One overcooling transient due to an assumed leak is defined as in Fig. 3, for which axisymmetric loading conditions are assumed. The figure shows a typical PTS transient with repressurization. The temperature and pressure start to decrease, but at a certain time, namely, about 7200 seconds after the transient begins, the system pressure increases rapidly. It is maintained and a slow heating occurs, which shows the typical characteristics of a PTS transient. In this case, the pressure is assumed to be a dominant factor.

Sensitivity Study

Several parametric studies are proposed to investigate the influence of certain parameters on the results. Of them, considered here is a postulated defect of those parameters such as underclad vs. surface cracking, defect aspect ratio ($a/c=1/3, 1/2, 1/1$), and elliptical vs. semielliptical cracking. An analysis matrix for the sensitivity of the postulated defect is shown in Table 1. In addition, the effects of the cladding are investigated for three conditions of Case 1 as follows:

- C1: No cladding. Cladding properties are assumed as identical to the base metal.
- C2: Cladding thermal conductivity is considered. Additional stress due to a steep temperature gradient in the cladding is evaluated.
- C3: Cladding is fully considered. Additional stresses due to a steep temperature gradient and a differential thermal expansion are evaluated.

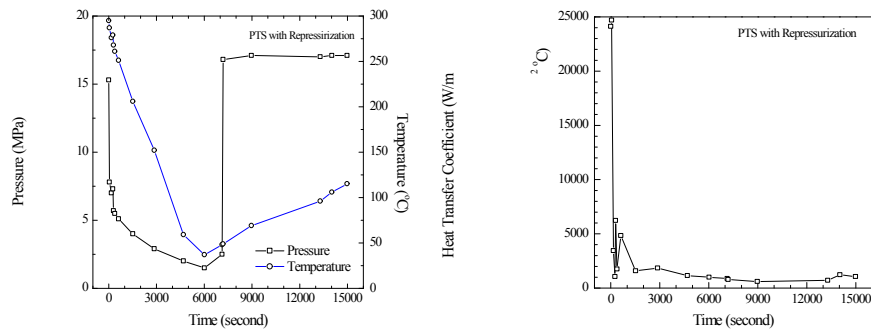


Figure 3. Transient histories of a PTS with repressurization

Table 1. Analysis matrix for a sensitivity study of a postulated defect

Case	Location	Shape	Aspect ratio (a/c)	Depth	
				a	a/t
1	surface	semi-elliptical	1/3	19.5	0.06
2	underclad	semi-elliptical	1/3	12.0	0.06
3	underclad	semi-elliptical	1/3	15.0	0.075
4	surface	semi-elliptical	1/2	19.5	0.06
5	surface	semi-elliptical	1/1	19.5	0.06
6	underclad	elliptical	1/3	12.0	0.06

Finite Element Modeling

In this paper, three-dimensional finite element analyses were performed for an assessment of various cracks in the RPV under PTS conditions. The model was designed using 20-node isoparametric quadratic brick elements with reduced Gaussian integration points and 20-node quarter point brick elements for the crack front point. Typical examples of the meshes are shown in Figs. 4 (semi-elliptical surface crack) and 5 (elliptical underclad crack), respectively. Three-dimensional finite element analyses, including thermal and mechanical calculations, were performed using the ABAQUS finite element analysis program. For each geometrical configuration, the stress intensity factor was calculated from the value of the J-integral obtained at the deepest point.

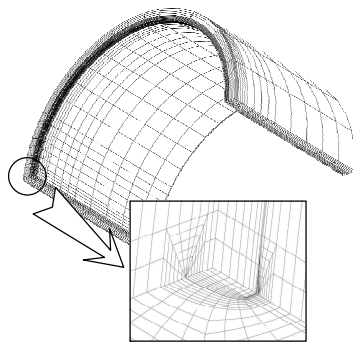


Figure 4. 3-Dimensional FE mesh for the semi-elliptical surface crack

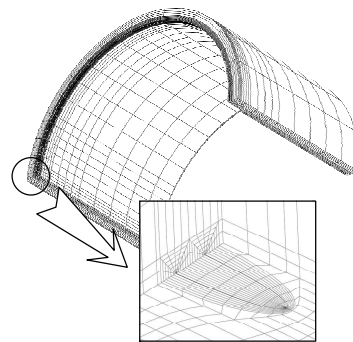


Figure 5. 3-Dimensional FE mesh for the elliptical underclad crack

Finite Element Analysis Results

Maximum Allowable RT_{NDT}

The maximum allowable nil ductile transition temperature for each crack geometric under a PTS event is presented in Table 2.

Effects of Crack Aspect Ratio

The effects of the crack aspect ratio on the stress intensity factor curves are shown in Fig. 6. The stress intensity factor curves decrease with an increase in the crack aspect ratio.

Influence of Crack Configuration

The effects of the crack location are shown in Fig. 7 for surface and underclad cracking. The stress intensity factor for a surface crack is much higher than that of an underclad crack with the same deepest point. In Case 2_C3, the stress intensity factor is so small that crack initiation does not occur during the transient event. Fig. 8 shows the stress intensity factor curves for the existing cracks in the base metal. The two results show a similar tendency. The stress intensity factor curves at two crack tip points are shown in Fig. 9 for the elliptical underclad crack. The stress intensity factor curve at point B is much higher than that of point A at the deepest point of a crack front. This means that it is possible to predict a crack initiation at point B.

Influence of Cladding Properties

Fig. 10 shows the stress intensity factor curves with different properties of the cladding. The curve of Case 1_C3 is much higher than that of Cases 1_C2 and C3. Accordingly, the maximum allowable RT_{NDT} of Case 1_C3 is the lowest.

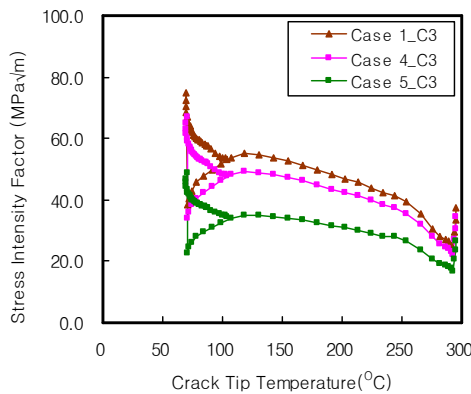


Figure 6. Comparison of the SIF with various crack aspect ratios

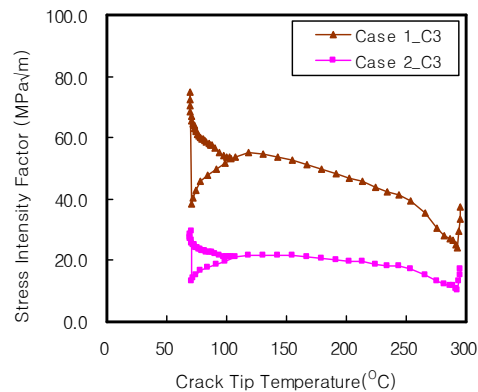


Figure 7. Comparison of the SIF with various crack locations

Conclusions

In this paper, three-dimensional finite element analyses were performed to evaluate the integrity of a RPV under PTS conditions, and the following conclusions were obtained.

1. Stress discontinuity takes place at a boundary line between the cladding and base metal.
2. As the crack aspect ratio increases with the same crack depth, the maximum allowable nil ductile transition temperature increases.

3. When the differences in the thermal conductivity and the thermal expansion coefficients of a cladding are fully considered, the stress intensity factor curve increases.
4. The stress intensity factor of a surface crack is much higher than that of an underclad crack with the same deepest point.
5. For an elliptical underclad crack, the possibility of a crack initiation was monitored at the boundary between the cladding and base metal.

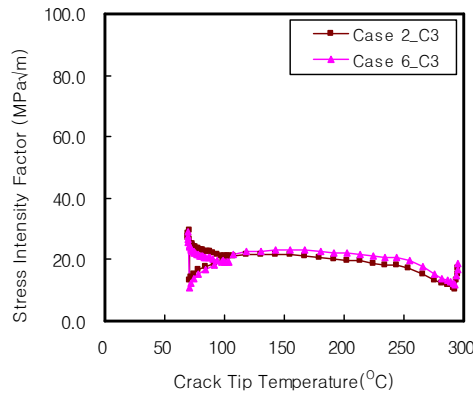


Figure 8. Comparison of the SIF with various crack shapes

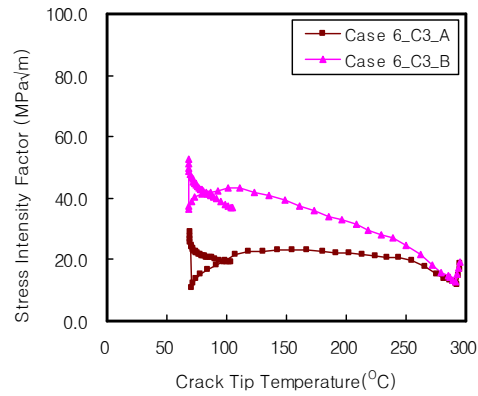


Figure 9. Comparison of the SIF with crack tip points

Table 2. Maximum Allowable RT_{NDT} by the maximum criteria at the deepest point

Case	Cladding	Maximum allowable RT_{NDT} (°C)
1	C1	69
	C2	71
	C3	56
2	C3	No intersection
3	C3	62
4	C3	No intersection
5	C3	89
6	C3 (Point A)	No intersection
	C3 (Point B)	78

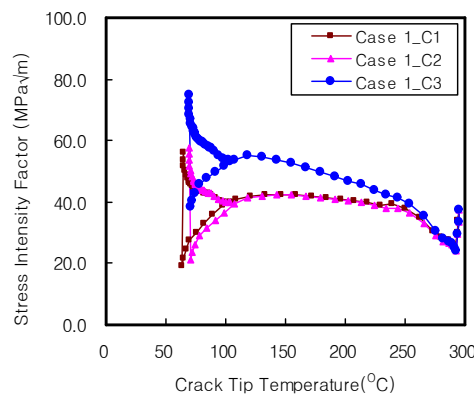


Figure 10. Comparison of the SIF with various thermal/mechanical properties of the cladding

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (2016M2C6A1930040).

References

- [1] U.S. Nuclear Regulatory Commission (1997): “Fracture toughness requirements for protection against pressurized thermal shock events”, 10 CFR 50.61.
- [2] The American Society of Mechanical Engineers (2007): “Protection against nonductile failure”, ASME Boiler and Pressure Vessel Code Section III, Appendix G.
- [3] The American Society of Mechanical Engineers (2007): “Rules for inservice inspection of nuclear power plant components”, ASME Boiler and Pressure Vessel Code Section XI, Appendix A.

Spectral quasi-linearization method for entropy generation using the Cattaneo-Christov heat flux model

Hiranmoy Mondal* and Precious Sibanda

School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal,
Private Bag X01, Scottsville, Pietermaritzburg-3209, South Africa

*Presenting and Corresponding author: hiranmoymondal@yahoo.co.in

Abstract

In this paper, we study entropy generation in Sikiadas nanofluid flow along a moving plate subject to an inclined magnetic field and a Cattaneo-Christov heat flux model that may predict the effects of thermal relaxation time in the boundary layer flow. The nonlinear transport equations are solved using a spectral-quasi linearization method. An analysis of the convergence of the method is presented, and the importance of various fluid and physical parameters on the behavior of the solutions is explored. It is shown that the method converges fast and gives accurate results. The results show that entropy generation increases with an increase in the Reynolds number.

Keywords: Cattaneo-Christov heat flux model; inclined magnetic field; Nanofluid; SQLM.

Introduction:

Heat transfer characteristics in a fluid have traditionally been studied using Fourier's law of heat conduction. It is important in many industrial and engineering processes including nuclear reactor cooling, space cooling, energy production, biomedical applications such as magnetic drug targeting, heat conduction in tissues etc. and many others. The temperature difference between two unlike bodies causes heat transfer mechanism. The heat transfer phenomenon was primarily described by Fourier [1] which is parabolic energy equation for temperature field. One of the major shortcomings of this model is that it produces a parabolic energy equation which means that an initial disturbance would instantly affect the system under consideration. After that Cattaneo [2] modifies the Fourier law of heat conduction in which he added the thermal relaxation term. The addition of thermal relaxation time causes heat transportation in the form of thermal waves with finite speed. A material invariant formulation of the Cattaneo's model was presented by Christov [3] through the consideration of Oldroyd's upper-convected derivative. The Cattaneo-Christov equations uniqueness and structural stability were discussed by Ciarletta and Straughan [4]. Mushtaq et.al [5] were studied Sakiadis flow of UCM fluid by considering Cattaneo-Christov heat flux model and concluded that temperature distribution is Non-Monotonic with an increasing thermal relaxation time. Han [6] explored the heat transfer phenomenon of viscoelastic fluid under the Cattaneo-Christov theory and Salahuddin et.al [7] analysed the MHD flow of Williamson fluid with variable thickness by considering Cattaneo-Christov heat flux model.

The addition of nano-sized metallic or metal oxide particles to base fluids such as oil or water leads to nanofluids. These fluids have enhanced thermo physical properties such as a higher thermal conductivity, viscosity, thermal diffusivity and convective heat transfer coefficients compared to base fluids. Dogonchi and Ganji [8] investigated the importance of the Cattaneo–Christov heat flux model for heat transfer in an MHD nanofluid flow between parallel plates. Sithole et al. [9] investigated entropy generation in a second grade fluid flow over a stretching sheet.

The innovation of our study in heat transfer fluids involves the addition of nano-sized metallic or metal oxide particles to base fluids such as oil or water. The resulting nanofluids have been found to possess enhanced thermo physical properties such as the thermal conductivity, viscosity, thermal diffusivity and convective heat transfer coefficients compared to the of base fluid. These novel properties ensure that nanofluids have great potential for useful as heat transfer fluids, in, for instance, microelectronic devices, fuel cells, engine cooling/vehicle thermal management, heat exchangers and in boiler flue gas temperature reduction. The boundary layer flow of a third-grade viscoelastic power-law non-Newtonian fluid over a porous wedge was investigated by Rashidi et al. [10]. He studied the impact of buoyancy and thermal radiation on magneto hydrodynamic nanofluid flow past a stretching sheet. Khan and Pop [11] obtained a numerical solution for the two-dimensional flow of nanofluid over a linearly stretching sheet. Buongiorno [12] presented a nonhomogeneous equilibrium mathematical model for convective transport in nanofluids. Kuznetsov and Nield [13] studied the nanofluid boundary layer flow past a vertical plate. De et al. [14] were investigate the flow of nanofluids. She concluded that Brownian motion and thermophoretic diffusion of nanoparticles are the most important mechanisms for the abnormal convective heat transfer enhancement.

The aim is to explore entropy generation and heat transfer characteristics in an MHD fluid flow for the well-known Sakiadis problem and a Maxwell nanofluid using the Cattaneo–Christov heat flux model. The conservation equations are solved numerically using the spectral quasilinearization method. The significance of physical and fluid parameters on the flow and entropy generation fields is discussed in detail.

Mathematical formulation:

The plate is assumed to be at a constant temperature T_w and T_∞ denotes the ambient fluid temperature. Christov's heat flux model is used. The uniform magnetic field is applied at an angle Λ to the positive direction of the y -axis. Making use of the standard boundary layer approximations, the equations governing the steady incompressible flow of nanofluid and heat transfer can be expressed as below:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \quad (1)$$

$$u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + \lambda_1 \left(\frac{u^2 \partial^2 u}{\partial x^2} + v^2 \frac{\partial^2 u}{\partial y^2} + 2uv \frac{\partial^2 u}{\partial x \partial y} \right) = \nu \frac{\partial^2 u}{\partial y^2} - \frac{\sigma}{\rho} B_0^2 \sin^2(\Lambda) u \quad (2)$$

$$\rho c_p \left(u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} \right) = -\nabla \cdot \mathbf{q} + (\rho c_p)_f \left[D_B \frac{\partial C}{\partial y} \frac{\partial T}{\partial y} + \frac{D_T}{T_\infty} \left(\frac{\partial T}{\partial y} \right)^2 \right] \quad (3)$$

$$u \frac{\partial C}{\partial x} + v \frac{\partial C}{\partial y} = D_B \frac{\partial^2 C}{\partial y^2} + \frac{D_T}{T_\infty} \frac{\partial^2 T}{\partial y^2} \quad (4)$$

In the energy equation, \mathbf{q} is the heat flux that satisfies the following equation

$$\mathbf{q} + \lambda_2 \left(\frac{\partial \mathbf{q}}{\partial t} + V \cdot \nabla \mathbf{q} - \mathbf{q} \cdot \nabla V + (\nabla \cdot V) \mathbf{q} \right) = -k \nabla T \quad (5)$$

In which $V = (u, v, 0)$ 2-Dimensional velocity vector, λ_2 is the relaxation time for heat flux and k is the thermal conductivity of the fluid. We considered flow is incompressible so that equation (5) takes the form as

$$\mathbf{q} + \lambda_2 \left(\frac{\partial \mathbf{q}}{\partial t} + V \cdot \nabla \mathbf{q} - \mathbf{q} \cdot \nabla V \right) = -k \nabla T \quad (6)$$

Eliminating \mathbf{q} from equations (3) and (5), we obtain the following equation (for more details see Christov [25] and Han et al [28])

$$u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + \lambda_2 \left(\left(u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} \right) \frac{\partial T}{\partial x} + \left(u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} \right) \frac{\partial T}{\partial y} + u^2 \frac{\partial^2 T}{\partial x^2} + v^2 \frac{\partial^2 T}{\partial y^2} + 2uv \frac{\partial^2 T}{\partial x \partial y} \right) = \alpha \frac{\partial^2 T}{\partial y^2} + \tau \left[D_B \frac{\partial C}{\partial y} \frac{\partial T}{\partial y} + \frac{D_T}{T_\infty} \left(\frac{\partial T}{\partial y} \right)^2 \right] \quad (7)$$

In addition, the boundary conditions are

$$u = U, \quad v = 0, \quad T = T_w, \quad D_B \frac{\partial C}{\partial y} + \frac{D_T}{T_\infty} \frac{\partial T}{\partial y} = 0 \quad \text{at } y = 0 \quad (8)$$

$$u \rightarrow 0, \quad T \rightarrow T_\infty, \quad C \rightarrow C_\infty \quad \text{as } y \rightarrow \infty \quad (9)$$

where x, y are the coordinates along plate and normal to the plate respectively, u, v are velocity components of along x and y axis respectively, λ_1 is the relaxation time of the fluid, c_p is the specific heat, ρ is the density of the fluid, ν is the kinematic viscosity, D_B and D_T is the Brownian diffusion coefficient and thermophoretic diffusion coefficient respectively. $\tau = \frac{(\rho c)_p}{(\rho c)_f}$ is the ratio between the effective heat capacity of the nanoparticles material and heat capacity of the fluid, T and C are fluid temperature and nanoparticles fraction, respectively, T_w and T_∞ are the temperature of the fluid at the wall and ambient temperature.

Introducing the following similarity transformations

$$\eta = y \sqrt{\frac{U}{\nu x}}, \quad u = U f'(\eta), \quad v = -\sqrt{\frac{U \nu}{x}} (f - \eta f'), \quad \theta(\eta) = \frac{T - T_\infty}{T_w - T_\infty}, \quad \phi(\eta) = \frac{C - C_\infty}{C_\infty} \quad (10)$$

By using similarity transformations equations (1), (2), (4) and (6) reduces to

$$f''' + \frac{1}{2} f f'' - \frac{\beta_1}{2} (2 f f f'' + f^2 f''' + \eta f'^2 f'') - M \sin^2(\alpha) f' = 0 \quad (11)$$

$$\frac{1}{Pr} \theta'' + \frac{1}{2} f \theta' - \frac{\beta_2}{2} (3 f f' \theta' + f^2 \theta'') + Nb \theta' \phi' + Nt \theta'^2 = 0 \quad (12)$$

$$\phi'' + \frac{1}{2} Le Pr f \phi' + \frac{Nt}{Nb} \theta'' = 0 \quad (13)$$

and the boundary conditions are

$$f' = 1, f = 0, \theta = 1, Nb\phi' + Nt\theta' = 0 \text{ at } \eta = 0 \quad (14)$$

$$f' \rightarrow 0, \theta \rightarrow 0, \phi \rightarrow 0 \text{ as } \eta \rightarrow \infty. \quad (15)$$

Here $M = \frac{\sigma B_0^2}{\rho U}$ is the magnetic parameter, \wedge is the inclined angle, $\beta_1 = \frac{U \lambda_1}{2x}$ is the viscoelastic fluid parameter, $\beta_2 = \frac{U \lambda_2}{2x}$ is the dimensionless thermal relaxation time, $Pr = \frac{\nu}{\alpha}$ is the Prandtl number, $Nb = \frac{\tau D_B C_\infty}{\nu}$ is the Brownian motion parameter, $Nt = \frac{\tau D_T (T_w - T_\infty)}{T_\infty \nu}$ thermophoresis parameter and $Le = \frac{\nu}{D_B}$ is the Lewis number.

Entropy Generation:

The volumetric entropy generation in the Cattaneo-Christov nanofluid, which is based on the second law of thermodynamics, is given by

$$S_{gen}''' = \frac{k_f}{T_\infty^2} \left(\frac{\partial T}{\partial y} \right)^2 + \frac{\sigma B_0^2}{T_\infty} u^2 + \frac{R D_B}{T_\infty} \frac{\partial T}{\partial y} \frac{\partial C}{\partial y} + \frac{R D_B}{C_\infty} \left(\frac{\partial C}{\partial y} \right)^2 \quad (16)$$

It is suitable to define the entropy generation number N_G as a ratio between the local volumetric entropy generation rate S_{gen}''' and a characteristic rate of entropy generation which is defined by S_0'''

$$S_0''' = \frac{k_f (T_w - T_\infty)}{T_\infty^2 x^2} \quad (17)$$

The entropy generation number N_G can be obtained as

$$N_G = \frac{S_{gen}'''}{S_0'''} = Re \theta'^2 + \frac{Re Br M}{\chi} f'^2 + \frac{Re \Sigma}{\chi} \theta' \phi' + \frac{Re \Sigma}{\chi^2} \phi'^2 \quad (18)$$

Where Re is the Reynolds number, Br is the Brinkman number, Σ is the constant and χ is the temperature difference parameter.

Equation (18) can be obtained as a summation of the entropy generation number caused by the heat transfer ($N1$) and the entropy generation number caused by both diffusive irreversibility and magnetic field ($N2$). That is $N_G = N1 + N2$, where

$$N1 = Re \theta'^2, N2 = \frac{Re Br M}{\chi} f'^2 + \frac{Re \Sigma}{\chi} \theta' \phi' + \frac{Re \Sigma}{\chi^2} \phi'^2 \quad (19)$$

The heat transfer irreversibility, diffusive irreversibility and the magnetic field all contribute to entropy generation. It is therefore worthwhile investigating the conditions under which heat

transfer dominates entropy generation. To investigate this question, the Bejan (Be) number is defined as the ratio of entropy generation due to heat transfer and the entropy generation number.

$$Be = \frac{N1}{N_G} \quad (20)$$

The Bejan number takes values in range $[0, 1]$. At the extreme when $Be = 1$ the irreversibility of heat transfer dominates. On the other extreme when $Be = 0$ the combined effects of diffusion and magnetic field dominates the irreversibility. When $Be = 0.5$, the contribution of heat transfer in entropy generation is the same as the combined contribution of diffusion and magnetic field in entropy generation. Additionally, the Bejan number Be is considered at the best values of the parameters at which the entropy generation its minimum.

Method of solution:

The nonlinear-coupled ordinary differential equations (11)-(13) subject to the boundary conditions (14) have been solved numerically using the spectral quasi linearization method (SQLM). The quasi linearization method is employed to linearize the equations before they are solved iteratively using the Chebyshev spectral collocation method. Applying the quasi linearization procedure to equations (11)-(15), the resultant equations are

$$\alpha_{1,r} f_{r+1}''' + \alpha_{2,r} f_{r+1}'' + \alpha_{3,r} f_{r+1}' + \alpha_{4,r} f_{r+1} = R_1 \quad (21)$$

$$\beta_{1,r} f_{r+1} + \beta_{2,r} f_{r+1}' + \beta_{3,r} \theta_{r+1}'' + \beta_{4,r} \theta_{r+1}' + \beta_{5,r} \phi_{r+1}' = R_2 \quad (22)$$

$$\gamma_{1,r} f_{r+1} + \gamma_{2,r} \theta_{r+1}'' + \gamma_{3,r} \phi_{r+1}'' + \gamma_{4,r} \phi_{r+1}' = R_3 \quad (23)$$

and the boundary conditions are

$$\begin{aligned} f_{r+1} = 0, \quad f_{r+1}' = 1, \quad \theta_{r+1} = 1, \quad Nb\phi_{r+1}' + Nt\theta_{r+1}' = 0, \quad \text{at } \eta = 0 \\ f_{r+1}' = 0, \quad \theta_{r+1} = 0, \quad \phi_{r+1} = 0 \quad \text{at } \eta \rightarrow \infty. \end{aligned} \quad (24)$$

Where

$$\alpha_{1,r} = \left(\frac{1}{2} f_r'' - \beta_1 f_r' f_r'' - \beta_1 f_r f_r''' \right), \quad \alpha_{2,r} = -\beta_1 \eta f_r' f_r'' - \beta_1 f_r f_r'' - M \sin^2(\Lambda),$$

$$\alpha_{3,r} = -\beta_1 f_r f_r' - \frac{\beta_1}{2} \eta f_r'^2 + \frac{1}{2} f_r, \quad \alpha_{4,r} = 1 - \frac{\beta_1}{2} f_r^2$$

$$\beta_{1,r} = \frac{1}{2} \theta_r' - \frac{\beta_2}{2} (3f_r' \theta_r' + 2f_r \theta_r''), \quad \beta_{2,r} = -\frac{3\beta_2}{2} f_r \theta_r', \quad \beta_{3,r} = \frac{1}{Pr} - \frac{\beta_2}{2} f_r^2,$$

$$\beta_{4,r} = \frac{1}{2} f_r - \frac{3\beta_2}{2} f_r f_r' + Nb\phi_r' + 2Nt\theta_r', \quad \beta_{5,r} = Nb\theta_r'$$

$$\gamma_{1,r} = \frac{1}{2} Pr Le \phi_r', \quad \gamma_{2,r} = \frac{Nt}{Nb}, \quad \gamma_{3,r} = 1, \quad \gamma_{4,r} = \frac{1}{2} Pr Le f_r'$$

$$R_1 = \frac{1}{2} f_r' f_r'' - \beta_1 (f_r'^2 f_r''' + \eta f_r'^2 f_r'' + 2 f_r' f_r' f_r'')$$

$$R_2 = \frac{1}{2} f_r' \theta_r' + Nt \theta_r'^2 + Nb \theta_r' \phi_r' - \beta_2 (3 f_r' f_r' \theta_r' + f_r'^2 \theta_r'')$$

$$R_3 = \frac{1}{2} Pr Le f_r' \phi_r'$$

The equations (21) to (24) constitutes a linear system of coupled differential equations with variable coefficients and can be solved iteratively using any numerical method for $r = 1, 2, 3, \dots$. In this work, as we discussed below, the Chebyshev spectral collocation method was used to solve the QLM scheme (21) to (24). Before applying the spectral method, it is convenient to transform the domain in the η direction is approximated to $[0, L]$ where L is the edge of the boundary limit (large enough), use the transformation of algebraic mapping $\eta = \frac{(\tau+1)L}{2}$ to map the physical domain into the computational domain $[-1, 1]$. This basic idea of this method is approximating the unknown functions by the Chebyshev interpolating polynomials in such a way that they are collocated at the Gauss-Lobatto points defined as

$$\tau_i = \cos\left(\frac{\pi i}{N}\right), \quad -1 \leq \tau \leq 1, \quad i = 0, 1, 2, \dots, N \quad (25)$$

where N is the number of collocation points. The derivative of f_{r+1} at the collocation points is represented as

$$\frac{\partial^p f_{r+1}}{\partial \eta^p} = \left(\frac{2}{L}\right)^p \sum_{k=0}^N D_{N,k}^p f_{r+1}(\tau_k) = \mathbf{D}^p \mathbf{F} \quad (26)$$

where $\mathbf{D} = \frac{2}{L} D$ and D is the Chebyshev spectral differentiation matrix

$\mathbf{F} = [f(\tau_0), f(\tau_1), \dots, f(\tau_N)]$. Similarly the derivatives of θ , and ϕ given by

$\theta^p = \mathbf{D}^p \Theta$ $\phi^p = \mathbf{D}^p \Phi$. $\theta^p = \mathbf{D}^p \Theta$ $\phi^p = \mathbf{D}^p \Phi$. where p is the order of derivative, and \mathbf{D} is the matrix of order $(N+1) \times (N+1)$. Substituting (24)-(25) into the equations (21)-(23) we obtain

$$\left[\alpha_{1,r} \mathbf{D}^3 + \alpha_{2,r} \mathbf{D}^2 + \alpha_{3,r} \mathbf{D} + \alpha_{4,r} \right] \mathbf{F}_{r+1} = \mathbf{R}_1 \quad (27)$$

$$\left[\beta_{1,r} \mathbf{D} + \beta_{2,r} \mathbf{I} \right] \mathbf{F}_{r+1} + \left[\beta_{3,r} \mathbf{D}^2 + \beta_{4,r} \mathbf{D} \right] \Theta_{r+1} + \left[\beta_{5,r} \mathbf{D} \right] \Phi_{r+1} = \mathbf{R}_2 \quad (28)$$

$$[\gamma_{1,r}\mathbf{I}]\mathbf{F}_{r+1} + [\gamma_{2,r}\mathbf{D}^2]\mathbf{\Theta}_{r+1} + [\gamma_{3,r}\mathbf{D}^2 + \gamma_{4,r}\mathbf{D}]\mathbf{\Phi}_{r+1} = \mathbf{R}_3 \quad (29)$$

Applying spectral method on the boundary conditions gives

$$f_{r+1}(\tau_N) = 0, \quad \sum_{k=0}^N D_{N,k} f_{r+1}(\tau_k) = 1, \quad \theta_{r+1}(\tau_N) = 1, \quad Nb \sum_{k=0}^N D_{N,k} \phi_{r+1}(\tau_k) + Nt \sum_{k=0}^N D_{N,k} \theta_{r+1}(\tau_k) = 0$$

$$\sum_{k=0}^N D_{0,k} f_{r+1}(\tau_k) = 0, \quad \theta_{r+1}(\tau_0) = 0, \quad \phi_{r+1}(\tau_0) = 0 \quad (30)$$

The above system of equations is written in the matrix form as

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} \mathbf{F}_{r+1} \\ \mathbf{\Theta}_{r+1} \\ \mathbf{\Phi}_{r+1} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \mathbf{R}_3 \end{bmatrix} \quad (31)$$

Where

$$A_{11} = \text{diag}[\alpha_{1,r}]\mathbf{D}^3 + \text{diag}[\alpha_{2,r}]\mathbf{D}^2 + \text{diag}[\alpha_{3,r}]\mathbf{D} + \text{diag}[\alpha_{4,r}]\mathbf{I}$$

$$A_{12} = \mathbf{0}, \quad A_{13} = \mathbf{0}$$

$$A_{21} = \text{diag}[\beta_{1,r}]\mathbf{D} + \text{diag}[\beta_{2,r}]\mathbf{I},$$

$$A_{22} = \text{diag}[\beta_{3,r}]\mathbf{D}^2 + \text{diag}[\beta_{4,r}]\mathbf{D},$$

$$A_{23} = \text{diag}[\beta_{5,r}]\mathbf{D}, \quad A_{31} = \text{diag}[\gamma_{1,r}]\mathbf{I},$$

$$A_{32} = \text{diag}[\gamma_{2,r}]\mathbf{D}^2, \quad A_{33} = \text{diag}[\gamma_{3,r}]\mathbf{D}^2 + \text{diag}[\gamma_{4,r}]\mathbf{D}$$

Where α , β and γ are $(N+1) \times (N+1)$ diagonal matrices, \mathbf{I} , $\mathbf{0}$ is a $(N+1) \times (N+1)$ Unit matrix and zero matrix respectively. The approximate solutions for \mathbf{F} , $\mathbf{\Theta}$ and $\mathbf{\Phi}$ are obtained by solving the matrix system (31).

Convergence analysis:

The coupled ordinary differential equation (11)–(13) with the boundary conditions (14)–(15) are solved using the SQLM. To validate the accuracy of the numerical results, the residual errors and the error norms are calculated. The residual error measures the extent to which the numerical solution approximates the genuine solution. The error norm is the difference between the approximate values at successive iterations and may be used to evaluate convergence and stability of the iteration scheme. We studied the change in the residual and error norms with several physical parameters such as the magnetic field, viscoelasticity, the thermal relaxation time, the Brownian motion and the thermophoresis parameter. The results are shown in Figs. 1 to 9.

Figs 1 to 3 show the residual errors in velocity profiles against iterations for different values of the magnetic field parameter, viscoelastic fluid and thermal relaxation time parameter. The residual errors converge after the fifth iteration with a residual error of 10^{-10} . The error norms decrease and smooth converge is achieved.

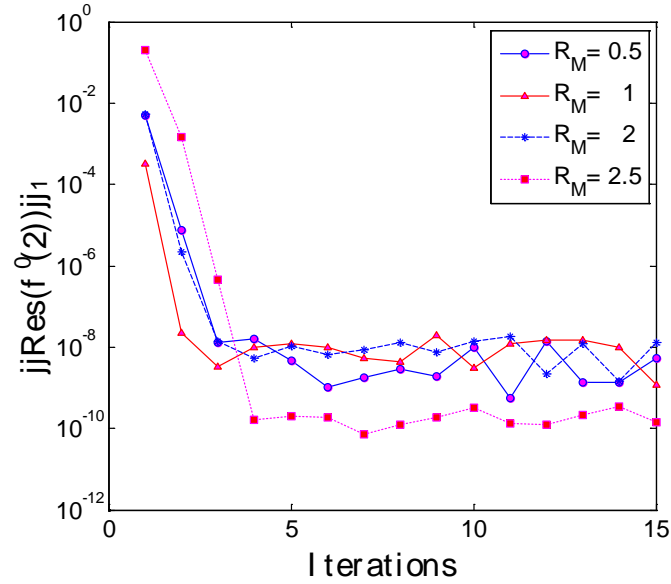


Fig.1: Residual error in the velocity profiles for different values of the magnetic parameter.

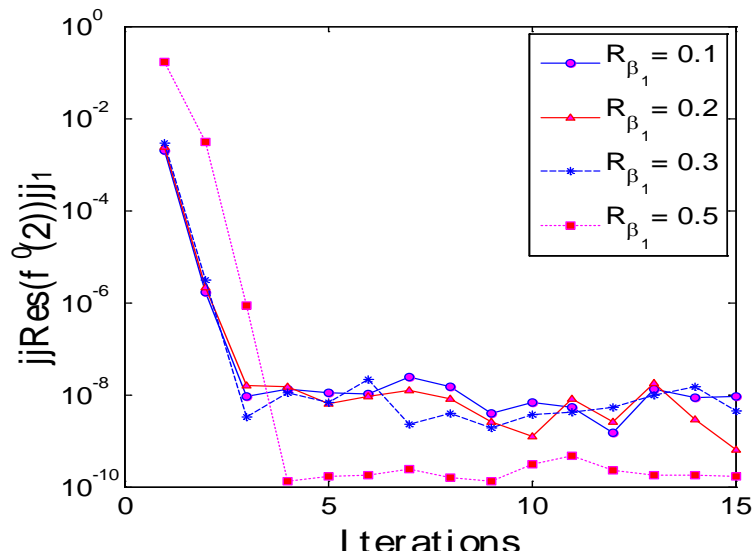


Fig. 2: Residual error in the velocity profiles for different values of β_1 .

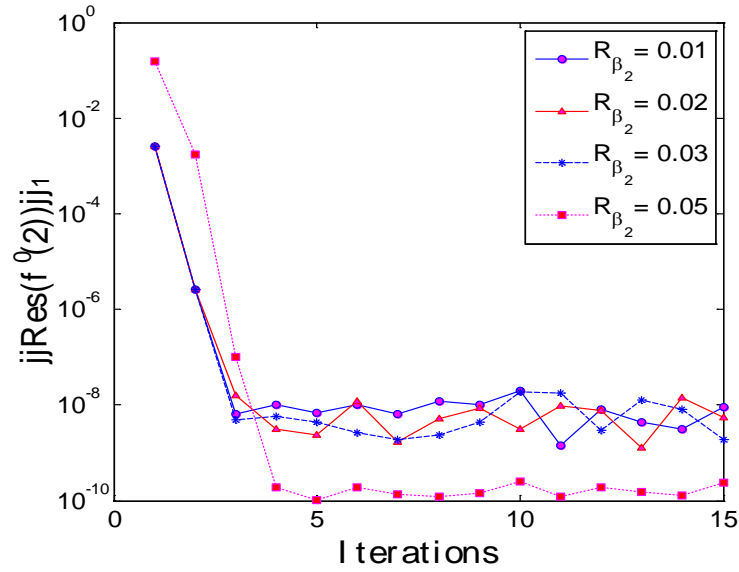


Fig.3: Residual error in the velocity profiles for different values of β_2 .

Figs. 4 –5 show the residual errors and solution error norms against the number of iterations for different values of the thermal relaxation time and thermophoresis parameter. The residual errors converge after the fourth iteration with a residual error 10^{-10} . The solution error norm decreases and convergence is rapid.

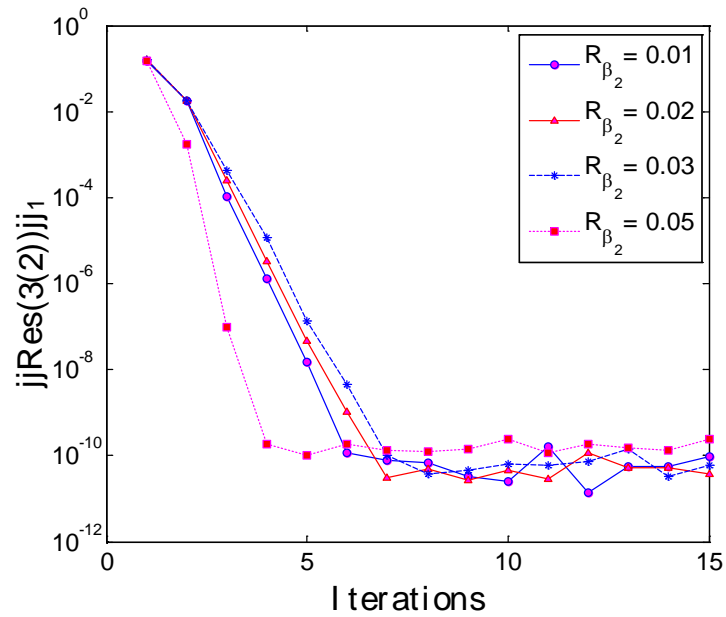


Fig.4: Residual error in the temperature profiles for different values of β_2 .

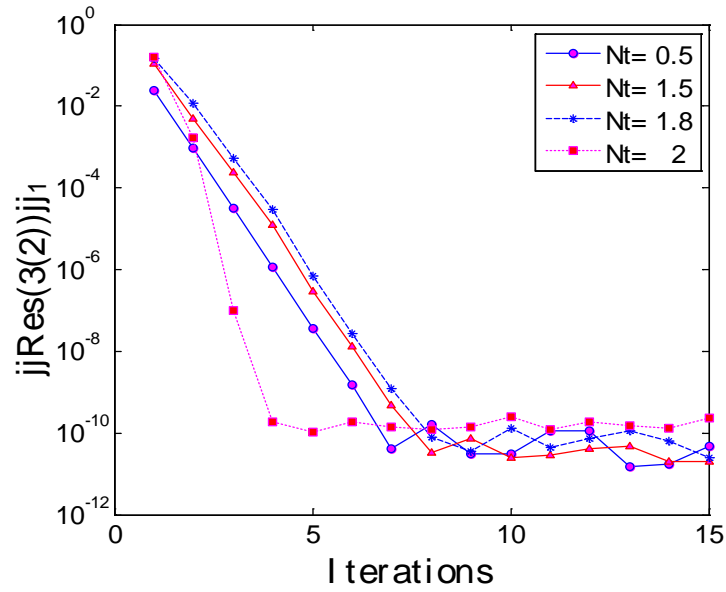


Fig.5: Residual error in the temperature profiles for different values of Nt .

Figs. 6-9 show the residual errors and solution error norms against iterations for different values of the viscoelastic fluid, dimensionless thermal relaxation time, Brownian motion and thermophoresis parameter. The residual errors converge after four iteration with a residual error is 10^{-10} .

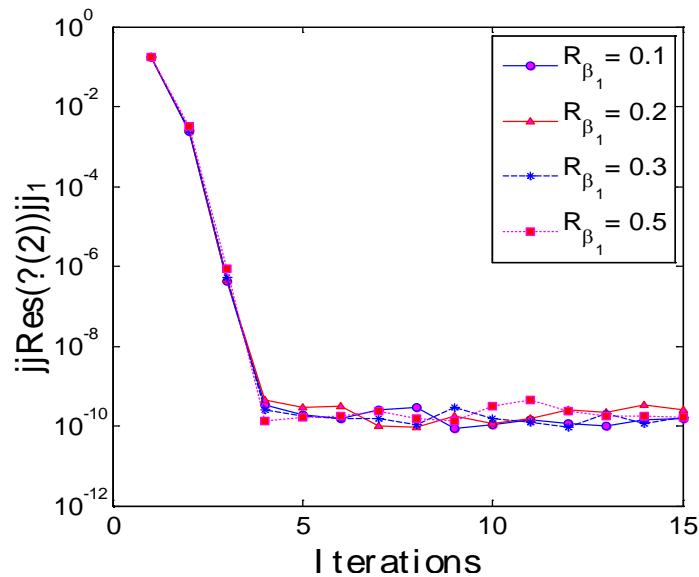


Fig.6: Residual error in concentration profiles for different values of β_1 .

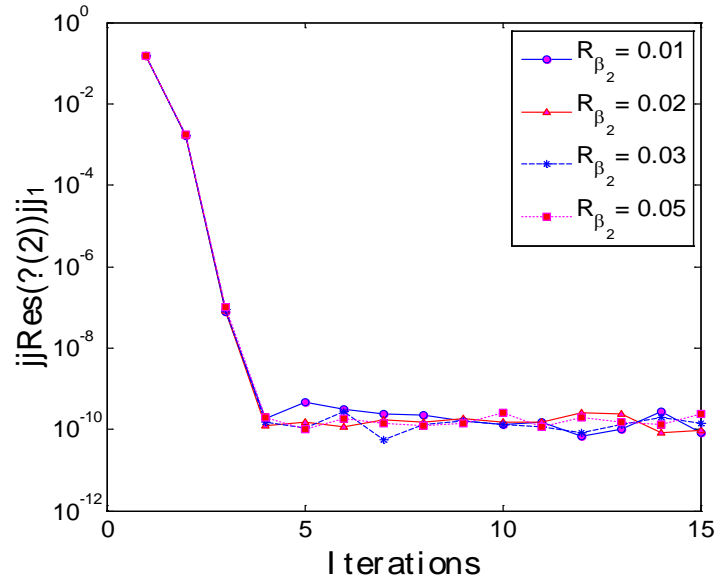


Fig.7: Residual error in concentration profiles for different values of β_2 .

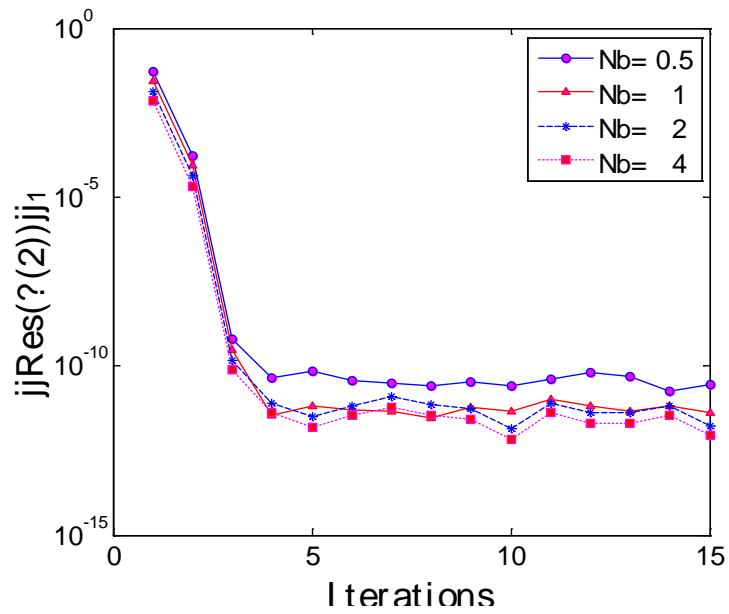


Fig.8: Residual error in concentration profiles for different values of Nb .

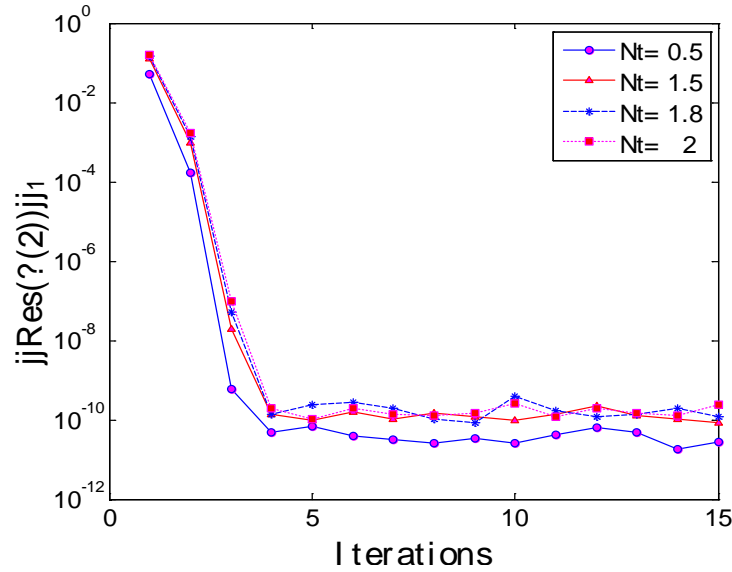


Fig.9: Residual error in concentration profiles for different values of Nt .

Fig. 10 shows the effect of the magnetic parameter on the entropy generation number. An increase in magnetic parameter results in an increase in entropy generation number. In the neighborhood of the sheet vicinity, the magnetic field has a significant impact on the entropy generation number. This tends to increase the resistance of the fluid motion, and consequently, the heat transfer rates. However, far from the sheet vicinity, the influence of magnetic parameter is insignificant.

Fig. 11 shows entropy generation with the Brinkman number, which represents a measure of the significance of the heat produced by viscous heating proportional to heat transported by molecular conduction. An increase in the Brinkman number tends to increase the entropy generation number especially near the sheet.

Fig. 12 shows the behavior of the entropy generation number with the temperature difference parameter. We observe that in the neighborhood of the sheet, the entropy generation number increases with the temperature difference parameter.

Fig. 13 relates the entropy generation number to the Reynolds number. We note that the Reynolds number has a significant impact on the entropy generation number as an increase in the Reynolds number leads to a significant increase in the entropy generation number, near the sheet. By increasing the Reynolds number, the fluid acceleration increases near the sheet.

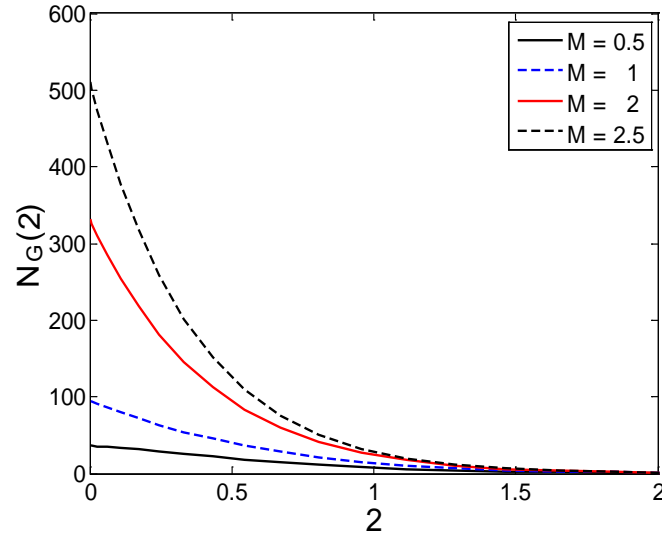


Fig. 10: Effects of the magnetic parameter on the entropy generation number N_G .

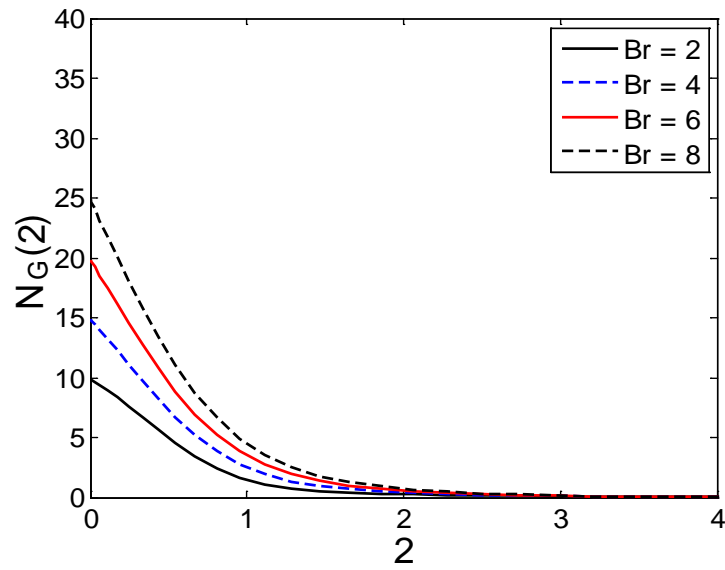


Fig. 11: Effects of Brinkman number on the entropy generation number N_G .

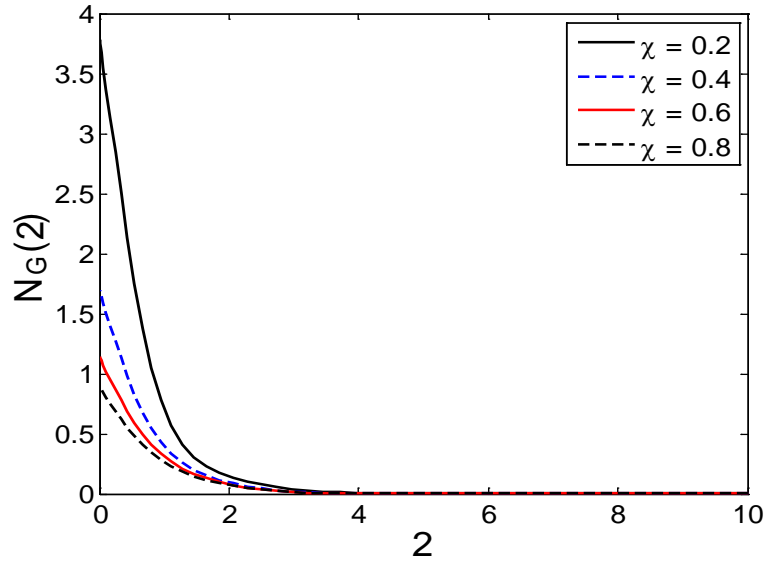


Fig. 12: Effects of temperature difference parameter on the entropy generation number N_G .

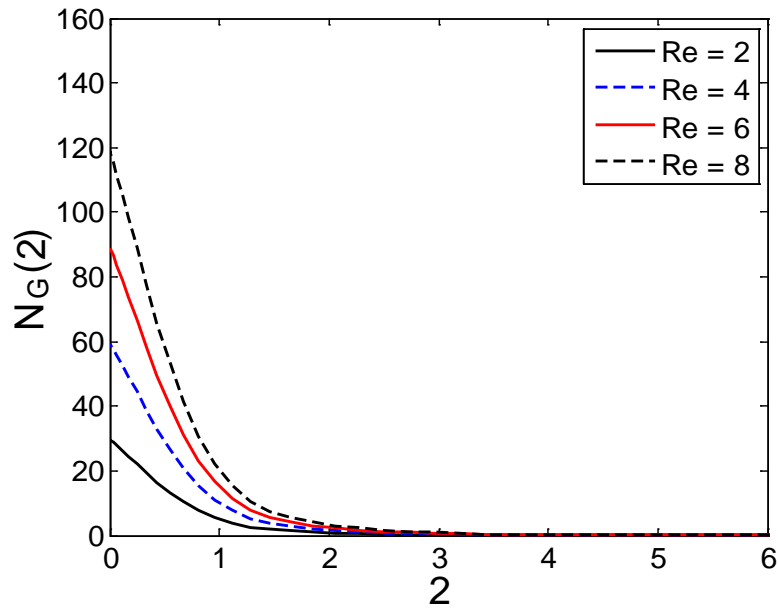


Fig. 13: Effects of Reynolds number on the entropy generation N_G .

Fig. 14 shows that the Bejan number is proportionally related to the magnetic parameter. As magnetic parameter increases, the entropy generation traced to diffusive irreversibility and magnetic field is totally controlled by the entropy generation due to heat transfer at the vicinity of the sheet. Figs.15 and 16 show the variations in the Bejan number with different values of the Brinkman number and the temperature difference parameter. We observe that an increase in the Brinkman number and temperature difference parameter leads to an increases in the Bejan number. An increase in the Brinkman number and the temperature difference parameter contribute to the increase in the magnitude of the diffusive irreversibility. However, the Brinkman number and the temperature difference parameter have no influence on heat

transfer irreversibility. Consequently, the irreversibility ratio increases and the Bejan number decreases.

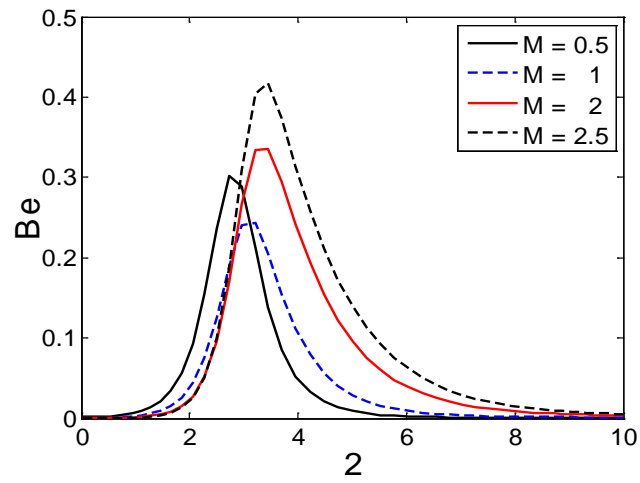


Fig.14: Effects of the magnetic parameter on the Bejan number (Be).

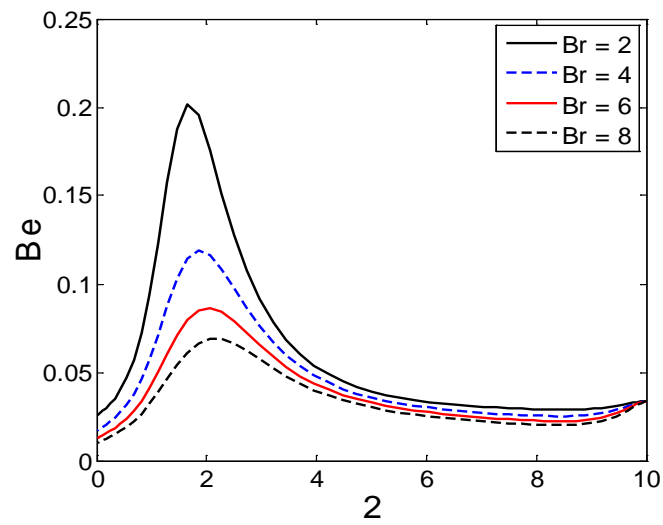


Fig.15: Effects of the Brinkman number on the Bejan number (Be).

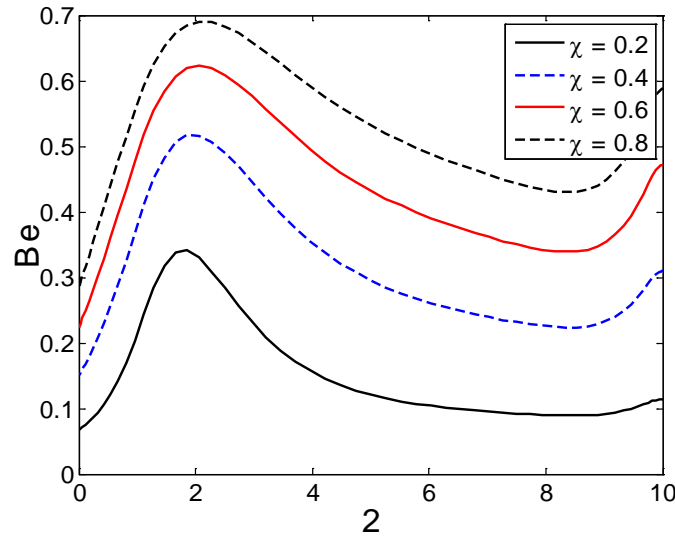


Fig. 16: Effects of temperature difference parameter on the Bejan number (Be).

Conclusions:

In this paper, we have investigated the influence of some fluid parameters on the entropy generation in Sakiadis nanofluid flow using Cattaneo-Christov heat flux model. The flow is over a stretching sheet subject to an inclined magnetic field. The transport equations were solved numerically using the spectral quasi-linearization method. The accuracy of the solutions was determined through an analysis of error norms and residual errors. Some key findings from the study include the following:

- The entropy generation number increases with an increase in temperature difference parameter, the Brinkman number and the Reynolds number.
- The Bejan number is strongly affected by variations in the temperature difference parameter and the Brinkman number.

Acknowledgment:

The authors would like to thanks Claude Leon Foundation Postdoctoral Fellowship, to the University of KwaZulu-Natal, South Africa, financially supported the research work.

References:

- [1]. Fourier, J.B. J. *Theorie analytique De La chaleur*, Paris 1822.
- [2]. Cattaneo C. (1948) Sulla conduzione del calore, *Atti Semin. Mat. Fis. Univ. Modena Reggio Emilia* **3**, 83–101.
- [3]. Christov, C.I. (2009). On frame in different formulation of the Maxwell—Cattaneo model of finite-speed heat conduction. *Mechanics Research Communications* **36**, 481–486
- [4]. Ciarletta M., Straughan, B. (2010). Uniqueness and structural stability for the Cattaneo—Christov equations. *Mechanics Research Communications* **37**: 445–447.
- [5]. Mushtaq, A., Abbasbandy, S., Mustafa, M., Hayat, T. and Alsaedi A. (2016) Numerical solution for Sakiadis flow of upper-convected Maxwell fluid using Cattaneo-Christov heat flux model, *AIP Advances* **6**, 015-208.
- [6]. Han S., Zheng, L., Li, C., and Zhang, X. (2014). Coupled flow and heat transfer in viscoelastic fluid with Cattaneo—Christov heat flux model. *Applied Mathematics Letters* **38**, 87–93

- [7]. Salahuddin, T., Malik, M. Y., Hussain, A., Bilal, S., and Awais, M. (2016). MHD flow of Cattaneo–Christov heat flux model for Williamson fluid over a stretching sheet with variable thickness, *Journal of Magnetism and Magnetic Materials*, **401**, 991–997.
- [8]. Dogonchi, A.S., Ganji, D.D. (2017) Impact of Cattaneo–Christov heat flux on MHD Nanofluid flow and heat transfer between parallel plates considering thermal radiation effect. *Journal of the Taiwan Institute of Chemical Engineers* **80**, 52–63.
- [9]. Sithole, H., Mondal, H., Sibanda, P. (2018) Entropy generation in a second grade magnetohydrodynamic nanofluid flow over a convectively heated stretching sheet with nonlinear thermal radiation and viscous dissipation. *Results in Physics* **9**, 1077–1085.
- [10]. Rashidi, M.M., Ganesh, N.V., Hakeem, A.K.A., and Ganga, B. (2014) Buoyancy effect on MHD flow of nanofluid over a stretching sheet in the presence of thermal radiation, *Journal of Molecular Liquids*, 198, 234–238.
- [11]. Khan W.A., Pop, I. (2010) Boundary-layer flow of a nanofluid past a stretching sheet. *International Journal of Heat and Mass Transfer* **53**, 2477–2483.
- [12]. Buongiorno, J. (2006) Convective transport in nanofluids. *Journal of Heat Transfer* **128**(3):240–250.
- [13]. Kuznetsov A.V., Nield, D.A. (2010) Natural convective boundary-layer flow of a nanofluid past a vertical plate. *International Journal of Thermal Science* **49**(2), 243–247.
- [14]. De, P., Mondal, H., and Bera, U.K. (2016) Dual solutions of heat and mass transfer of nanofluid over a stretching/shrinking sheet with thermal radiation. *Meccanica* **51**(1), 117–124.

Homogenization of graphene sheet reinforced composites considering material and geometrical uncertainty

Dimitrios Savvas¹ and †*George Stefanou²

¹Institute of Structural Analysis & Antiseismic Research, National Technical University of Athens, 9 Iroon Polytechniou, Zografou Campus, 15780 Athens, Greece.

²Institute of Structural Analysis & Dynamics of Structures, Department of Civil Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece.

*Presenting author: gstefanou@civil.auth.gr

†Corresponding author: gstefanou@civil.auth.gr

Abstract

This paper presents a computational procedure for the determination of the stochastic material properties of defective graphene sheets. The lattice of graphene is modeled using the molecular structural mechanics approach, which is a continuum based nanoscale modeling technique, where the C-C covalent bonds are replaced by energetically equivalent beam elements. The statistical properties for each component of the elasticity matrix are obtained by performing Monte Carlo simulations on randomly generated finite element models of defective graphene sheets. Moreover, computational homogenization of graphene sheet reinforced composites is performed considering material and geometrical uncertainty. The results demonstrate the effect of combined uncertainty on the homogenized properties of the composite material.

Keywords: Graphene, Vacancy defects, Microstructural uncertainty, Homogenization, Composites.

Introduction

Graphene is an allotrope of carbon consisting of a single layer of carbon atoms arranged in a hexagonal lattice exhibiting superb mechanical and physical properties (approximately Young's modulus 1 TPa, ultimate strength 130 GPa, thermal conductivity $3000 \text{ W m}^{-1} \text{ K}^{-1}$, electric conductivity $2.5 \times 10^5 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$). These exceptional properties of graphene along with its high aspect ratio make it ideal reinforcement in composite materials. This is illustrated among others in Aluko et al. [1], where the elastic response of graphene nanoplatelet (GNP) reinforced composites was shown to increase with increased GNP volume fraction, dispersion, and strain rates. Bending and buckling analyses of functionally graded polymer composite plates reinforced with GNP conducted in Song et al. [25] have shown that an addition of a very small amount of GNPs into the polymer matrix can significantly reduce the bending deflections and increase the critical buckling load.

In order to study the mechanical behavior of graphene reinforced polymers, the characterization of the mechanical properties of isolated graphene is important. For this purpose, a lot of experimental [4][11][18] and numerical [16][13][10][24] studies have been conducted. However, there are limited studies computing the mechanical properties of defective graphene lattices (e.g. [2][28][19][20]). A direct experimental evidence for the existence of defects (e.g. topological defects, vacancies and adatoms) in graphene layers has been provided by Hashimoto et al. [5].

In this paper, statistical properties for the components of the elasticity tensor of graphene lattices, which contain randomly dispersed single vacancy defects are computed first. Then, composites containing randomly dispersed graphene sheets to which homogenized stiffness properties have been assigned, are analyzed. The presented numerical results demonstrate the effect of combined material and geometrical uncertainty on the homogenized properties of graphene sheet reinforced composite materials.

Computation of random material properties of graphene sheets

Structure of graphene

Graphene is an allotrope of carbon in the form of a two-dimensional, atomic-scale, hexagonal lattice in which one atom forms each vertex. Each atom has four bonds, one σ -bond with each of its three neighbors and one π -bond that is oriented out of plane. The distance L_{C-C} between the carbon atoms is about 1.42 Å. Graphene's stability into a single layer of carbon atoms is due to its tightly packed carbon atoms and a sp^2 orbital hybridization. The thickness of a monolayer graphene is about 0.34 nm, which corresponds to the interlayer spacing of graphite and the thickness of one carbon atom.

Graphene is the basic structural element of other allotropes, including graphite and carbon nanotubes which can be formed via stacking and wrapping of the graphene's layers, respectively. Specifically, the helicity of a carbon nanotube is characterized by the roll-up vector \vec{C}_h . This is called the chiral vector and it is defined as:

$$\vec{C}_h = n\vec{\alpha}_1 + m\vec{\alpha}_2 \quad (1)$$

where $\vec{\alpha}_1 = [\sqrt{3}L_{C-C}, 0]$ and $\vec{\alpha}_2 = 0.5[\sqrt{3}L_{C-C}, 3L_{C-C}]$ are basis vectors of length $\sqrt{3}L_{C-C}$ defined on the hexagonal lattice of graphene (see Fig. 1) The pair of indices (n, m) define the chiral angle $\theta_{\vec{C}_h}$ as

$$\theta_{\vec{C}_h} = \tan^{-1}\left(\frac{\sqrt{3}m}{2n+m}\right) \quad (2)$$

As it can be deduced from Eq. (2), $\theta_{\vec{C}_h} = 0^\circ$ for zig-zag direction ($m=0$) and $\theta_{\vec{C}_h} = 30^\circ$ for armchair direction ($m=n$). For all other directions of the chiral vector, $\theta_{\vec{C}_h} \in (0^\circ, 30^\circ)$.

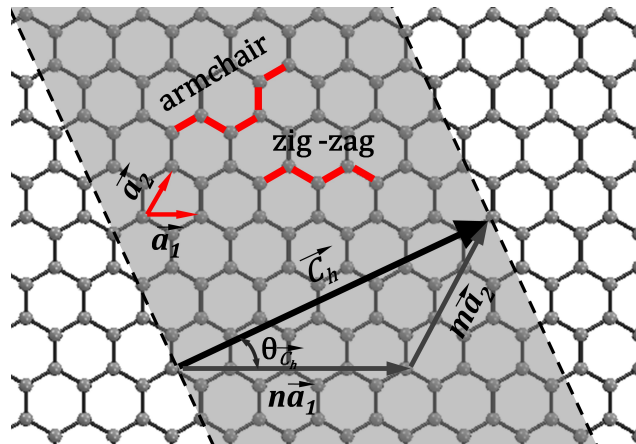


Figure 1. Graphene lattice structure and definition of chiral vector \vec{C}_h .

The molecular structural mechanics approach

The molecular structural mechanics (MSM) approach is a continuum based nanoscale modeling technique developed by Li and Chou [12], which has attracted great attention because of its simplicity and effectiveness. In this framework, the potential energy produced by the C-C atomic interactions is equating to the sum of energies produced by the deformations of a beam element, which substitute the C-C chemical bond. Thus, the method results in a space frame model built by connected beam elements, which is equivalent to the atomistic model of the graphene's lattice. In principle, this approach provides a linkage between molecular mechanics and continuum structural mechanics by which geometry and material properties of the beam elements are obtained.

In the context of molecular mechanics, graphene can be regarded as a molecular system consisting of carbon atoms. The lattice deformation under a specific load is governed by the atomic motions which are regulated by a force field. This force field, which is generated by electron-nucleus and nucleus-nucleus interactions, is usually expressed in the form of a steric potential energy. The general expression of this total steric potential energy is a sum of energies due to valence or bonded and non-bonded interactions as follows:

$$U = U_r + U_\theta + U_\phi + U_\omega + U_{vdW} + U_{es} \quad (3)$$

where U_r , U_θ , U_ϕ , U_ω are the bond-stretching energy, the bond-angle variation energy, the dihedral-angle torsion energy and the inversion (out of plane torsion) energy, respectively. U_{vdW} and U_{es} are associated with non-bonded van der Waals and electrostatic interactions, respectively, which are usually negligible and therefore are omitted. The corresponding interatomic interactions are schematically represented in Fig. 2.

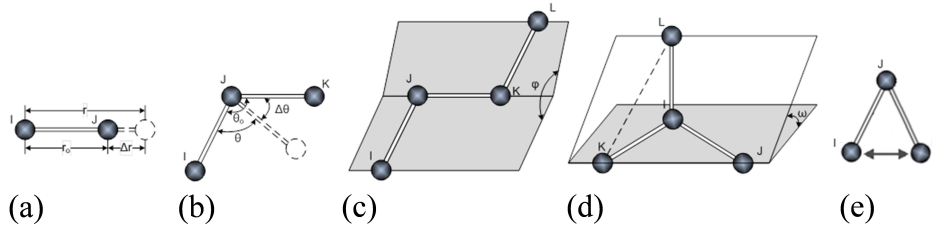


Figure 2. Interatomic interactions in molecular mechanics: a) stretching, b) bending, c) dihedral angle torsion, d) out of plane torsion and e) van der Waals.

By adopting a quadratic potential force field to account for linear covalent C-C interactions and a circular beam element in the context of the finite element (FE) method, the following relationships between structural mechanics parameters and molecular mechanics force field constants are derived

$$d = 4 \sqrt{\frac{k_\theta}{k_\tau}}, \quad E = \frac{k_r L_{C-C}}{4\pi k_\theta}, \quad G = \frac{k_r^2 k_\tau L_{C-C}}{8\pi k_\theta^2} \quad (4)$$

Substituting the force field constant values $k_r=938 \text{ kcal} \cdot \text{mole}^{-1} \cdot \text{\AA}^{-2}=6.52 \times 10^{-7} \text{ N} \cdot \text{nm}^{-1}$, $k_\theta=126 \text{ kcal} \cdot \text{mole}^{-1} \cdot \text{rad}^{-2}=8.76 \times 10^{-10} \text{ N} \cdot \text{nm} \cdot \text{rad}^{-2}$ and $k_\tau=40 \text{ kcal} \cdot \text{mole}^{-1} \cdot \text{rad}^{-2}=2.78 \times 10^{-10} \text{ N} \cdot \text{nm} \cdot \text{rad}^{-2}$, defined by Cornell et al. [3], the values of diameter $d=0.147 \text{ nm}$, Young's modulus $E=5.49 \text{ TPa}$ and shear modulus $G=0.871 \text{ TPa}$ of the beam elements are obtained.

Random material properties of defective graphene sheets

Various structural defects are generated during the production or chemical functionalization process of graphene sheets such as Stone-Wales (SW), single vacancy (SV) and double vacancy (DV) defects [5][14][26]. The most common types of defects are the vacancy defects which are characterized by the number of the absent atoms from the lattice of graphene. In this paper, we will examine graphene sheets containing only SV defects which are formed by removing one carbon atom and its three adjacent bonds from the lattice.

The FE model of a defective graphene sheet is constructed using 2d Bernoulli beams with properties obtained from the MSM approach described in the previous section. The defects are uniformly distributed on the lattice. Each defect is treated as an individual scattering center (independently of the defect type or the number of atoms that forms it). The defect density is defined as $N_d/N_a \times 100$ where N_d and N_a are the number of defects and number of atoms in the pristine graphene, respectively. A FE model of a defective graphene sheet containing 3% randomly dispersed SV defects is shown in Fig. 3.

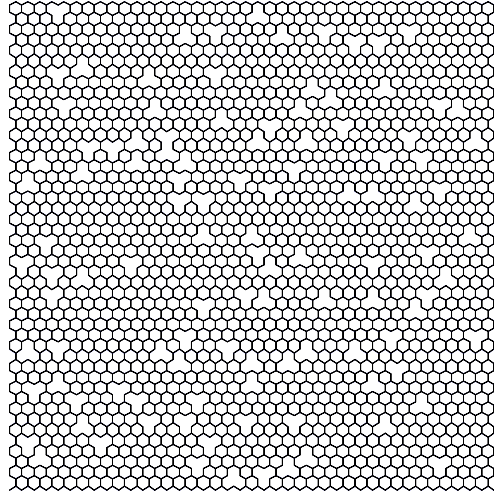


Figure 3. FE model of a defective graphene sheet containing 3% randomly dispersed SV defects.

The homogenized in-plane anisotropic stress-strain relation of the graphene material can be expressed by:

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ & C_{22} & C_{23} \\ \text{sym} & & C_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} \quad (5)$$

with $[\sigma_1 \ \sigma_2 \ \sigma_3]^T = [\sigma_{xx} \ \sigma_{yy} \ \sigma_{xy}]^T$ and $[\varepsilon_1 \ \varepsilon_2 \ \varepsilon_3]^T = [\varepsilon_{xx} \ \varepsilon_{yy} \ \varepsilon_{xy}]^T$ the stress and strain tensors respectively and \mathbf{C} the homogenized anisotropic elasticity tensor. Notice that, due to symmetry, only six components C_{ij} , with $i, j=1,2,3$ are needed to fully determine the homogenized elasticity matrix. In order to solve for these unknown parameters, three independent uniform strain (Dirichlet) boundary conditions are applied on a square graphene sheet of size $10 \times 10 \text{ nm}^2$, which are derived from the following set of strain deformation cases:

$$\boldsymbol{\varepsilon} = \left\{ \begin{bmatrix} \varepsilon_1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \varepsilon_2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \varepsilon_3 \end{bmatrix} \right\} \quad (6)$$

Substituting the first strain vector in Eq. (5), the components $C_{11}=\sigma_1/\varepsilon_1$, $C_{21}=\sigma_2/\varepsilon_1$ and $C_{31}=\sigma_3/\varepsilon_1$ can be directly calculated, while the remaining unknown components of \mathbf{C} are computed in a similar way by applying all other strain vectors of Eq. (6). The statistical properties for each component C_{ij} of the elasticity matrix are obtained by performing 1000 Monte Carlo simulations on randomly generated FE models of defective graphene sheets containing 3% randomly dispersed SV defects. The mean and standard deviation of the average axial stiffness ($C_{ii}/2$, $i=1, 2$) are calculated as $\mu=910.73$ GPa and $\sigma=29.95$ ($\text{COV}=\sigma/\mu \approx 3\%$), respectively and of the shear stiffness C_{33} as $\mu=413.63$ GPa and $\sigma=16.41$ ($\text{COV} \approx 4\%$), respectively. The histograms of $C_{ii}/2$ and C_{33} are depicted in Fig. 4.

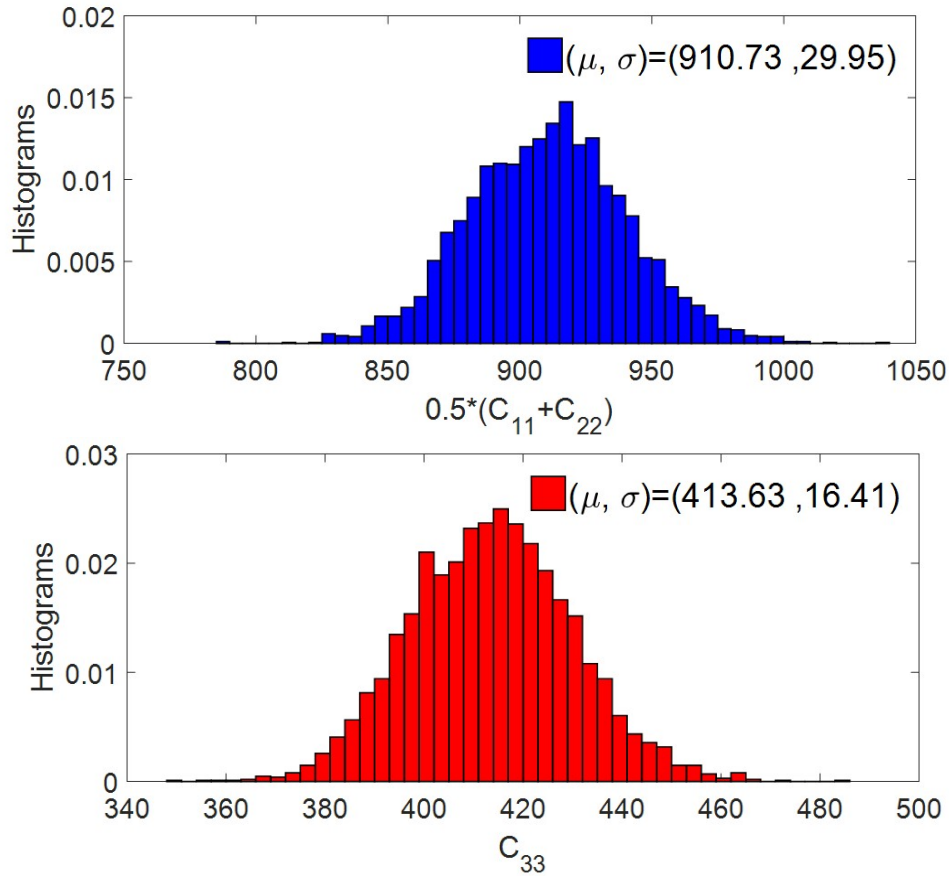


Figure 4. Histograms for average axial stiffness and shear stiffness of graphene sheet with 3% randomly dispersed SV defects obtained from 1000 Monte Carlo simulations.

Determination of effective material properties for graphene sheet reinforced composites

The computational procedure for the determination of effective material properties for graphene sheet reinforced composites is illustrated in Fig. 5. In the first step, 1000 Monte Carlo simulations are performed in order to compute the statistical properties for the components of the elasticity tensor of defective graphene lattices containing 3% randomly dispersed SV defects. The graphene lattices are analyzed using the MSM approach described previously. Then, homogenized graphene sheets with material properties considered as

random variables are randomly dispersed into a polymer matrix. The resulting composites are analyzed using the extended finite element method (see Savvas et al. [23]) and homogenized mechanical properties are computed by performing Monte Carlo simulations as described in the next section.

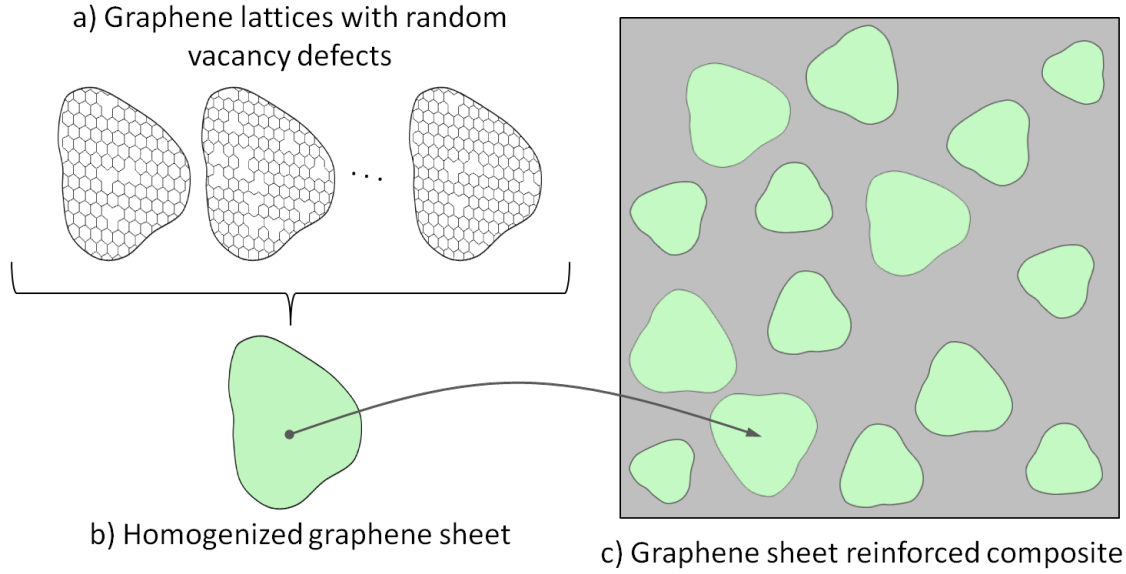


Figure 5. a) Graphene lattices with random vacancy defects, b) Homogenized graphene sheet derived from Monte Carlo simulation and homogenization, c) Graphene sheet reinforced composite material with random dispersion of inclusions.

Computational homogenization

The homogenization scheme adopted in this paper is based on the fundamental assumption of statistical homogeneity of the heterogeneous medium [6] which means that all statistical properties of the state variables are the same at any material point and thus a representative volume element (RVE) can be identified. It has been shown in the literature that the size of the RVE is related to a scale factor $\delta = l_{\text{meso}}/l_{\text{micro}}$ (under the condition $l_{\text{micro}} \ll l_{\text{meso}}$ with l_{micro} the length scale of the reinforcement (e.g. diameter, thickness, length) and l_{meso} the length scale of the examined volume element [17] [27][21][22]. Satisfaction of this condition leads to homogenized material properties independent of the type of the boundary conditions imposed on the model. Note also that the relation $l_{\text{meso}} \ll l_{\text{macro}}$, with l_{macro} denoting the characteristic length over which the macroscopic loading varies in space, must always hold for complete separation of scales. In this case, the uniformity of microscopic stress and strain fields near the boundary surface $\partial\Omega$ of the mesoscale model Ω , required in the context of Hill's homogenization theory, is valid [8].

The Hill-Mandel homogeneity condition postulates that the strain energy computed on a material point of the macro-continuum medium has to be equal to that computed over the mesoscale volume element in an average sense:

$$\bar{\sigma} : \bar{\epsilon} = \frac{1}{|V|} \int_{\Omega} (\sigma : \epsilon) dx \quad (7)$$

where V is the volume of the mesoscale volume element and the macroscopic stresses and strains are computed as:

$$\bar{\boldsymbol{\sigma}} = \frac{1}{|V|} \int_{\Omega} \boldsymbol{\sigma}(\mathbf{x}) d\mathbf{x}, \quad \bar{\boldsymbol{\varepsilon}} = \frac{1}{|V|} \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{x}) d\mathbf{x} \quad (8)$$

Eq. (8) is valid provided that the following constraint is satisfied:

$$\frac{1}{|V|} \int_{\partial\Omega} (\mathbf{t} - \bar{\boldsymbol{\sigma}} \cdot \mathbf{n})(\mathbf{u} - \bar{\boldsymbol{\varepsilon}} \cdot \mathbf{x}) dS = 0 \quad (9)$$

which is a priori satisfied by both kinematic uniform boundary conditions (KUBCs) and static uniform boundary conditions (SUBCs) as shown in Huet [9]. The constraint of Eq. (9) can be also satisfied by orthogonal uniform mixed boundary conditions (OUMBCs) [7] and periodic boundary conditions (PBCs) [15].

Miehe and Koch [15] proposed a computational procedure to define homogenized stresses and overall tangent moduli of microstructures undergoing small strains. They have shown that homogenized properties can be defined in terms of discrete forces and stiffness properties on the boundary of discretized microstructures. Using these deformation-driven algorithms, the homogenized elasticity tensor of a mesoscale model can be calculated by solving a kinematic uniform or a static uniform boundary value problem.

Specifically, for the case of kinematic uniform boundary conditions, a prescribed uniform strain tensor $\bar{\boldsymbol{\varepsilon}} = [\bar{\varepsilon}_{11}, \bar{\varepsilon}_{22}, 2\bar{\varepsilon}_{12}]$ is applied on the boundary $\partial\Omega$ of a discretized mesoscale model Ω through displacements in the form:

$$\mathbf{u}_b = \mathbf{D}_b^T \bar{\boldsymbol{\varepsilon}} \quad (10)$$

where \mathbf{D}_b is a geometric matrix which depends on the coordinates of the boundary node b and is defined as:

$$\mathbf{D}_b = \frac{1}{2} \begin{bmatrix} 2x_b & 0 \\ 0 & 2y_b \\ y_b & x_b \end{bmatrix} \text{ with } (x_b, y_b) \in \partial\Omega \quad (11)$$

Note that the stiffness matrix \mathbf{K} of the extended finite element model of the graphene sheet reinforced composite can be rearranged into sub-matrices associated with interior nodes i and boundary nodes b . Thus the static problem is denoted by:

$$\begin{bmatrix} \mathbf{K}_{ii} & \mathbf{K}_{ib} \\ \mathbf{K}_{bi} & \mathbf{K}_{bb} \end{bmatrix} \begin{bmatrix} \mathbf{U}_i \\ \mathbf{U}_b \end{bmatrix} = \begin{bmatrix} \mathbf{F}_i \\ \mathbf{F}_b \end{bmatrix} \quad (12)$$

Then the homogenized elasticity tensor $\mathbf{C}(\theta)$ can be calculated in terms of the condensed stiffness matrix $\tilde{\mathbf{K}}_{bb} = \mathbf{K}_{bb} - \mathbf{K}_{bi} \mathbf{K}_{ii}^{-1} \mathbf{K}_{ib}$ in the form:

$$\mathbf{C}(\theta) = \frac{1}{|V|} \mathbf{D} \tilde{\mathbf{K}}_{bb} \mathbf{D}^T \quad (13)$$

where $\mathbf{D} = [\mathbf{D}_1 \ \mathbf{D}_2 \ \cdots \ \mathbf{D}_M]$ with M the total number of boundary nodes. Note that θ denotes the randomness of the computed homogenized elasticity tensor due to material uncertainty related to the random properties assigned to graphene sheets and to geometrical uncertainty related to the random dispersion of the inclusions (graphene sheets) within the composite material.

Numerical results

In this section, the two sources of microstructural uncertainty (material and geometrical) are considered simultaneously and their effect on the homogenized constitutive properties of the composite material is assessed. Note that the composites are assumed to be under plane stress conditions. The statistical properties of the homogenized material for the graphene sheets have been calculated in a previous section. The material of the matrix is considered isotropic and linearly elastic with Young's modulus $E_m=1$ GPa and Poisson's ratio $\nu_m=0.3$. Fig. 6 illustrates the empirical distribution and the statistical convergence of mean and COV of the axial stiffness $C_{ii}/2$ ($i=1,2$) of the graphene sheet reinforced composite material. Similar results are presented in Fig. 7 for the shear stiffness C_{33} . Note that all composites are reinforced with 40% vf of graphene sheets with arbitrary shape (see Fig. 5). It is noted that the first two statistical moments of the average axial and shear stiffness of the above composites seem to converge sufficiently within 1000 Monte Carlo simulations.

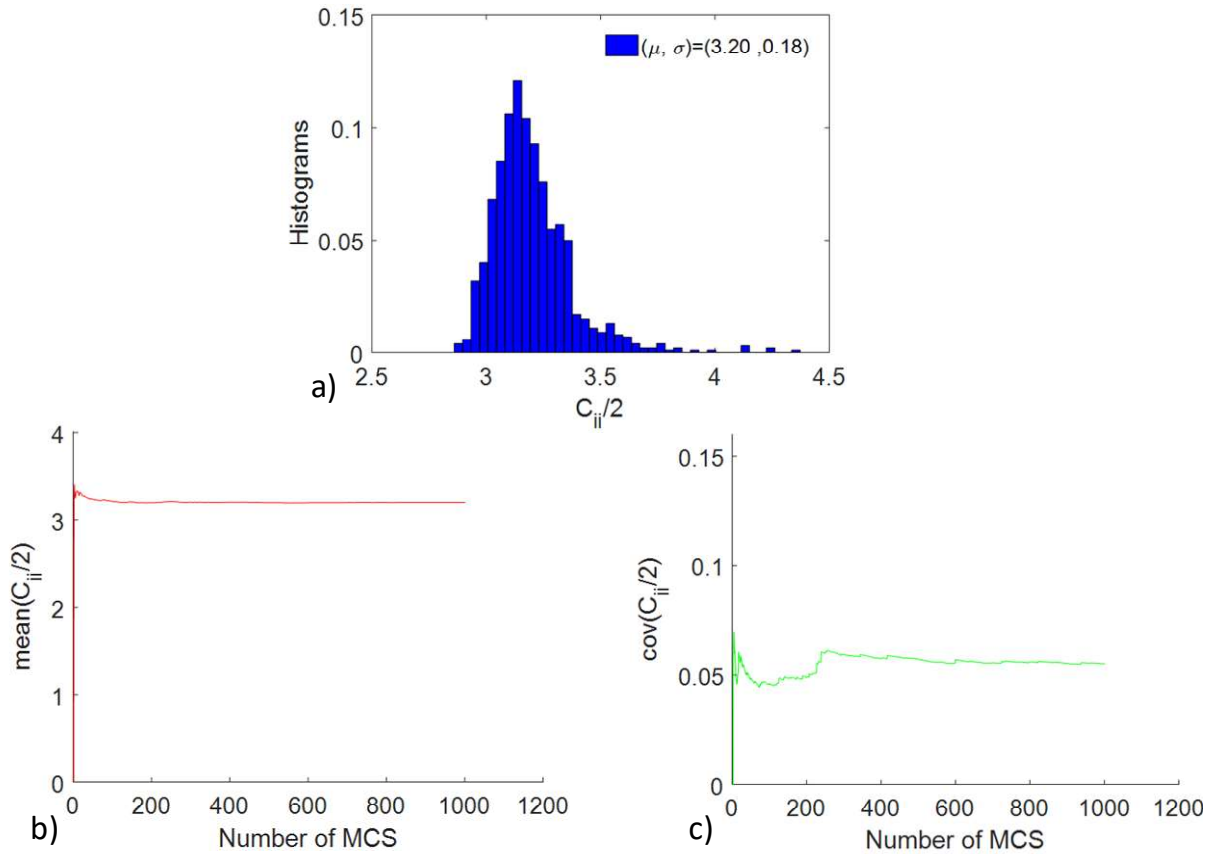


Figure 6. a) Histogram of average axial stiffness $C_{ii}/2$, b) Statistical convergence of mean($C_{ii}/2$) with respect to the number of Monte Carlo simulations, c) Statistical convergence of COV($C_{ii}/2$) with respect to the number of Monte Carlo simulations.

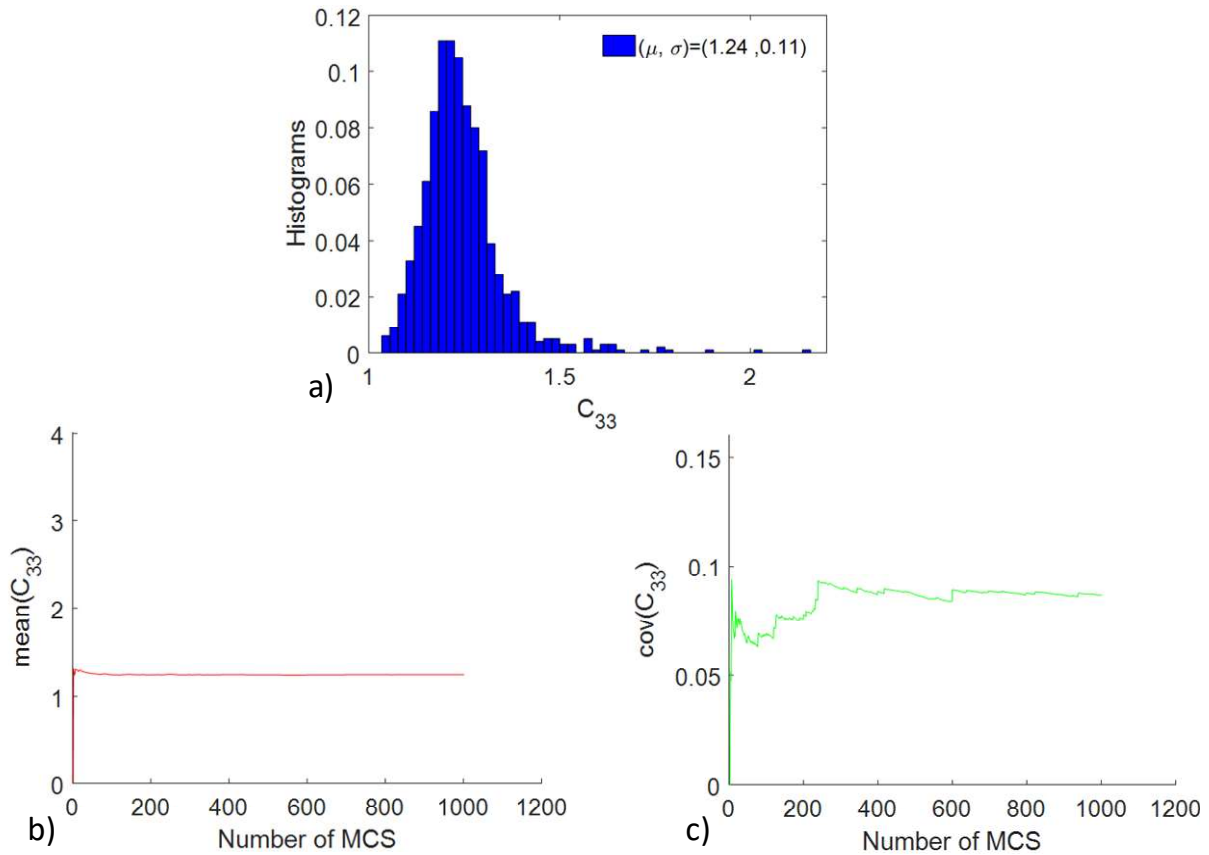


Figure 7. a) Histogram of shear stiffness C_{33} , b) Statistical convergence of $\text{mean}(C_{33})$ with respect to the number of Monte Carlo simulations, c) Statistical convergence of $\text{COV}(C_{33})$ with respect to the number of Monte Carlo simulations.

Conclusions

In this paper, the effect of the two sources of microstructural uncertainty (material and geometrical) on the homogenized elastic properties of graphene sheet reinforced composites was assessed. The lattice of graphene was modeled using the molecular structural mechanics approach. The statistical properties for each component of the elasticity matrix were obtained by performing Monte Carlo simulations on randomly generated finite element models of defective graphene sheets. Subsequently, computational homogenization of graphene sheet reinforced composites was performed considering material and geometrical uncertainty. A magnification of uncertainty was observed, as the COV of the material properties of the composite was larger than the corresponding COV of the reinforcements, especially in the case of shear stiffness.

References

- [1] Aluko, O., Gowtham, S., and Odegard, G. (2017). Multiscale modeling and analysis of graphene nanoplatelet/carbon fiber/epoxy hybrid composite. *Composites Part B: Engineering*, **131**:82-90.
- [2] Askari, D. and Ghasemi-Nejhad, M. N. (2011). Effects of vacancy defects on mechanical properties of graphene/carbon nanotubes: A numerical modeling. *Journal of Computational and Theoretical Nanoscience*, **8**(4):783-794.
- [3] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, **117**(19):5179-5197.
- [4] Frank, I., Tanenbaum, D. M., van der Zande, A. M., and McEuen, P. L. (2007). Mechanical properties of suspended graphene sheets. *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, **25**(6):2558-2561.

- [5] Hashimoto, A., Suenaga, K., Gloter, A., Urita, K., and Iijima, S. (2004). Direct evidence for atomic defects in graphene layers. *Nature*, **430**(7002):870.
- [6] Hashin, Z. (1983). Analysis of composite materials: A survey. *Journal of Applied Mechanics*, **50**(2):481-505.
- [7] Hazanov, S. and Huet, C. (1994). Order relationships for boundary conditions effect in heterogeneous bodies smaller than the representative volume. *Journal of the Mechanics and Physics of Solids*, **42**(12):1995-2011.
- [8] Hill, R. (1963). Elastic properties of reinforced solids: some theoretical principles. *Journal of the Mechanics and Physics of Solids*, **11**(5):357-372.
- [9] Huet, C. (1990). Application of variational concepts to size effects in elastic heterogeneous bodies. *Journal of the Mechanics and Physics of Solids*, **38**(6):813-841.
- [10] Jiang, J.-W., Wang, J.-S., and Li, B. (2009). Young's modulus of graphene: A molecular dynamics study. *Physical Review B*, **80**:113405.
- [11] Lee, C., Wei, X., Kysar, J. W., and Hone, J. (2008). Measurement of the elastic properties and intrinsic strength of monolayer graphene. *Science*, **321**(5887):385-388.
- [12] Li, C. and Chou, T.-W. (2003). A structural mechanics approach for the analysis of carbon nanotubes. *International Journal of Solids and Structures*, **40**(10):2487-2499.
- [13] Liu, F., Ming, P., and Li, J. (2007). Ab initio calculation of ideal strength and phonon instability of graphene under tension. *Physical Review B*, **76**(6):064120.
- [14] Liu, L., Qing, M., Wang, Y., and Chen, S. (2015). Defects in graphene: generation, healing, and their effects on the properties of graphene: a review. *Journal of Materials Science & Technology*, **31**(6):599-606.
- [15] Miehe, C. and Koch, A. (2002). Computational micro-to-macro transitions of discretized microstructures undergoing small strains. *Archive of Applied Mechanics*, **72**(4-5):300-317.
- [16] Odegard, G. M., Gates, T. S., Nicholson, L. M., and Wise, K. E. (2002). Equivalent-continuum modeling of nano-structured materials. *Composites Science and Technology*, **62**(14):1869-1880.
- [17] Ostoja-Starzewski, M. (2006). Material spatial randomness: From statistical to representative volume element. *Probabilistic engineering mechanics*, **21**(2):112-132.
- [18] Poot, M. and van der Zant, H. S. (2008). Nanomechanical properties of few-layer graphene membranes. *Applied Physics Letters*, **92**(6):063111.
- [19] Rafiee, R. and Eskandariyun, A. (2017). Comparative study on predicting young's modulus of graphene sheets using nano-scale continuum mechanics approach. *Physica E: Low-dimensional Systems and Nanostructures*, **90**:42-48.
- [20] Savvas, D. and Stefanou, G. (2018). Determination of random material properties of graphene sheets with different types of defects. *Composites Part B: Engineering*, **143**:47-54.
- [21] Savvas, D., Stefanou, G., Papadopoulos, V., and Papadrakakis, M. (2016a). Effect of waviness and orientation of carbon nanotubes on random apparent material properties and rve size of cnt reinforced composites. *Composite Structures*, **152**:870-882.
- [22] Savvas, D., Stefanou, G., and Papadrakakis, M. (2016b). Determination of rve size for random composites with local volume fraction variation. *Computer Methods in Applied Mechanics and Engineering*, **305**:340-358.
- [23] Savvas, D., Stefanou, G., Papadrakakis, M., and Deodatis, G. (2014). Homogenization of random heterogeneous media with inclusions of arbitrary shape modeled by XFEM. *Computational Mechanics*, **54**(5):1221-1235.
- [24] Shokrieh, M. M. and Rafiee, R. (2010). Prediction of Young's modulus of graphene sheets and carbon nanotubes using nanoscale continuum mechanics approach. *Materials & Design*, **31**(2):790-795.
- [25] Song, M., Yang, J., and Kitipornchai, S. (2018). Bending and buckling analyses of functionally graded polymer composite plates reinforced with graphene nanoplatelets. *Composites Part B: Engineering*, **134**:106-113.
- [26] Tian, W., Li, W., Yu, W., and Liu, X. (2017). A review on lattice defects in graphene: Types, generation, effects and regulation. *Micromachines*, **8**(5):163.
- [27] Trovalusci, P., Ostoja-Starzewski, M., De Bellis, M. L., & Murrall, A. (2015). Scale-dependent homogenization of random composites as micropolar continua. *European Journal of Mechanics-A/Solids*, **49**, 396-407.
- [28] Tserpes, K. I. (2012). Strength of graphenes containing randomly dispersed vacancies. *Acta Mechanica*, **223**(4):669-678.

Meshless Method with reduced integration – high performance

*Wilber Vélez¹, Tiago Oliveira¹, Elvis Pereira¹ and †Artur Portela¹

¹Department of Civil and Environmental Engineering, University of Brasília, BR.

*Presenting author: wilbervelez@hotmail.com

†Corresponding author: aportela@unb.br

Abstract

The work theorem establishes an energy relationship between a statically – admissible stress field and an independent kinematically – admissible strain field in a local region. The Meshless Method with reduced integration (IMFM) is derived through a weighted-residual formulation that leads to the work theorem of structures theory.

In the formulation of the IMFM, the kinematically – admissible strain field is chosen as the one corresponding to an arbitrary rigid – body displacement; as a consequence, the domain term is canceled out and the work theorem is reduced to regular local boundary terms only. The Moving Least Squares (MLS) approximation of the elastic field is used to construct the trial function in this local meshless formulation. IMFM has a highly performance in problems with irregular nodal arrangement leading to accurate numerical results in two-dimensional elasticity problems.

This paper presents the size effect of the irregularity nodal arrangement (c_n) on energy and displacement relative error to solve the Timoshenko cantilever beam using different parameters of the local support domain (α_s), the local quadrature domain (α_q) and three different nodal discretization. Results are compared with the exact solution and the Meshless Local Petrov–Galerkin (MLPG) method and optimal results have been obtained for 2D problems plane stress.

Keywords: Meshless Method with reduce integration (IMFM), Meshless Local Petrov-Galerkin (MLPG), work theorem, Moving Least Squares (MLS), irregular nodal arrangement.

Introduction

Numerical methods based in grid, like Finite Element Method (FEM), are widely used for scientific researches. Grid-based methods required high quality meshes to solve fracture mechanics problems with material discontinuity, large deformation where excessive mesh distortion takes place and other situations. The meshless methods were generated with the expectation of providing more adaptive, accurate and stable numerical solutions that can deal with problems where conventional methods are not suitable [1]. Generally, their formulation is based in the weighted-residual method [2].

Different meshless methods have been developed during the last 20 years [3]. Some methods based on a weighted-residual weak-form formulation were applied in solid mechanics such as the Diffuse Element Method (DEM) [4], the Reproducing Kernel Particle Method (RKPM) [5], and the Element-free Galerkin (EFG) [6], were the first. After, other methods emerged based on local weighted-residual weak forms, such as the Meshless Local Petrov-Galerkin Method (MLPG) [7,8], the Meshless Local Boundary Integral Equation (MLBIE) [9], the Local Point Interpolation Method (LPIM) [10], Local Radial Point Interpolation Method

(LRPIM) [11], the meshless Finite Volume Method (FVM) [12], the Rigid-Body Displacement Mesh-Free (RBDMF) and the Generalized-Strain Mesh-Free (GSMF) [13].

The Meshless Method with reduce integration (IMFM) formulation presented by [14], the kinematically – admissible strain field is chosen as the one corresponding to an arbitrary rigid – body displacement; as a consequence, the domain term is canceled out and the work theorem is reduced to regular local boundary terms only. The Moving Least Squares (MLS) approximation of the elastic field is used to construct the trial function in this local meshless formulation. IMFM and the popular MLPG using the MLS approximation, this fact allowing to have more precise conclusions when comparing the two methods.

This paper is focused on the size effect of the configuration parameters of the local support domain (α_s), the local quadrature domain (α_q) and irregularity nodal arrangement (c_n). It presents a comparison of the energy and displacement relative error for three different irregular nodal distribution to solve the Timoshenko cantilever beam. The results are compared with the exact solution and the Meshless Local Petrov–Galerkin (MLPG) method, optimal parameters have been determined.

Methodology

Let Ω be the domain of a body and Γ its boundary subdivided in Γ_u and Γ_t that is $\Gamma = \Gamma_t \cup \Gamma_u$, as represented in Fig. 1. The general fundamental boundary value problem of linear elastostatics aims to determine the distribution of stresses σ , strains ε and displacements u , throughout the body, when it has constrained displacements \bar{u} , on Γ_u and is loaded by an external system of distributed surface and body forces with densities denoted, respectively by \bar{t} , on Γ_t and b , in Ω .

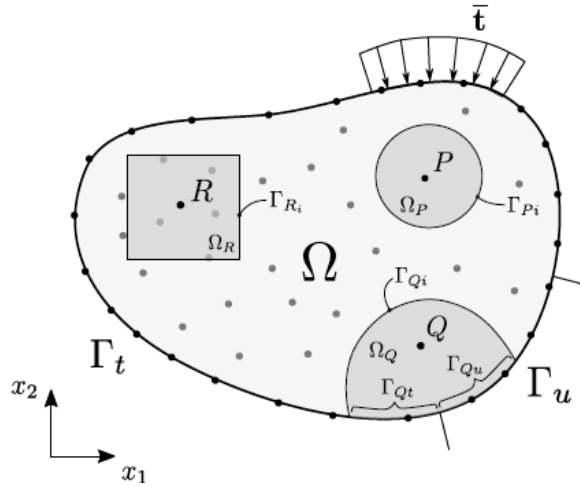


Figure 1. Meshless discretization of the global domain Ω and the local domains Ω_P , Ω_Q and Ω_R , with boundary $\Gamma = \Gamma_u \cup \Gamma_t$ represented.

The solution of these problem is a totally admissible elastic field that simultaneously satisfies the kinematic admissibility and the static admissibility. If this solution exists, it can be shown that it is unique, provided linearity and stability of the material are admitted [15, 16].

In the domain of the body, loaded by a system of external distributed surface and body forces with densities denoted, respectively by \bar{t} , on the boundary Γ_t and b , in the domain Ω , consider a statically admissible stress field σ , that is any stress field that satisfies equilibrium with the system of applied external forces which therefore satisfies

$$\mathbf{L}^T \boldsymbol{\sigma} + \mathbf{b} = 0 \quad (1)$$

in the domain Ω , with boundary conditions

$$\mathbf{t} = \mathbf{n}\boldsymbol{\sigma} = \bar{\mathbf{t}}, \quad (2)$$

on the static boundary Γ_t , in which $\boldsymbol{\sigma}$: Stress components, \mathbf{L} : Matrix differential operator, \mathbf{t} : represents the traction components, $\bar{\mathbf{t}}$: Prescribed tractions values, \mathbf{n} : Unit normal components to the boundary.

In the domain Ω , with boundary $\Gamma = \Gamma_t \cup \Gamma_u$, consider an arbitrary local domain Ω_Q , assigned to a reference point $Q \in \Omega_Q$, with local boundary $\Gamma_Q = \Gamma_{Qi} \cup \Gamma_{Qt} \cup \Gamma_{Qu}$, in which Γ_{Qi} is the interior local boundary, while Γ_{Qt} and Γ_{Qu} are local boundaries that share the global boundaries, respectively the static boundary Γ_t and the kinematic boundary Γ_u , as represented in Fig. 1. The work theorem is derived as a local form that is valid in an arbitrary domain Ω_Q , associated to the reference point Q .

The general work theorem of the theory of structures establishes an energy relationship, valid in an arbitrary local domain $\Omega_Q \cup \Gamma_Q \in \Omega \cup \Gamma$, between two independent elastic fields that can be defined in the body which are, respectively a statically admissible stress field that satisfies equilibrium with a system of external distributed surface and body forces, and a kinematically admissible strain field that satisfies compatibility with a set of constrained displacements. Derived as a weighted residual statement, the work theorem serves as a unifying basis for the formulation of numerical models Continuum Mechanics [17].

Expressed as an integral local form, defined in the local domain $\Omega_Q \cup \Gamma_Q$, which can be written in a compact form as

$$\int_{\Gamma_Q} \mathbf{t}^T \mathbf{u}^* d\Gamma + \int_{\Omega_Q} \mathbf{b}^T \mathbf{u}^* d\Omega = \int_{\Omega_Q} \boldsymbol{\sigma}^T \boldsymbol{\varepsilon}^* d\Omega \quad (3)$$

in which the stress field $\boldsymbol{\sigma}$ and the strain field $\boldsymbol{\varepsilon}^*$ are not linked by any constitutive relationship and therefore, they are independent of each other.

Kinematic formulations consider, in the work theorem, a particular and convenient specification of the kinematically admissible strain field, leading thus to an equation of mechanical equilibrium that is used in numerical models, to generate the respective stiffness matrix of each model. A simple case of local equilibrium equations, based on a kinematically admissible strain field generated by a rigid-body displacement.

Bearing in mind the key feature of the work theorem, which is the complete independence of the statically admissible stress field $\boldsymbol{\sigma}$ and the kinematically admissible strain field $\boldsymbol{\varepsilon}^*$, the strain field can be conveniently defined by a rigid-body displacement that is

$$\mathbf{u}^*(\mathbf{x}) = \mathbf{c}, \quad (4)$$

where \mathbf{c} is a constant vector that conveniently leads to null strains that is

$$\boldsymbol{\varepsilon}^*(\mathbf{x}) = \mathbf{0}. \quad (5)$$

When considered the kinematically admissible strain field generated by arbitrary rigid-body displacement, Eq. (4), the local form of the work theorem, Eq. (3), leads to the next expression

$$\int_{\Gamma_Q - \Gamma_{Q'}} \mathbf{t} \, d\Gamma + \int_{\Gamma_{Q'}} \bar{\mathbf{t}} \, d\Gamma + \int_{\Omega_Q} \mathbf{b} \, d\Omega = 0 \quad (6)$$

which states an integral form of mechanical equilibrium, of tractions and body forces, in the local domain $\Omega_Q \cup \Gamma_Q$, are represented in Fig. 2.

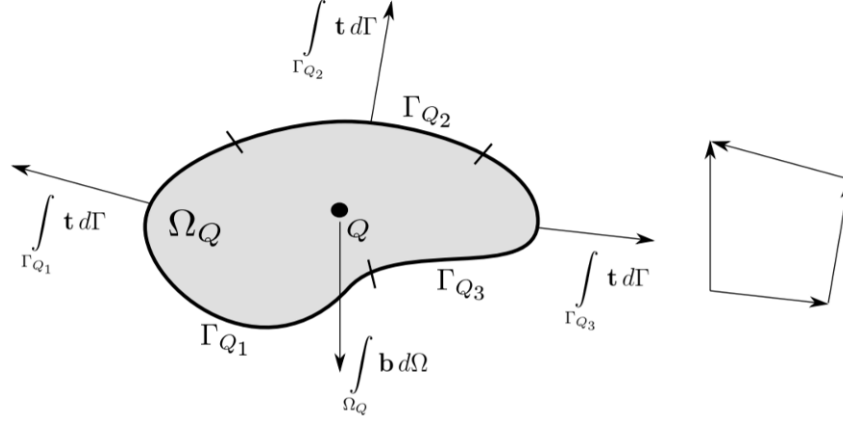


Figure 2. Schematic representation of the equilibrium of tractions and body forces, pointwisely defined at collocation points of a local form of the work theorem associated with a field node Q .

The modelling strategy adopted in this paper, is based in the application of the work theorem, in the set of kinematically admissible strain fields, to solve the actual elastic problem. Consider the local form of the work theorem, Eq. (3). To derive the equilibrium equations of the numerical model, the kinematic formulation of the local form is carried out through the specification of an appropriate kinematically admissible strain field ε^* . This paper considers the arbitrary rigid-body displacement formulation that leads to the local form of equilibrium Eq. (6), in which are then used to generate the stiffness matrix of the meshfree numerical model.

The statically admissible stress field σ , which is required to satisfy equilibrium with a system of external forces, is assumed as the stress field that settles in the body, when it is loaded by the actual system of external distributed surface and body forces, with the actual displacement constraints.

Recall that the elastic field that settles in the body is the only totally admissible elastic field that satisfies the given problem. Therefore, besides satisfying static admissibility, through Eq. (1) and (2), that is the same as satisfying equilibrium through Eq. (6), generated by the weak form Eq. (3) of the work theorem, this unique totally admissible elastic field also satisfies kinematic admissibility defined as

$$\varepsilon = \mathbf{L}\mathbf{u}, \quad (7)$$

in the Defining the Statically Admissible Stress Field domain Ω , with boundary conditions

$$\mathbf{u} = \bar{\mathbf{u}}, \quad (8)$$

on the kinematic boundary Γ_u , in which the displacement \mathbf{u} is assumed continuous with small derivatives, to allow for geometrical linearity of the strain field ε . Hence, Eq. (8), which specifies the constraints of the actual unique solution of the elastic problem must be fulfilled.

For the sake of simplicity, this paper considers the formulation of the meshfree numerical methods in the absence of body forces. Consequently, the equations of equilibrium are always defined only on the boundary of the local domain.

The essential feature of meshfree numerical methods is that they perform the discretization of the problem domain and boundaries with a set of scattered field nodes that do not require any mesh for the approximation of the field variables. The meshfree method IMFM, presented in this paper, is based on the moving least-squares (MLS) approximation, introduced by [18].

The MLS approximation is considered to be one of the best methods to approximate data with a good accuracy. Circular or rectangular local supports centered at each nodal point can be used. In the region of a sampling point X , the domain of definition of MLS approximation is the subdomain Ω_x , where the approximation is defined, as showed in the Fig. 3

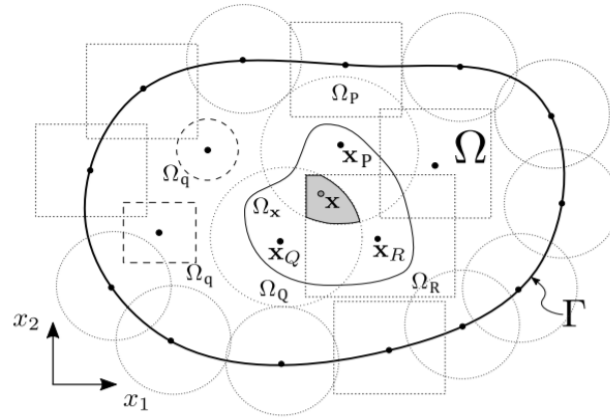


Figure 3. Representation of a global domain Ω and boundary Γ in the meshless discretization, with X_i nodes distributed within the body.

In general, this local domain is a circular or rectangular region, centered at the respective node, where the rigid-body displacement formulation of the work theorem is defined as a local form of mechanical equilibrium.

The local character of the MLS approximation is a consequence of the compact support of each node, where the respective MLS shape functions are defined. The size of the compact support, in turn, sets out, in a neighborhood of a sampling point, the respective domain of MLS approximation at this point. The domain of definition contains all the nodes whose MLS shape functions do not vanish at this sampling point. Therefore, the domain of influence of each node, is the union of the MLS domains of definition of all points in the local domain of the node.

Finally, local meshfree formulations use a node-by-node stiffness calculation to generate, in the domain of influence of the local node, the respective rows of the global stiffness matrix.

In the absence of body forces, the local form of the work theorem with the rigid-body displacement, Eq. (6), can be written simply as

$$\int_{\Gamma_Q - \Gamma_{Q_i}} \mathbf{t} \, d\Gamma = - \int_{\Gamma_{Q_i}} \bar{\mathbf{t}} \, d\Gamma \quad (9)$$

which represents mechanical equilibrium of boundary tractions of the local domain Ω_Q , associated with the field node $Q \in \Omega_Q$.

General numerical methods can be effectively formulated through a reduced integration of the equilibrium Eq. (9) which, in the simplest linear case, leads to a point-wise discrete, form that improves the accuracy and the computational efficiency, as numerical results clearly demonstrate.

Assuming a variation linear of the tractions along each boundary segment of the local domain, the local integral form of equilibrium can be evaluated with a single quadrature point, centered on each segment of the boundary. Applying this linear integration process in the Eq. (9), the following expression is obtained

$$\frac{L_i}{n_i} \sum_{l=1}^{n_i} t_{x_l} = -\frac{L_t}{n_t} \sum_{j=1}^{n_t} \bar{t}_{x_j} \quad (10)$$

In which n_i and n_t denote the total number of the integration points, one per segment, defined on the interior local boundary $\Gamma_{Qi} = \Gamma_Q - \Gamma_{Qt} - \Gamma_{Qu}$, with length L_i , and the local static boundary Γ_{Qt} , with length L_t . This integrated equation represents a point-wise discrete form of the mechanical equilibrium of boundary tractions, evaluated at a set of points on the boundary of the local domain Ω_Q , as showed in the Fig. 4.

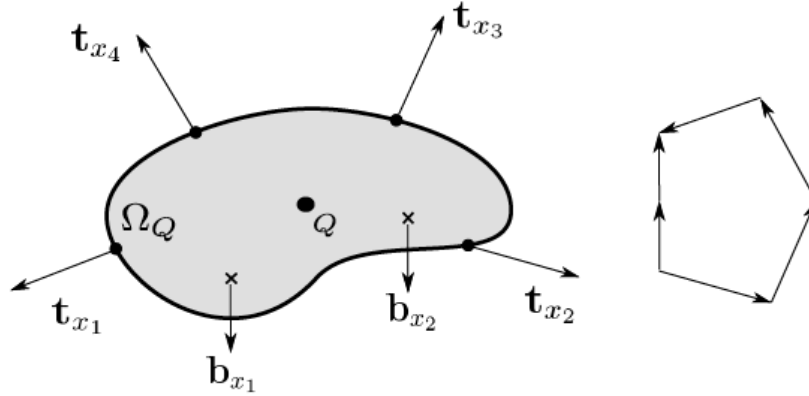


Figure 4. Schematic representation of the equilibrium of tractions and body forces, pointwisely defined at collocation points of a local domain associated with a field node Q .

Consider a meshfree discretization of the body. The local meshless method with reduce integration is used to compute the respective system of algebraic equations, in node by node process, throughout traction evaluation at each central point of the boundary segments of the integrated local form assigned to each node with rectangular local domain. Figure 5 represents these local domains with four boundary segments and one point on each segment.

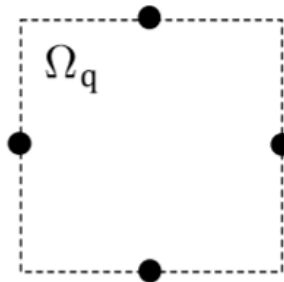


Figure 5. Schematic representation of rectangular local domain with one point on each side.

Discretization of Eq. (10) is carried out with the MLS approximation, for the local domain Ω_Q , in terms of the nodal unknowns $\hat{\mathbf{u}}$, thus leading to the system of two linear algebraic equations

$$\frac{L_t}{n_i} \sum_{l=1}^{n_i} \mathbf{n}_{x_l} \mathbf{D} \mathbf{B}_{x_l} \hat{\mathbf{u}} = - \frac{L_t}{n_t} \sum_{j=1}^{n_t} \bar{\mathbf{t}}_{x_j} \quad (11)$$

that can be written as

$$\mathbf{K}_Q \hat{\mathbf{u}} = \mathbf{F}_Q \quad (12)$$

in which \mathbf{K}_Q , the nodal stiffness matrix associated with the local Ω_Q , is a $2 \times 2n$ matrix given by

$$\mathbf{K}_Q = \frac{L_t}{n_i} \sum_{l=1}^{n_i} \mathbf{n}_{x_l} \mathbf{D} \mathbf{B}_{x_l} \quad (13)$$

and \mathbf{F}_Q is the respective force vector given by

$$\mathbf{F}_Q = - \frac{L_t}{n_t} \sum_{j=1}^{n_t} \bar{\mathbf{t}}_{x_j}. \quad (14)$$

Consider that the problem has a total of N field nodes Q , each one associated with the respective local region Ω_Q . Assembling Eq. (15), for all M interior and static – boundary field nodes leads to the global system of $2M \times 2N$ equations

$$\mathbf{K} \hat{\mathbf{u}} = \mathbf{F}. \quad (15)$$

Finally, the remaining equations are obtained from the $N - M$ boundary field nodes on the kinematic boundary. For a field node on the kinematic boundary, a direct interpolation method is used to impose the Kinematic boundary condition as

$$u_k(x_j) = \sum_{i=1}^n \phi_i(x_j) \hat{u}_{ik} = \bar{\mathbf{u}}_k, \quad (16)$$

Or, in matrix form as

$$\mathbf{u}_k = \Phi_k \hat{\mathbf{u}} = \bar{\mathbf{u}}_k, \quad (17)$$

with $k = 1, 2$, where $\bar{\mathbf{u}}_k$ is specified nodal displacement component. Equations (17) are directly assembled into the global system of equations (15).

It can be easily anticipated high computational efficiency, with very accurate results, of this local formulation with linear reduced integration. As a matter of fact, the nodal stiffness matrix is effectively computed, in Eq. (13), with only 4 integration points (1 integration point on each side of the local boundary), which basically implies a very short processing time to run the analysis. In addition, the reduced integration leads to high accuracy of the results, which plays a key role in the behavior of IMFM, since it implies a reduction of the nodal stiffness which, in turn, leads to an increase of the solution accuracy and, which is most important, presents no instabilities. For more information about the formulation for these method, see [14].

Parameters of the Meshfree Discretization

This section presents some numerical results for Cantilever beam and the Plate with a circular hole for different nodal configurations. The effects of the size of local support and quadrature domain are analyzed and compared with exact solution.

For a generic node i , the size of the local support Ω_s and the local domain of integration Ω_q are respectively given by

$$r\Omega_s = \alpha_s d_i, \quad (18)$$

$$r\Omega_q = \alpha_q d_i, \quad (19)$$

in which d_i represents the distance of the node i , to the nearest neighboring node; for the analysis is performed for two different values of the local support domain size ($\alpha_s = 2.00$ and $\alpha_s = 2.11$), and the local quadrature domain size (α_q) which vary 0.45 to 0.55 with increments 0.01.

Irregular nodal arrangement

The nodal irregularity is generated by changing randomly the coordinates of the nodal regularity distribution by small distance, this movement can be calculated by

$$x_{1i}' = x_{1i} \pm c_n d_{x_{1i}}, \quad (20)$$

$$x_{2i}' = x_{2i} \pm c_n d_{x_{2i}}, \quad (21)$$

in which C_n is a parameter that controls the nodal irregularity and vary randomly in the range of 0.0 and 0.4. For nodes located in the boundary there are restrictions that depend on the position of the node.

Numerical Examples

Displacement and energy norms can be used for error estimation and can be computed, respectively as

$$\|u\| = \left[\int_{\Omega} u^T u d\Omega \right]^{1/2} \quad (22)$$

$$\|\varepsilon\| = \left[\frac{1}{2} \int_{\Omega} \varepsilon^T D \varepsilon d\Omega \right]^{1/2} \quad (23)$$

The relative error for $\|u\|$ and $\|\varepsilon\|$ is given, respectively by

$$r_u = \frac{\|u_{num} - u_{exact}\|}{\|u_{exact}\|} \quad (24)$$

$$r_\varepsilon = \frac{\|\varepsilon_{num} - \varepsilon_{exact}\|}{\|\varepsilon_{exact}\|} \quad (25)$$

Cantilever Beam

A Cantilever beam showed in Fig. 6, is subjected to a parabolic traction at the free end. The main properties are tabulated in Table 1 and the problem is solved for plane stress case.

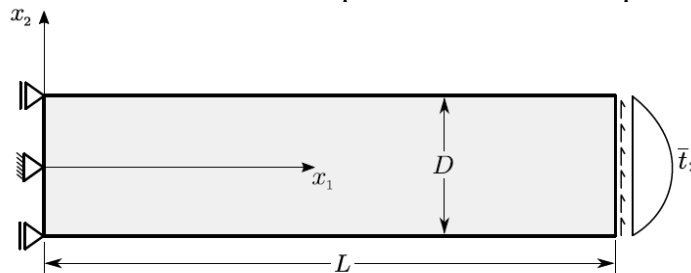


Figure 6. Cantilever beam

Table 1. Properties of Cantilever Beam

Parameters	Values
Height, D	12 m
Length, L	48 m
Thickness, t	1 m
Load, P	1000 N
Modulus of Elasticity, E	30 MPa
Poisson's Ratio, ν	0.3

The parabolic traction and the moment of inertia is given by

$$\bar{t}_2(x_2) = -\frac{P}{2I} \left(\frac{D^2}{4} - x_2^2 \right), \quad (26)$$

$$I = \frac{D^3}{12} \quad (27)$$

The exact solution of the problem is given by [19]. The equations for the exact displacement are:

$$u_1(x_1, x_2) = -\frac{Px_2}{6EI} \left[(6L - 3x_1) + (2 + \nu) \left(x_2^2 - \frac{D^2}{4} \right) \right] \quad (28)$$

$$u_2(x_1, x_2) = \frac{P}{6EI} \left[3\nu x_2^2 (L - x_1) + (4 + 5\nu) \frac{D^2 x_1}{4} + (3L - x_1) x_1^2 \right] \quad (29)$$

And the exact stress components are given by

$$\sigma_{11}(x_1, x_2) = -\frac{P(L - x_1)x_2}{I} \quad (30)$$

$$\sigma_{12}(x_1, x_2) = -\frac{P}{2I} \left(\frac{D^2}{4} - x_2^2 \right) \quad (31)$$

$$\sigma_{22}(x_1, x_2) = 0 \quad (32)$$

The IMFM is used for solving this problem, both a regular ($c_n = 0.0$) and irregular ($c_n = 0.1, 0.2, 0.3$ and 0.4) nodal distribution are employed with a discretization of $21 \times 9 = 189$ nodes.

In the first discretization the nodes located in the boundary have a regular distribution and nodes located inside of the beam have an irregular distribution; it is called Configuration A as is showed in Fig. 7 (presented by [20]) and Fig. 8.

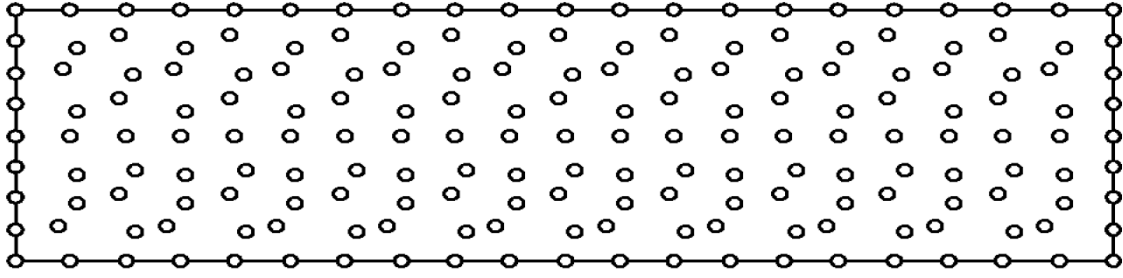


Figure 7. Irregular internal nodal arrangement for the Cantilever beam (MLPG [20]) – Configuration A

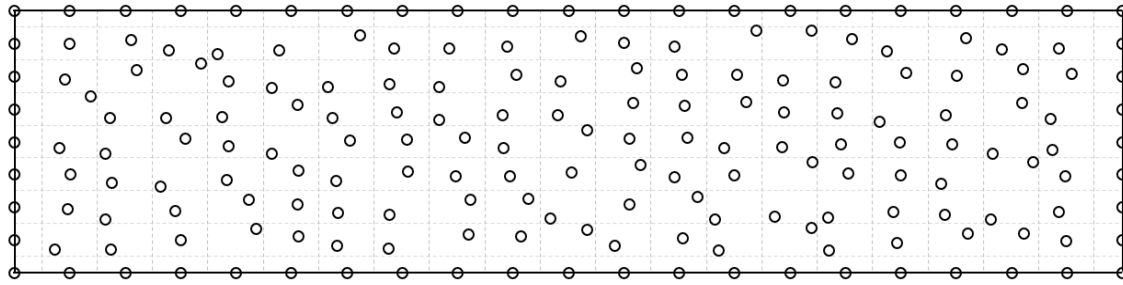


Figure 8. Irregular internal nodal arrangement for the Cantilever beam (IMFM) – Configuration A

Second discretization shows all nodes with an irregular distribution called Configuration B as is showed in Fig 9.

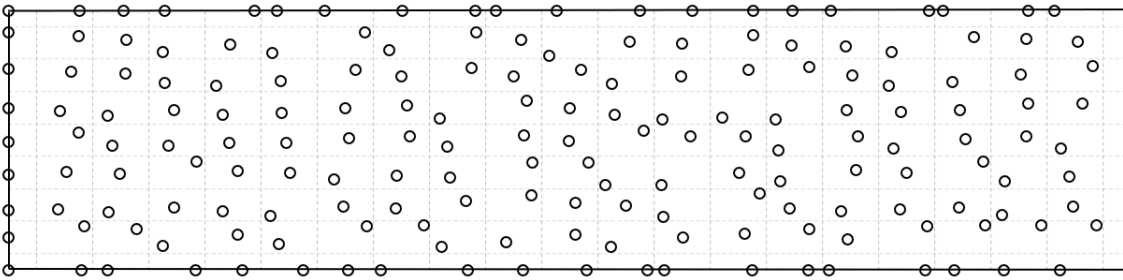


Figure 9. Irregular nodal arrangement for the Cantilever beam (IMFM) – Configuration B

Influence of the local support domain size (α_s)

This parameter must be greater than 1.0, the algorithm of MLS approximation may be singular and the shape function cannot be constructed for the small values. Initially, the influence of α_s in the solution is obtained for the local quadrature domain fixed ($\alpha_q = 0.5$), this value presented an optimal result for regular nodal distribution by [14].

Figure 10 shows the variation of relative error as a function of the size of the parameter of nodal irregularity which vary between 0.0 to 0.4 with 0.1 increments. Results are presented for two values of local support domain ($\alpha_s = 2.00$ and $\alpha_s = 2.11$) using both nodal configuration (A and B).

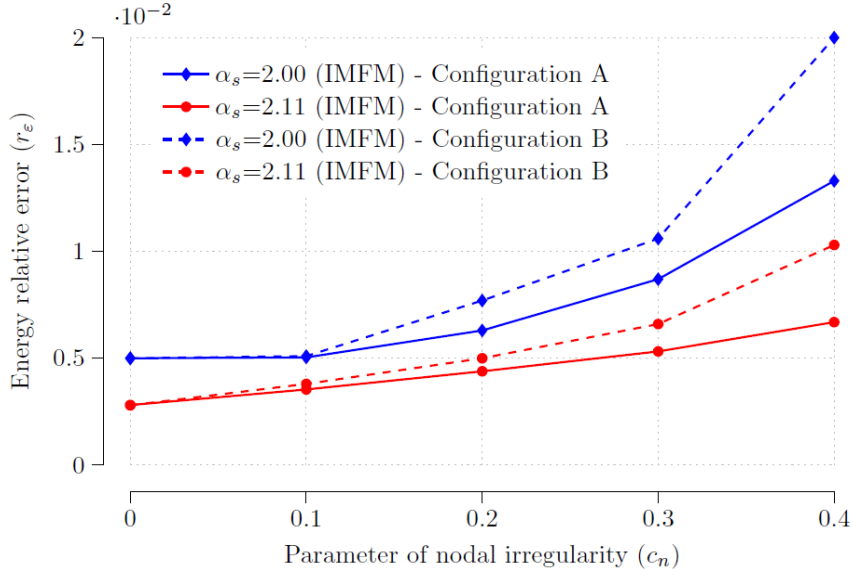


Figure 10. Effect of irregularity on the nodal arrangement on energy relative error with $\alpha_q = 0.5$, carried out with $21 \times 9 = 189$ nodes.

Figure 10 shows that the value of 2.11 for the local support domain (α_s) presents low relative energy errors for model performed with both configurations. Additionally, the magnitude of the energy relative error increase with the nodal irregularity, presenting a directly proportional behavior between the two variables.

The energy relative error for two different methods is presented in Fig. 11, the same irregular nodal discretization called Configuration A is showed in the Fig. 7 for MLPG and the Fig. 8 for IMFM. The energy relative error for MLPG and IMFM have the same expression presented in Eq. 26 and 28. These values for MLPG were obtained in [20].

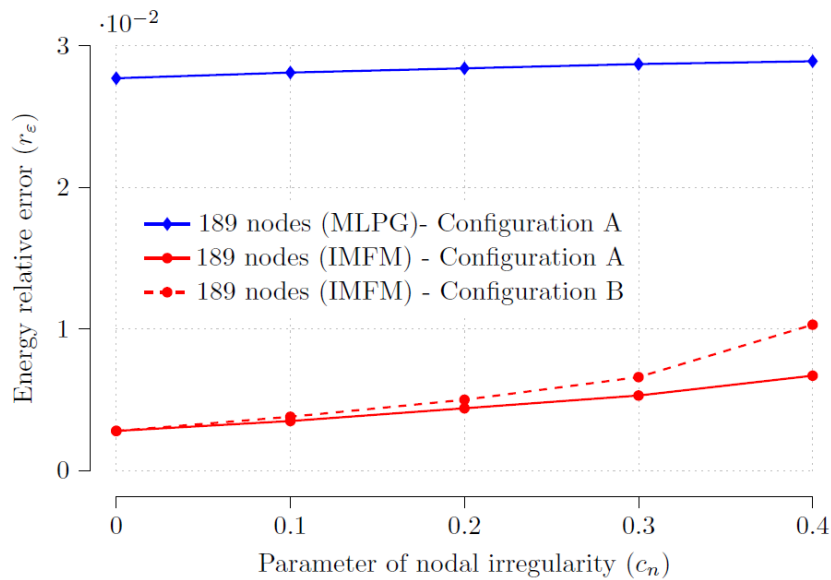


Figure 11. Effect of irregularity on the nodal arrangement on energy relative error for MLPG and IMFM, carried out with $21 \times 9 = 189$ nodes.

For the irregular nodal discretization IMFM and MLPG presented similar behavior for the configuration A, the IMFM presented more accuracy.

Influence of the local quadrature domain size (α_q)

This parameter must be less than 1.0, the reason is to ensure that the local sub – domains of the internal nodes are entirely within the solution domain, without being intersected by the global boundary. The influence of α_q is obtained for the local support domain fixed ($\alpha_s = 2.11$).

Figure 12 shows the variation of relative error as a function of the size of the local support domain which vary between 0.45 to 0.55 with 0.01 increments. Results are presented for four values of parameter of nodal irregularity which vary between 0.0 to 0.4 with 0.1 increments; using the nodal configuration A.

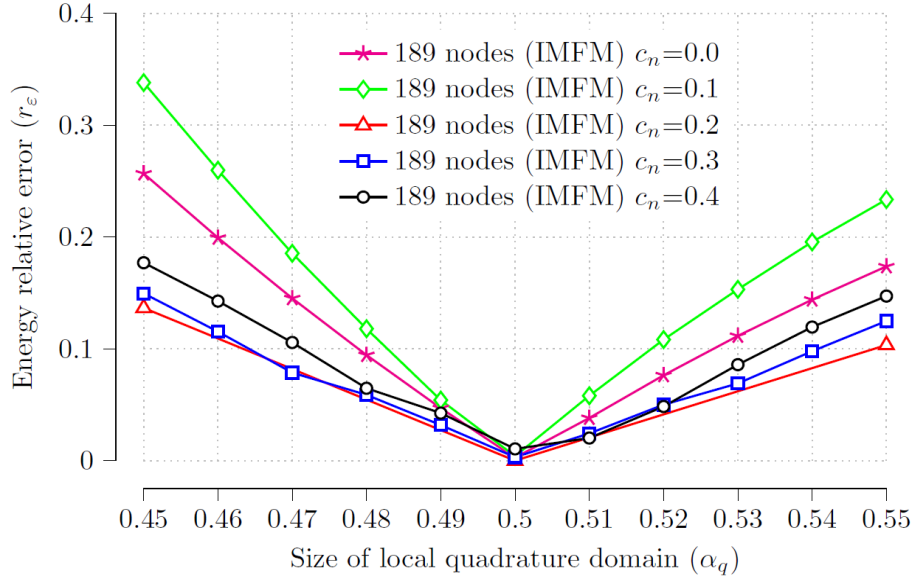


Figure 12. Effect of irregularity on the nodal arrangement on energy relative error with $\alpha_s = 2.11$, carried out with $21 \times 9 = 189$ nodes.

Figure 12 shows that the value of 0.5 for the local quadrature domain (α_q) presents low relative energy errors for all values of parameter of nodal irregularity. Similar results are obtained for displacement relative error.

Influence nodal discretization

Other different irregular nodal discretization was carried out to know the influence on the accuracy on the energy and displacement relative error. Two additional configurations were presented with $11 \times 5 = 55$ nodes and $33 \times 17 = 561$ nodes.

Figure 13 and 14 show respectively, the variation of energy and displacement relative error as a function of the size of the parameter of nodal irregularity which vary between 0.0 to 0.4 with 0.1 increments. Results are presented for three nodal discretization using values fixed of the local support domain ($\alpha_s = 2.11$) and the local quadrature domain ($\alpha_q = 0.5$), using both nodal configuration (A and B).

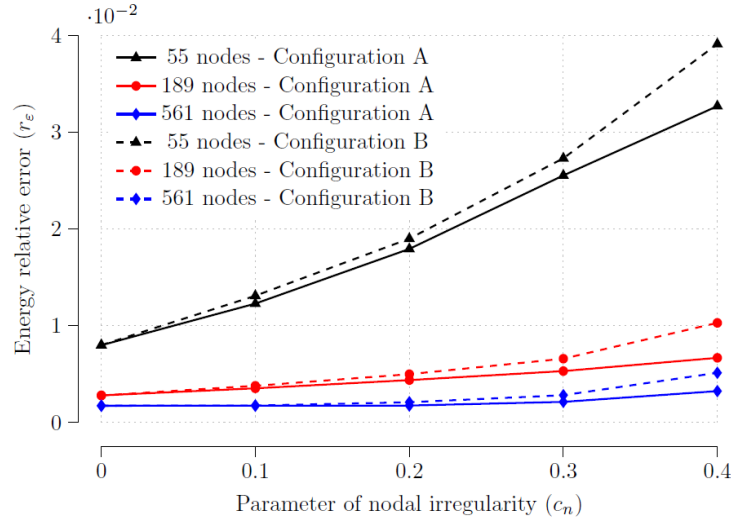


Figure 13. Effect of irregularity on the nodal arrangement on energy relative error with $\alpha_s = 2.11$ and $\alpha_q = 0.5$, carried out with $11 \times 5 = 55$ nodes, $21 \times 9 = 189$ nodes and $33 \times 17 = 561$ nodes.

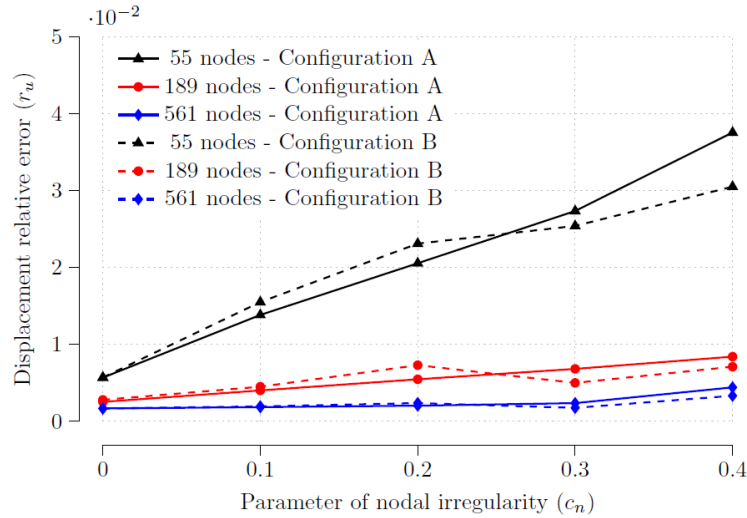


Figure 14. Effect of irregularity on the nodal arrangement on displacement relative error with $\alpha_s = 2.11$ and $\alpha_q = 0.5$, carried out with $11 \times 5 = 55$ nodes, $21 \times 9 = 189$ nodes and $33 \times 17 = 561$ nodes.

The Figure 13 and 14 shows that the energy and displacement relative error decreases with finer nodal distributions for both configuration. This result evidences that the meshless methods with reduce integration has the potential to be a very method including nodal arrangement with greater irregularity.

Conclusions

The effect of the nodal irregularity is very little on energy and displacement relative errors for different configuration and discretization. This fact reveals that the meshless methods with reduce integration (IMFM) and meshless Local Petrov Galerkin (MLPG) are stable for irregular nodal arrangements, but the IMFM presented more accuracy.

The nodal irregularity and the energy relative error presented a directly proportional relation for the cantilever beam. A similar behavior is showed for different mesh discretization, using a fixed value $\alpha_q = 0.5$ on the local quadrature domain.

The local quadrature domain ($\alpha_q = 0.5$) for regular and irregular nodal arrangement, different mesh discretization and any value of the local support domain presented the most accurate results for the cantilever beam.

The discretization with the nodes located in the boundary have a regular distribution and nodes located inside of the beam have an irregular distribution presents lower energy and displacement relative error that the discretization which all nodes have an irregular distribution.

Acknowledgments

The program PECC – *Pós-Graduação em Estruturas e Construção Civil*, Department of Civil and Environmental Engineering, Faculty of Technology, University of Brasília and CNPq – *Brazilian National Counsel of Technological and Scientific Development* for his PhD scholarship.

References

- [1] Daxini, S. D. and Prajapati, J. M. (2014). A review on recent contribution of the meshfree methods to structure and fracture mechanics applications, *the scientific word journal* **2014**, 13 pages.
- [2] Finalyson, B. A. (1972). The Method of Weighted Residuals and Vibrational Principles. Academic Press.
- [3] Chen, J. S., Hillman, M. and Chi, S. W. (2017) Meshfree Methods: Progress made after 20 years, *Journal of Engineering Mechanics* **143**, 4.
- [4] Nayroles, B., Touzot, G. and Villon, P. (1992). Generalized the Finite Element Method: Diffuse Approximation and Diffuse Elements, *Computational Mechanics* **10**, 307–318.
- [5] Li, W. K., Jun, S., and Zhang, Y.F. (2007). Reproducing Kernel Particle Methods, *International Journal for Numerical methods in Engineering* **20**, 1081–1106.
- [6] Belytshko, T., Lu, Y. Y. and Gu, L. (1994). Element – free Galerkin methods, *International Journal for Numerical Methods in Engineering* **37(2)**, 229–256.
- [7] Atluri, S. N. and Zhu, T. (1998). A new Meshless Local Petrov-Galerkin (MLPG) approach in computational mechanics, *Computational Mechanics* **22(2)**, 117–127.
- [8] Atluri, S. N. and Shen, S. (2002) The Meshless Local Petrov – Galerkin (MLPG) Method: A simple and Less-costly Alternative to the Finite Element and Boundary Element Methods, *CMES: Computer Modeling in Engineering and Sciences* **3(1)**, 11-51.
- [9] Zhu, T., Zhang, J. and Atluri, S. N. (1998). A Local Boundary Integral Equation (LBIE) Method in Computational Mechanics and a Meshless Discretization Approach, *Computational Mechanics* **21**, 223–235.
- [10] Liu, G. R. and Gu, Y. T. (2001). A Local Point Interpolation Method for Stress Analysis of Two-Dimensional Solids, *Structural Engineering and Mechanics* **11(2)**, 221–236.
- [11] Liu, G. R., Yan, L., Wang, J. G. and Gu, Y. T. (2002). Point Interpolation Method Based on Local Residual Formulation Using Radial Basis Functions, *Structural Engineering and Mechanics* **14**, 713–732.
- [12] Atluri, S. N. and Han, Z. D. and Rajendran, A. M. (2004). A New Implementation of the Meshless Finite Volume Method Through the MLPG Mixed Approach, *CMES: Computer Modeling in Engineering and Sciences* **6**, 491-513.
- [13] Oliveira, T. and Portela, A. (2016). Weak – Form Collocation – a Local Meshless Method in Linear Elasticity. *Engineering Analysis with Boundary Elements* **73**, 144 – 160.
- [14] Santana, E., Oliveira, T., Vélez, W. and Portela, A. (2017). Meshfree Method with Reduced Integration and Automatic Parameter Optimization, *Computers and Structures* (submitted).
- [15] Fredholm, I. (1906) Solution d'un problème fondamental de la théorie de l'élasticité. *Arkiv for Matematik Astronomi och Fysik* **28(1)**, 1-8

- [16] Fichera, G. (2006). Linear Elliptic Differential Systems and Eigenvalue Problems. Springer.
- [17] Brebbia C. A. and Tottenham, H. (1985). Variational Basis of Approximate Models in Continuum Mechanics. Southampton and Springer Verlag.
- [18] Atluri, S. N. and Zhu, T. (2000). New Concepts in Meshless Methods, *International Journal for Numerical Methods in Engineering* **6**, 537-556.
- [19] Timoshenko, S. P. and Goodier, J. N. (1970). *Theory of Elasticity*, 3rd edn, McGraw-Hill, New York, USA.
- [20] Liu, G. R. (2003) Meshfree methods moving beyond the finite element method. CRC Press, 1st edition.

Analysis of the results obtained from the application of the two-stage method with calculations of some statically indeterminate trusses

***Janusz Rębielak¹**

¹Chair of Structures and Construction Engineering, Faculty of Architecture, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland.

*Presenting author: j.rebielak@wp.pl

Abstract

The paper presents results of calculations of forces in members of selected types of statically indeterminate trusses carried out by application of the two-stage method of computations of such structural systems. The method makes possible to do the simple and approximate calculations of the complex trusses in two stages, in each of which is calculated a statically determinate truss being an appropriate counterpart of the basic form of the statically indeterminate truss structure. Systems of the statically determinate trusses considered in the both stages are defined by cancelation of members, number of which is equal to the statically indeterminacy of the basic truss. In the paper are presented outcomes obtained in the two-stage method applied for two different shapes of trusses and carried out for various ways of removing of appropriate members from the basic trusses. The results are compared with outcomes gained due to application of a suitable computer software for computation of the same types of trusses and for the same structural conditions.

Keywords: Truss, Calculus of vectors, Superposition method, Statically indeterminate system, Cremona's method, Approximate solution.

Introduction

Values of forces acting in members of the statically indeterminate systems are computed by the application of various methods like for instance, the force method, the displacement method, the iteration methods like the method of successive approximations, and the finite elements method, etc. Procedures of these methods are nowadays adapted in numerous and various types of computer software [1]-[7]. Concept of computational procedure used in the two-stage method is described in previous papers [8]-[10] and it is in simplified way shown in Fig. 1.

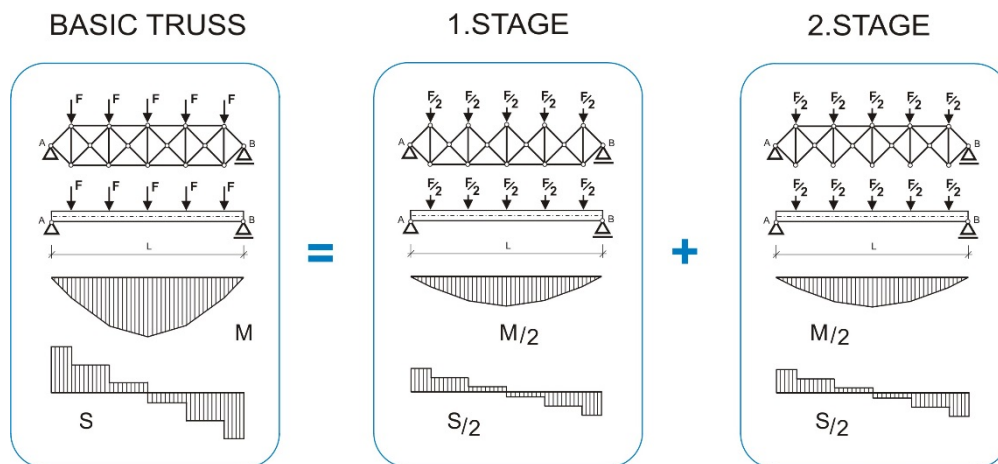


Figure 1. Concept of two-stage method of calculation of statically indeterminate trusses

The point of the two-stage method is to define shapes of the statically determinate trusses calculated in the intermediate stages, which shapes are appropriately compatible with the basic system. It is done after cancellation suitable members from the geometry of the basic truss, number of which equals the degree of static indeterminacy of the basic truss system. Then in each stage of this method it is calculated the statically determinate truss loaded by forces having half values of the load forces applied to the basic truss. Final values of forces acting in a member of the basic truss is the resultant of forces calculated in both stages acting in the members having appropriate positions in the calculated trusses. Because in the two-stage method there are not taken into account differences between stiffness of members joined in the same node, that is why it is an approximate method of calculation of the statically indeterminate trusses.

The two-stage method uses in general the principles of geometry of strains as well as the static equilibrium regarding reactions and the generalized internal forces. These principles are based on the background of statics of the rigid body therefore they can be applied for solutions of statically determinate systems. Methods used for calculations of statically indeterminate systems apply concepts of virtual work [11]-[14] what causes, that they give the exact values of forces really acting in members of such complex structural systems.

Subject of static calculations and analyses

Although the two-stage method gives the approximate results but on basis of conclusions of previous research one can state that its accuracy is good enough for the engineering practice. In the initial analyses there have been considered simple forms of trusses similar to the geometry schemes shown in Fig. 1. These trusses have been mostly loaded by concentrated forces applied in symmetric way along their lengths only to the top chord or only to the bottom chord. Moreover in the first stage members of the top chord were canceled, while in the second stage members of the bottom chord were rejected. Subject of the current research refers to the same truss geometry but the investigated trusses are loaded in way presented in Fig. 1a and in Fig. 1d. Moreover it is assumed, that in each stage are canceled members at the same time from the top chord and from the bottom chord. There are defined following types of structural configurations of the calculated trusses.

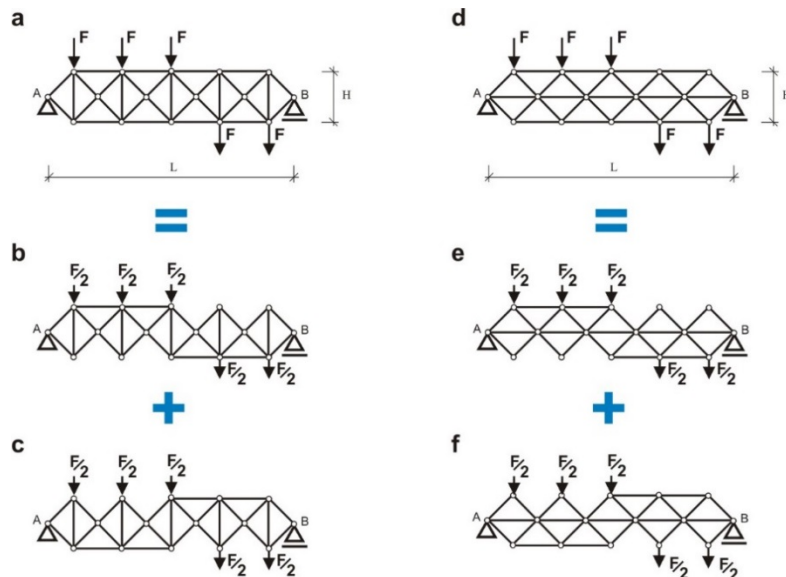


Figure 2. Basic static schemes of two groups of calculated trusses

If the truss has some vertical members and in each stage two members are excluded from the top and bottom chord then this structural configuration, shown in Fig. 2b and in Fig. 2c, is marked by symbol $V \frac{2}{2}$. When the truss has in its chords horizontal members and like previously two members of external chords are deleted in each stage then that type of structural configuration is marked by symbol $H \frac{2}{2}$, compare Fig. 2e and Fig. 2f. In this

research there are considered also other subtypes of such configurations, in which three members of the top chord are excluded and only one member of the bottom chord is rejected in suitable stages. These structural configurations are marked respectively by symbols V 3/1 and H 3/1. Various types of considered static schemes of the calculated trusses are assumed purposely in order to recognize more precisely the features of two-stage method and exactness of obtained results. Values of forces calculated in the two-stage method for the presented truss systems are compared with outcomes gained by application of a computer software for the same static and structural conditions for all the calculated truss systems. Basic structures are of 5.00 m clear span and their construction depth is equal to 1.00 m.

The both compared types of truss systems consists of the same number of members and nodes. Number of nodes is defined by symbol “w”, while symbol “p” defines number of members. The condition for the inner statically determinacy of the plane truss is defined as follows:

$$p = 2 \cdot w - 3 \quad (1)$$

The considered truss systems are built by number of nodes $w = 16$, what implies that the statically determinate truss created by means of this number of nodes has to be constructed by the following number of members:

$$29 = 2 \cdot 16 - 3 \quad (2)$$

The basic truss systems are in each case built by the number of members $p = 33$, what means that the calculated structures are the fourfold statically indeterminate systems. It implies that in order to create the statically determined system it is necessary to delete 4 appropriate members from area of the basic truss. Static systems of considered trusses calculated according to the rules of the two-stage method in its stages, for the structural configuration marked by symbols V 2/2 and H 2/2, are shown in Fig. 2. Final values of the forces calculated for the both basic system will be resultants of forces defined in each stage for members of suitable positions. The concept of the assumed calculation method is compatible with rules of calculus of vectors, principle of superposition and with the three fundamental conditions of equilibrium presented below:

$$\sum_{i=1}^n F_{ix} = 0 \quad (3)$$

$$\sum_{i=1}^n F_{iy} = 0 \quad (4)$$

$$\sum_{i=1}^n M_i = 0 \quad (5)$$

It is assumed that the both basic trusses are subjected to the same type of load. In this investigation three concentrated forces are applied to the nodes of the top chord located in the close vicinity to the support node A, while remaining two concentrated forces are applied to the bottom chord located closer to the support node B, compare Fig. 2. It is assumed that the unit load forces F have value equal to 1.00 kN.

The same static systems of basic trusses have been subjected to the static calculation carried out by the application of the Autodesk Robot Structural Analysis Professional 2017. The computer software is used for the exact calculation of the force values acting in members of the statically indeterminate systems. Static calculations were made by assumption that the truss consists of steel tubular members having diameter of 30.00 mm, the thickness of the section equals to 4.00 mm and the steel material has the Young's modulus equal to 210 GPa.

Values of forces calculated for structural configurations V 2/2 and V 3/1

The truss system with vertical members is the subject of the first group of static calculations carried out by means of the two-stage method for two selected types of rejection of members, which was shortly discussed above. In the first case two members of are appropriately deleted from the top and from the bottom chords of the basic truss system. Results obtained in the first stage of calculations, together with Cremona's polygon of forces, are presented in Fig. 3.

Outcomes of second stage of computations are shown in Fig. 4. Final values of forces calculated for the truss configuration V 2/2 are presented in Fig. 7a.

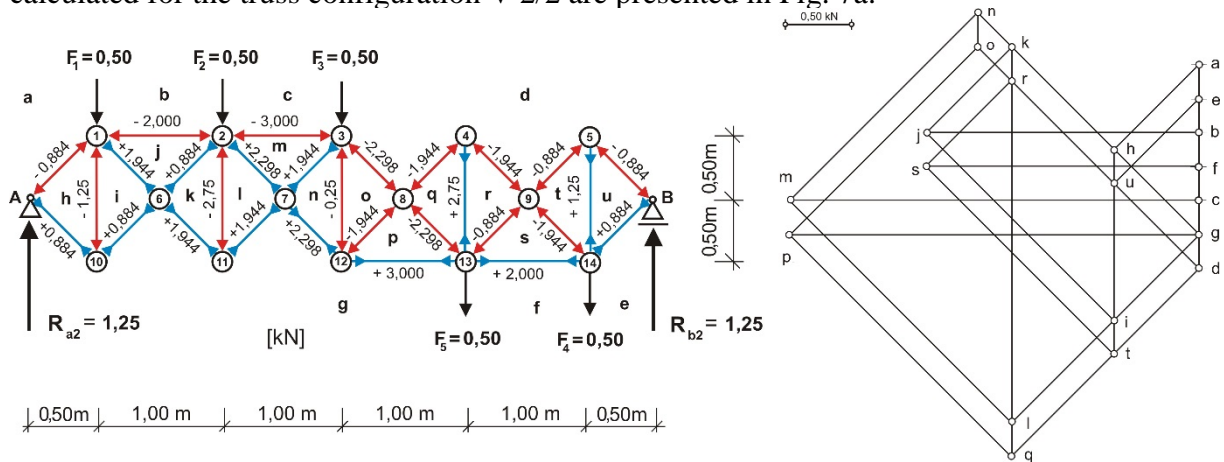


Figure 3. Values of forces determined in the first stage of calculations for the truss type configuration V 2/2 together with Cremona's polygon of forces

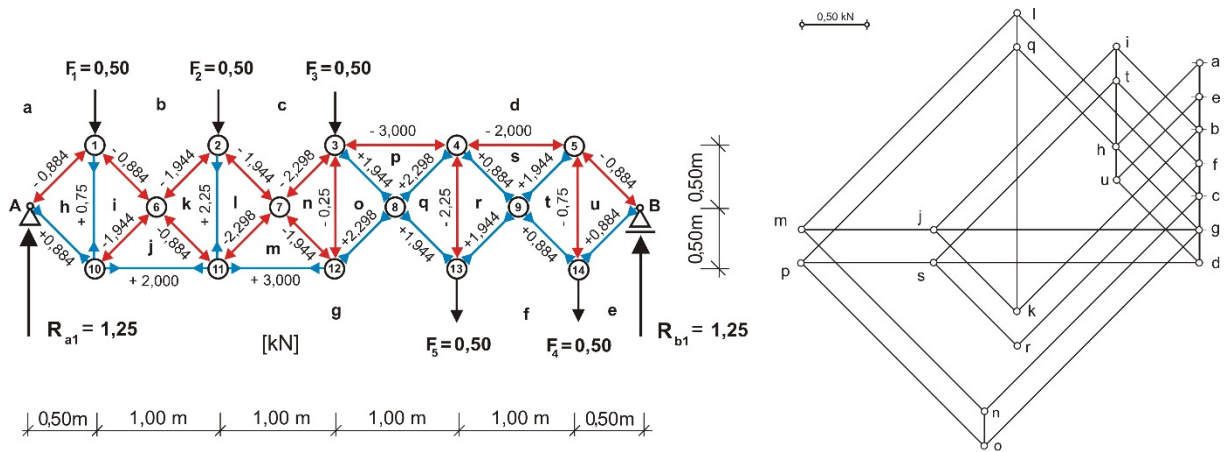


Figure 4. Values of forces defined in the second stage of calculations for the truss type configuration V 2/2 together with Cremona's polygon of forces

The same shape of truss system is now calculated in the two-stage method for structural configuration denoted by symbol V 3/1. Results of such calculations are shown in Fig. 5 and in Fig. 6. The final force values computed for this configuration are also presented in Fig. 7a.

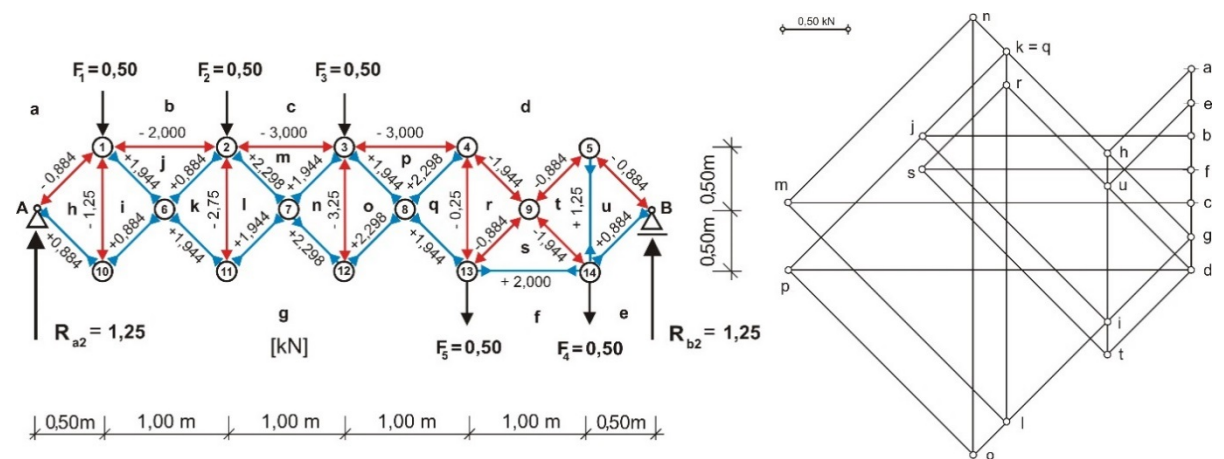


Figure 5. Values of forces determined in the first stage of calculations for the truss type configuration V 3/1 together with Cremona's polygon of forces

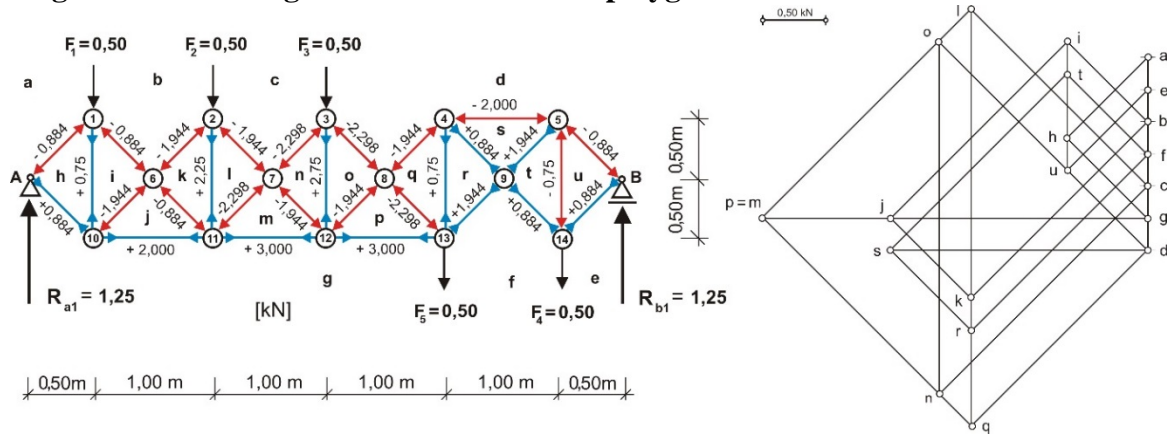


Figure 6. Values of forces determined in the second stage of calculations for the truss type configuration V 3/1 together with Cremona's polygon of forces

After application of computation procedures appropriate to the two-stage method values of forces defined in the same members of the basic statically indeterminate truss are of the same values, what does not depend on the considered types of configurations. It implies that they are exactly equal, when in process of calculation are removed 2 corresponding members from external chords, type configuration V 2/2, and when from the top chord are deleted 3 members and 1 member from the bottom chord, which configuration is marked by symbol V 3/1. Truss of the same static scheme, compare Fig. 1a, has been calculated by application of the computer software Autodesk Robot Structural Analysis Professional 2017. Outcomes of the computer calculations are presented in Fig. 7b.

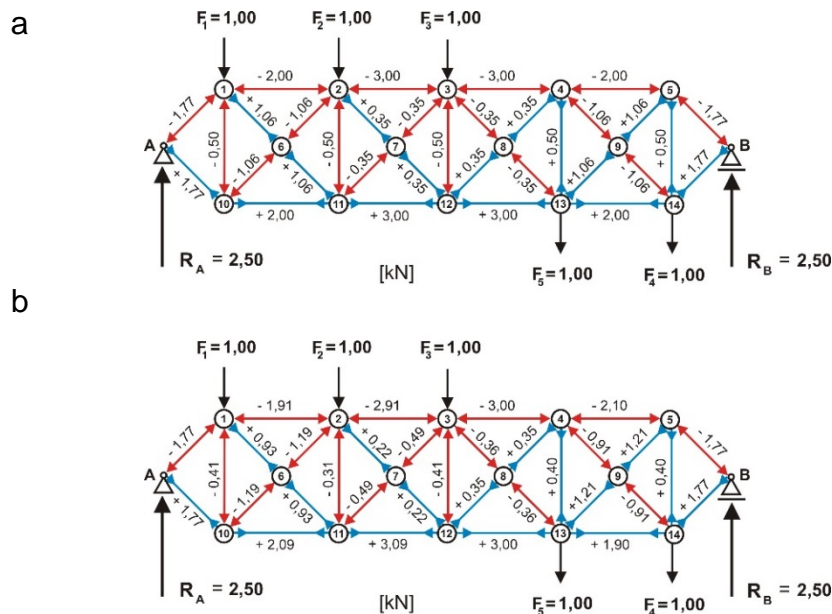


Figure 7. Values of forces in members for the truss type configuration V 2/2 together calculated, a) in the two-stage method, b) by means of computer software

Differentiations of the force values calculated in the both compared methods in members of external chords are rather small. For example the compression force defined in the two-stage method in member situated between nodes 4 and 5 equals -2.00 kN, while by application of the computer software it is equal to -2.10 kN, which constitutes about 5 % of the biggest value. Bigger differentiations one can notice between force values defined in vertical members and in the cross braces. For instance the force value calculated in the two-stage

method in vertical member located between nodes 2 and 11 equals -0.50 kN, while by the computer software it is equal to -0.31 kN. Similar bigger differentiations one can notice only between values of forces, calculated in the both methods, however having the very small absolute values.

Values of forces calculated for structural configurations H 2/2 and H 3/1

The second group of calculations has been carried out by application of the two-stage method for the statically indeterminate truss having static system shown in Fig. 2d. In Fig. 8 and in Fig. 9 are presented results gained for the truss schemes, where in both the stages are removed two members from external chords of the basic truss, truss configuration denoted H 2/2. Values of forces defined for configuration marked by symbol H 3/1 are shown in Fig. 10 and in Fig. 11.

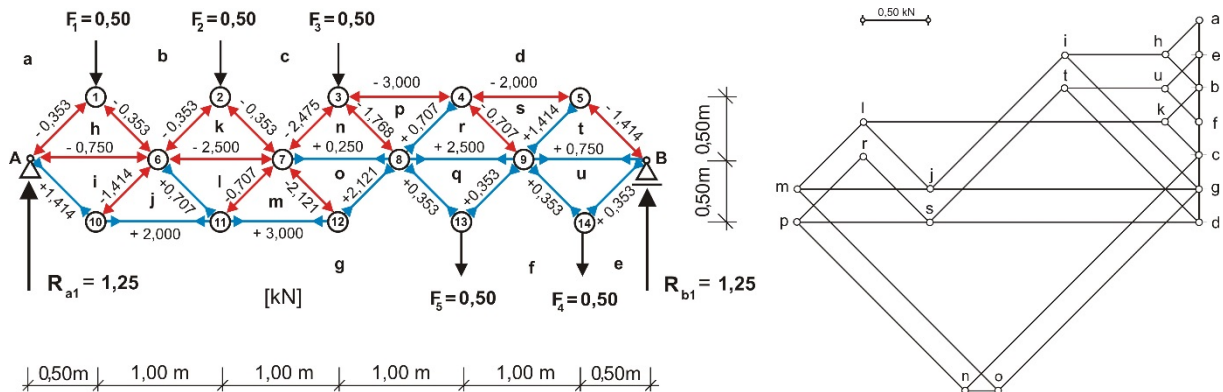


Figure 8. Values of forces determined in the first stage of calculations of the truss type configuration H 2/2 together with Cremona's polygon of forces

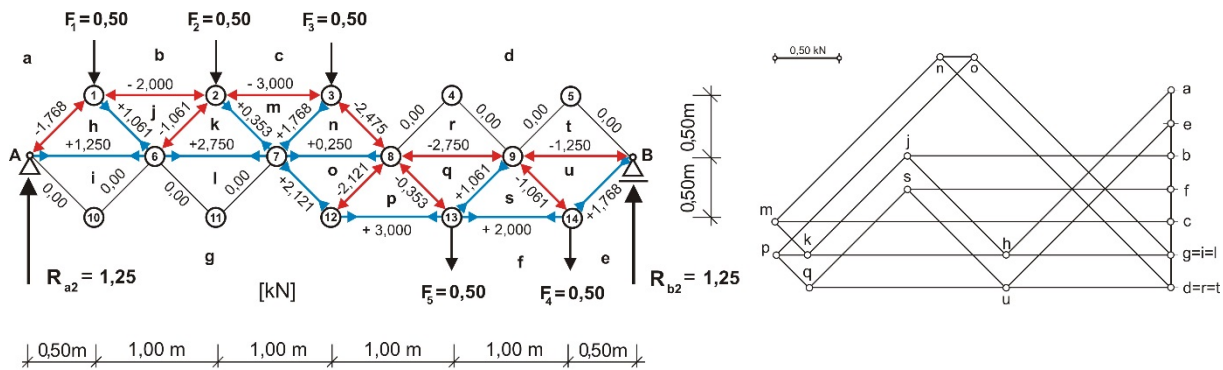


Figure 9. Values of forces determined in the second stage of calculations of the truss type configuration H 2/2 together with Cremona's polygon of forces

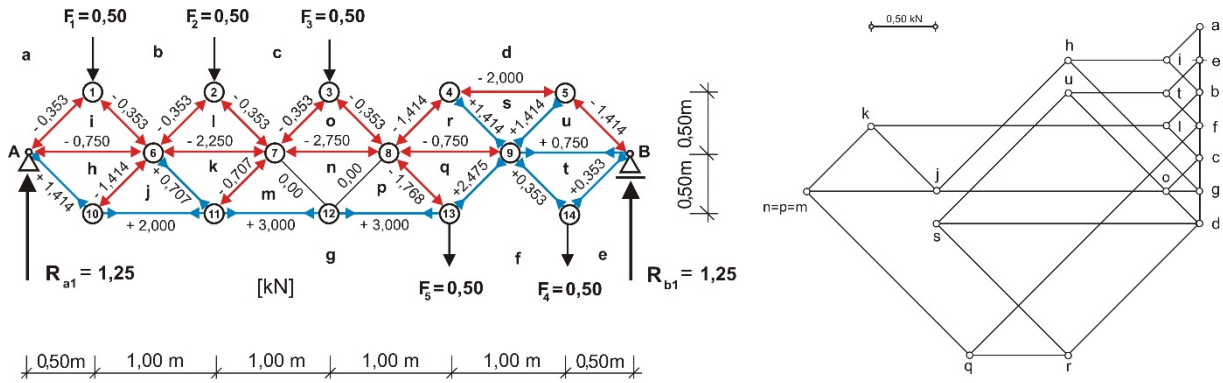


Figure 10. Values of forces determined in the first stage of calculations of the truss type configuration H 3/1 together with Cremona's polygon of forces

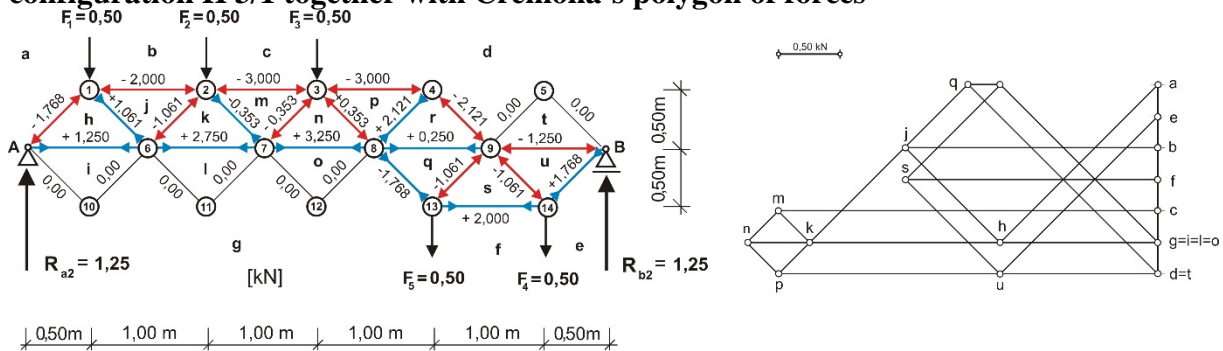


Figure 11. Values of forces determined in the second stage of calculations of the truss type configuration H 3/1 together with Cremona's polygon of forces

All final results of the calculated statically indeterminate truss are presented in Fig. 12. Values of forces defined in members of the basic truss by means of the two-stage method using procedure of cancelation of two members from each of the external chords, type of configuration H 2/2, are shown in Fig. 12a. Results obtained in this method for the procedure of excluding three appropriate members from the top chord and cancelation of single member from the bottom chord, type of configuration H 3/1, are shown in Fig. 12b. Values of forces determined by application of suitable computer software are presented in Fig. 12c.

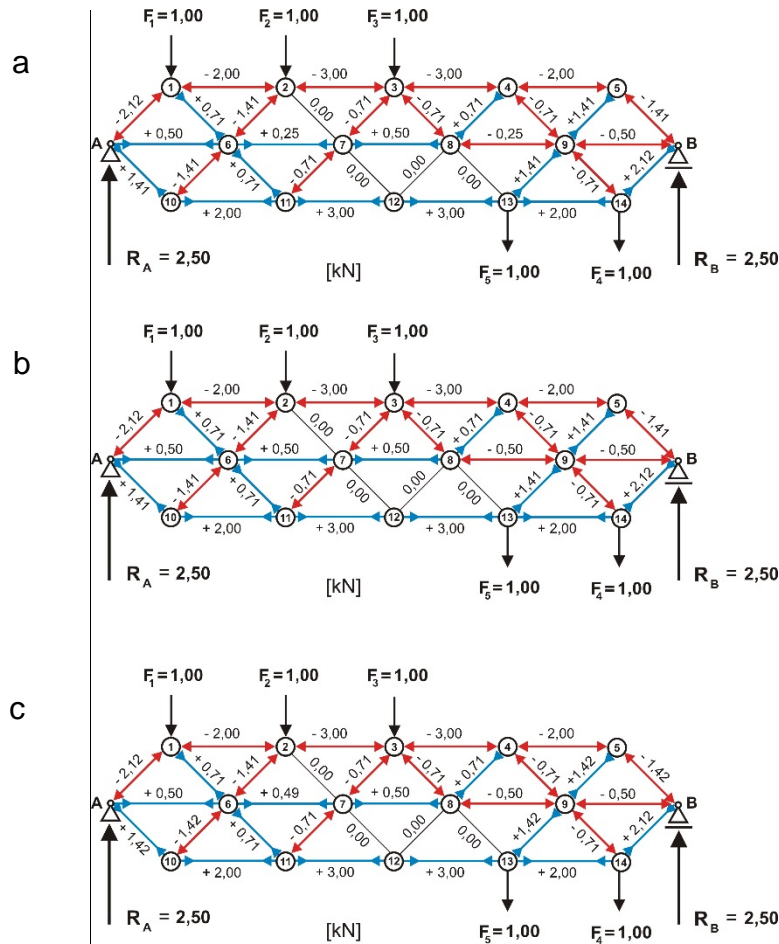


Figure 12. Values of forces defined in members of the H truss system by application of two-stage method, a) for the type configuration H 2/2, b) for the type configuration H 3/1, c) calculated by application of computer software

From analysis of all results obtained in compared calculations carried out in the two-stage method for various types of rejection of members from the outer chords follows, that values of forces defined in members of these chords and in cross braces of the basic truss are mostly of the same or of nearly the same values, compare Fig. 12a and Fig. 12b. Moreover they are also mostly of the same values of forces, which are calculated by application of the Autodesk Robot Structural Analysis Professional 2017, see Fig. 12c. Slightly bigger differentiation one can notice in values of forces defined in members of the middle chord. For instance value of force determined in member located between nodes 6 and 7 by application of the two-stage method equals +0.25 kN for the configuration H 2/2, see Fig. 12a, while in the same member for the configuration H 3/1 it is equal to +0.50 kN, see Fig. 12b. In general one can state, that results obtained in the two-stage method for the configuration H 3/1 are very similar to values of forces calculated in the basic statically truss by help of the computer software.

Results of all calculations, presented above, can testify the usefulness of the two-stage method for calculations of statically indeterminate trusses. It belongs to a group of recently developed methods of approximate solutions of such systems [15] invented on basis of various types of basic principles.

Conclusions

Values of forces obtained by application of the two-stage method for the calculation of the statically indeterminate trusses are approximated but mostly very similar to the force values determined by use of the exact methods applied in computer software. Bigger differences one can notice in values of very small forces, but they can be considered of small importance in the engineering practice. Accuracy of these results in a very small degree depends of the way

of the necessary cancellation of selected members from area of the basic truss in order to create the intermediate forms of statically determinate systems computed in the both stages. Other characteristics of the two-stage method will be subjects of the next research.

References

- [1] Timoshenko, S.P. (1966) *History of strength of materials*, Arkady, Warszawa, - in Polish
- [2] Makowski, Z.S. (1981) *Analysis, design and construction of double-layer grids*, Applied Science Publishers, London.
- [3] Zienkiewicz, O.C. and Taylor R.L. (2000) *The finite element method*, Oxford Press, UK.
- [4] Allen E., Zalewski W. and Boston Structures Group (2010) *Form and forces. Designing efficient, expressive structures*, John Wiley & Sons, Hoboken, New Jersey.
- [5] Kolendowicz, T. (1993) *Theory of structures for architects*, Arkady, Warszawa, - in Polish
- [6] Hibbeler, R.C. (1994) *Mechanics of materials*, Prentice Hall.
- [7] Hibbeler, R.C. (1995) *Structural analysis*, Prentice Hall.
- [8] Rębielak, J. (2014) A two-stage method for an approximate calculation of statically indeterminate trusses, *Journal of Civil Engineering and Architecture*, **78**, 567-572.
- [9] Rębielak, J. (2016) Simple method of approximate calculation of statically indeterminate trusses, *Proceedings of The 7th International Conference on Computational Methods (ICCM2016)*, August, 1-4, 2016, University of California, Berkeley, USA, 748-753.
- [10] Rębielak, J. (2017) Examples of applications of two-stage method in calculations of statically indeterminate trusses, *Proceedings of The 8th International Conference on Computational Methods (ICCM2017)*, July, 25-29, 2017, Guilin, China, 644-649.
- [11] Niezgodziński, M.E. and Niezgodziński, T. (1979) *Strength of materials*, State Scientific Publishing House, Warszawa, - in Polish
- [12] Przewłocki, J. and Górski, J. (2006) *Basis of theory of structures*, Arkady, Warszawa, - in Polish
- [13] Dyląg, Z. and Krzemińska-Niemiec, E. and Filip, F. (1989) *Theory of structures*, State Scientific Publishing House, Warszawa, - in Polish
- [14] Cywiński, Z. (1976) *Theory of structures in problems. Vol.II. The rudiments of statically indeterminate systems*, State Scientific Publishing House, Warszawa - Poznań, - in Polish
- [15] Tan, Z-Q., Jiang, X-D., He, Y-S., Ban, S-H., Xu, R. and Xi R-Q. (2018) Generalized variational principles for solutions of statically indeterminate trusses under mechanical-thermal-assembly loadings, *Journal of Engineering Mechanics*, Volume 144, Issue 1.

Two-stage method applied in calculations for statically indeterminate truss of larger span

***Janusz Rębielak¹**

¹Chair of Structures and Construction Engineering, Faculty of Architecture, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland.

*Presenting author: j.rebielak@wp.pl

Abstract

The paper presents application of two-stage method to calculations of the force values acting in members of selected type of statically indeterminate truss having longer span than trusses considered in previous scientific articles. The calculations are carried out in two stages. The two-stage method applies principles of calculus of vectors as well as rules of superposition. At each stage of this method it is calculated a statically determinate truss system, shape of which is determined by reduction of the appropriate number of members from the basic statically indeterminate truss. The number of deleted members equals the statically indeterminacy of the basic truss. Thus in both stages there are calculated the statically determinate trusses. The final values of forces acting in the members of the basic truss are resultants of forces calculated for the counterpart members in each stage. Basic shape of the statically indeterminate truss has horizontal layer of members located in the middle of construction depth. The basic truss is subjected to symmetrical and nonsymmetrical type of the load. There are discussed differences between results obtained for these two types of the load and between force values calculated in the two-stage method and by application of a suitable computer software.

Keywords: Truss, Statically indeterminate system, Superposition method, Calculus of vectors, Cremona's method, Approximate solution.

Introduction

Truss systems are applied for a long time in the structures of roofs and floors. In comparison to bending beams of the same clear span and the same conditions of load the trusses are more lightweight. Members of the truss are connected together by means of theoretically articulated joints, due to which the members are subjected to act only of axial forces. The assumption that members are connected in articulated nodes is one of the basic requirements of the truss system. This assumption makes possible the application of principles of the calculus of vectors to the methods of calculation of values of forces in statically determinate trusses, like for instance the Cremona's method. Some analytical methods are used for this purpose, Ritter's method is one of them [1]-[4]. More complex trusses are mostly the statically indeterminate systems and they are also more efficient structures than the statically determinate trusses. In both systems the basic structural assumptions remain the same but due to their complexity, however the way of force distribution inside statically indeterminate trusses also depends on the stiffness of members joined in each particular node [5,6]. Therefore the processes for calculation of forces in members of these systems are also more complex. In such cases the distribution of forces in a single node depends, among others, on mutual ratios of the stiffness of members connected in it. There are numerous other factors having influence on force distribution between members especially in statically indeterminate systems, which have to be taken into consideration in various methods of the static calculations [7]-[11]. The degree of difficulty of the calculation procedure increases enormously with the growing number of members and nodes of the structure. It is especially important for the space trusses also called the space structures or the space frames [12]. At present the new numerical methods used into modern software make the calculation processes of complex statically indeterminate trusses very fast and enormously efficient. In certain types

of the structural analysis it is not necessary to take into consideration the exact values of forces acting in the truss members.

Concept of the two-stage method and subject of calculations

The two-stage method was worked out during initial static analyses of a simple tension-strut truss, being initially a statically indeterminate system. If the basic tension-strut structure is overloaded some of its members are excluded from the force transmission and the remaining structure become a statically determinate system. The number of excluded tension members is equal to the statically indeterminacy of the basic truss system. Then the reduced structure can be calculated by one of very simple methods, for instance by help the Cremona's method. The new calculation procedure should keep rules of the calculus of vectors, together with principle of superposition and the below given basic conditions of equilibrium of coplanar force systems:

$$\sum_{i=1}^n F_{ix} = 0 \quad (1)$$

$$\sum_{i=1}^n F_{iy} = 0 \quad (2)$$

$$\sum_{i=1}^n M_i = 0 \quad (3)$$

Calculations of the force values are carried out in two stages. In each stage the considered truss is shaped on pattern of the basic truss by removing of number of members equal to degree of statically indeterminacy of the basic truss. Moreover in each stage the load forces are of the half values and they have to be applied to the suitable nodes. Final values of forces acting in particular members are resultants of forces calculated in the both stages for appropriate members and having corresponding localizations into the truss area.

The point of the two-stage method is described in papers [13,14,15], where are presented results of calculations of simple forms of trusses having relatively very small clear spans. The paper presents outcomes obtained by application of the two-stage method to defining values of forces acting in the statically indeterminate truss of longer clear span than trusses calculated previously. The basic truss system is subjected to symmetrical and nonsymmetrical type of load. The concentrated load forces are applied to suitable nodes of the top chord. Static analysis of the assumed form of the basic truss is undertaken in order to estimate accuracy of the force values calculated in two-stage method for trusses of complex shapes and larger clear spans.

Methods, results of calculations and comparison analyses

Some characteristic results gained by the application of the two-stage method can be closely recognized after analyses of the force values calculated for truss of geometry shown in Fig. 1.

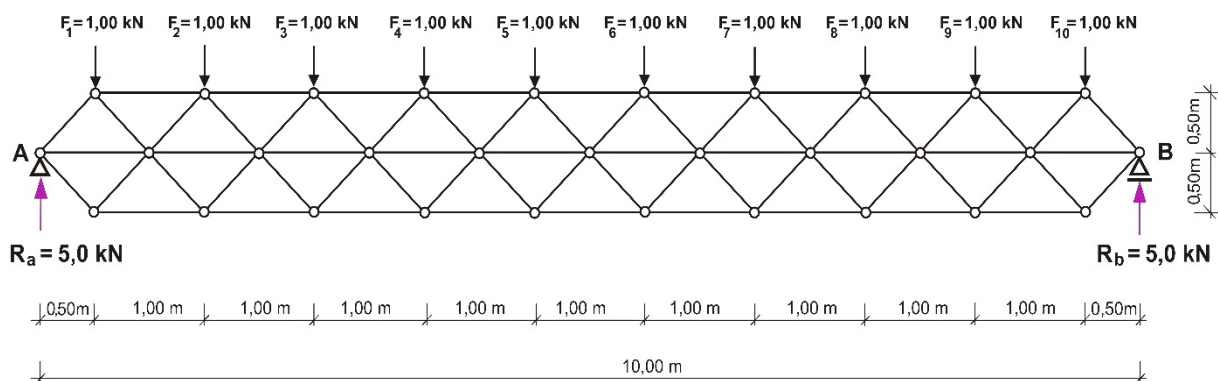


Figure 1. Static scheme of the basic truss subjected to symmetrical load

The basic truss has clear span of 10.00 meters, its construction depth equals 1.00 meter and it is loaded by concentrated forces, each of value equal to 1.00 kN. Here are considered the two ways of the force application. The first one is a symmetric loading, while concentrated forces are applied to each node of the upper chord. The second one is called asymmetric, when these forces will be applied to half the number of the upper chord nodes, only on one side of the truss. The basic truss is calculated by application of the two-stage method and these results are compared with outcomes gained by usage of the software Autodesk Robot Structural Analysis Professional 2017, designed for the precise calculation of force values in the statically indeterminate systems. It has been assumed that the truss consists of steel tubular members having diameter of 30.00 mm, the thickness of the section equals to 4.00 mm and the steel material has the Young's modulus equal to 210 GPa.

Analysis of basic truss loaded in symmetric way

The condition for the inner static determinacy of the plane truss is as follows:

$$p = 2 \cdot w - 3 \quad (4)$$

where symbol “p” defines number of members, while “w” determines number of nodes. The considered shape of the truss system shown in Fig.1 is created by number of nodes $w = 16$, what implies that the statically determinate truss created by number of members $p = 68$, which are connected together by means of number of nodes $w = 31$. A statically determined truss consisting of this number of nodes has to be created by a number of members determined by the equation below:

$$59 = 2 \cdot 31 - 3 \quad (5)$$

It implies that the considered truss is the nine-fold indeterminate system ($68 - 59 = 9$). Therefore in each stage of the two-stage method one should remove nine appropriate members. In the first stage there were deleted nine members from the top chord of the basic truss. Because the truss is of symmetric form and moreover it is symmetrically loaded the calculation process can be limited only to half of the truss. The load forces are of halve values, they equal to 0.50 kN and they are applied to the top chord nodes. The values of forces calculated in the first stage by the application of Cremona's method is presented in Fig. 2.

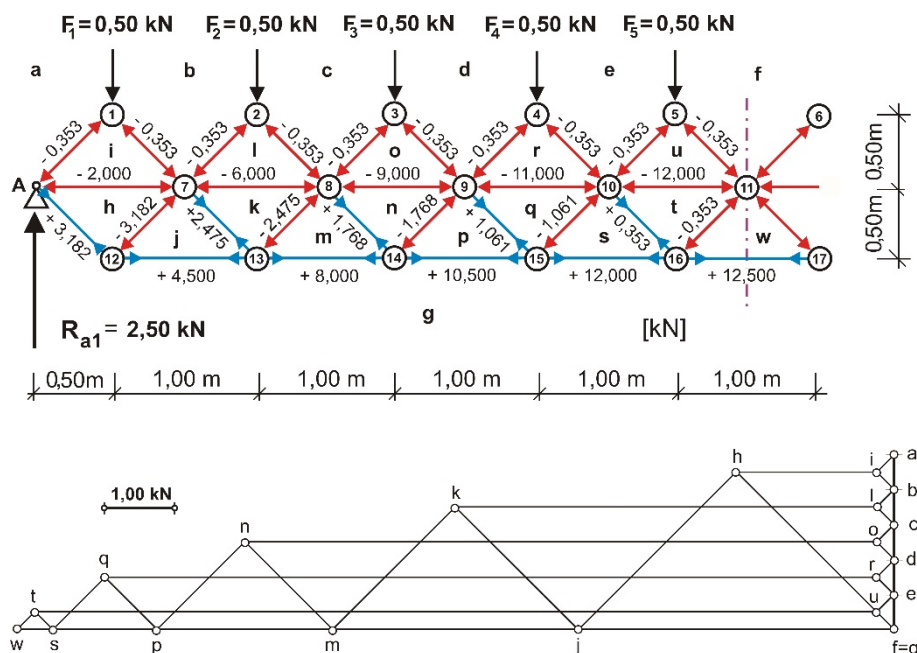
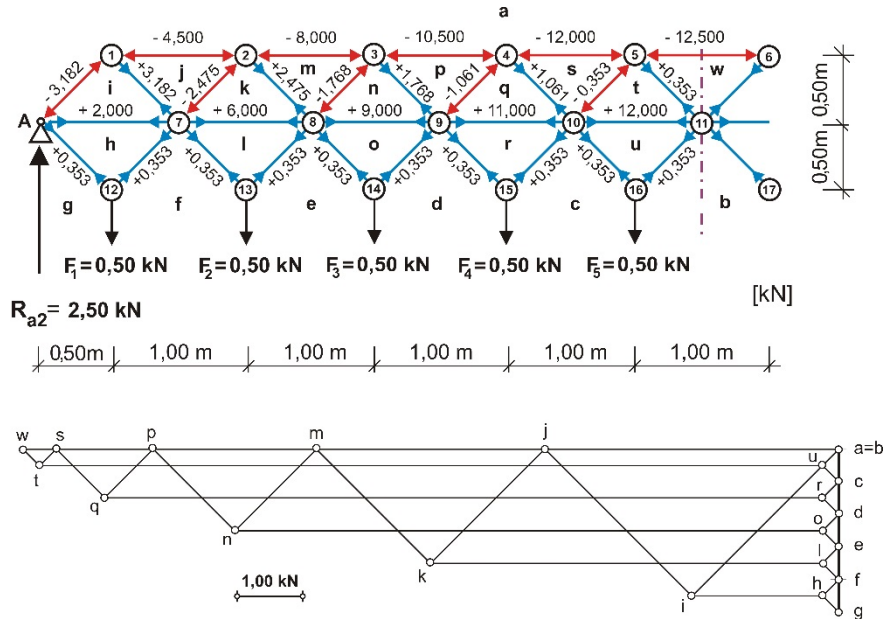


Figure 2. Results calculated in the first stage of calculations for the basic truss loaded in symmetric way together with Cremona's polygon of forces

According to previously explained rules in the second stage of this method one should eliminate nine members from the bottom chord of the basic truss. From analysis of the geometry of this truss and the general conditions of equilibrium it follows, that load forces of values equal 0.50 kN have to be applied to the bottom chord nodes. Scheme of the calculated



truss system together with results gained by means of Cremona's method are shown in Fig. 3.

Figure 3. Results calculated in second stage for basic truss loaded in symmetric way together with Cremona's polygon of forces

Resultant values of forces in members of the truss are calculated by the application of the two-stage method are shown in Fig. 4a. The results calculated for the same truss by the application of Autodesk Robot Structural Analysis Professional 2017 are presented in Fig. 4b.

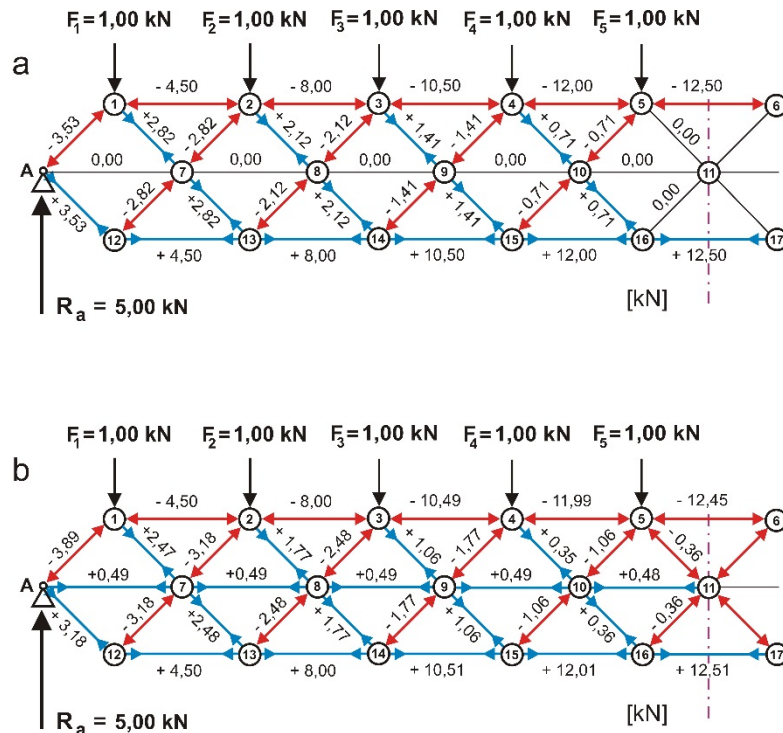


Figure 4. Final values of forces calculated for the symmetrical loaded truss, a) in the two-stage method, b) by application of computer software

Let us compare outcomes gained by usage of various methods for the same members. For instance, the value of the force acting in the member placed between node No 5 and node No 6 calculated in the first stage of the two-stage method equals zero, see Fig. 2, because it has been removed in this stage. In the second stage the value of the compression force is determined as -12.50 kN, which is why the final, resultant force value in this member is equal to -12.50 kN, see Fig. 4a. By application of the computer software the compression force value in member located between nodes No 5 and No 6 is defined as -12.45 kN, see Fig. 4b. The difference in relation to the smaller force value is rather small and it comes to 0.40 %. Bigger differences of the force values are noticed in other members like, for example, in the member connected to the support and placed between the support node A and No 1. In the first stage of calculations the compression force has value of -0.353 kN, see Fig. 2, while in the second stage its value is calculated as equal to -3.182 kN, see Fig. 3. Therefore the final value of the compression force defined in this member by means of the two-stage method equals -3.53 kN, see Fig. 4a. The value of the compression force calculated by the application of Autodesk Robot Structural Analysis Professional 2017 in the same member is equal to -3.89 kN, see Fig. 4b. The real difference between these two values equals 0.36 kN, which gives the relative difference coming to 10.1 %. A bigger relative differentiation can be noticed between values of forces defined in both compared methods for members of the middle layer located e.g. between nodes No 9 and No 10. In the two-stage method the force acting in this member is of zero value, while calculated by means of the computer software it is a tension force of value equals +0.49 kN. From initial analysis of the gained results follows the conclusion, that the biggest relative differences of forces appear in members subjected to act of relatively small absolute values.

Analysis of basic truss loaded in asymmetrically way

The statically indeterminate truss system shown in Fig. 5 has the same geometric and structural parameters like the truss structure presented in Fig. 1, what implies that it is a nine-fold indeterminate system, but it is now loaded in an asymmetric way. In this case the truss is loaded by five concentrated forces, each of 1.00 kN value, applied only to five successive nodes located only on the right site, in close vicinity to the support node B. In the first stage of the two-stage method nine members of the top chord have to be removed from the basic truss in order to make the truss a statically determinate system. The concentrated forces are of half value, equal to 0.50 kN, are applied to the same nodes of the upper chord.

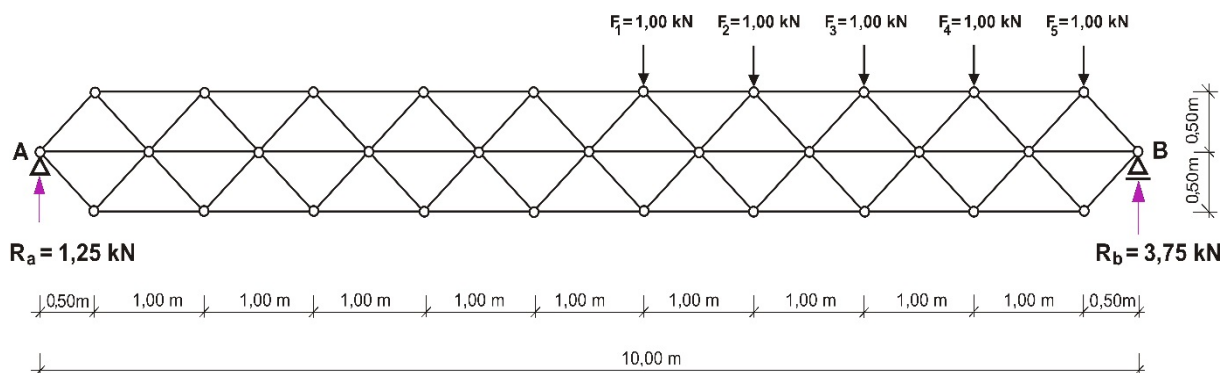


Figure 5 . Static scheme of the basic truss subjected to asymmetric way of load

Results of this stage of the calculations together with a suitable Cremona's polygon of forces are shown in Fig. 6. In the second stage of the two-stage method there are removed also nine members, but this time from the lower chord. Like previously the load forces of half values are applied to the corresponding nodes of the lower chord. All values of forces calculated in the second stage of calculation of the asymmetric loaded truss are presented in Fig. 7. The final results, being resultants of the force values calculated in both stages for the counterpart members, are shown in Fig. 8a. The procedure for determining the force values in particular

truss members is the same as described above. The results defined for the same truss by the application of the computer software Autodesk Robot Structural Analysis Professional 2017 are presented in Fig. 8b.

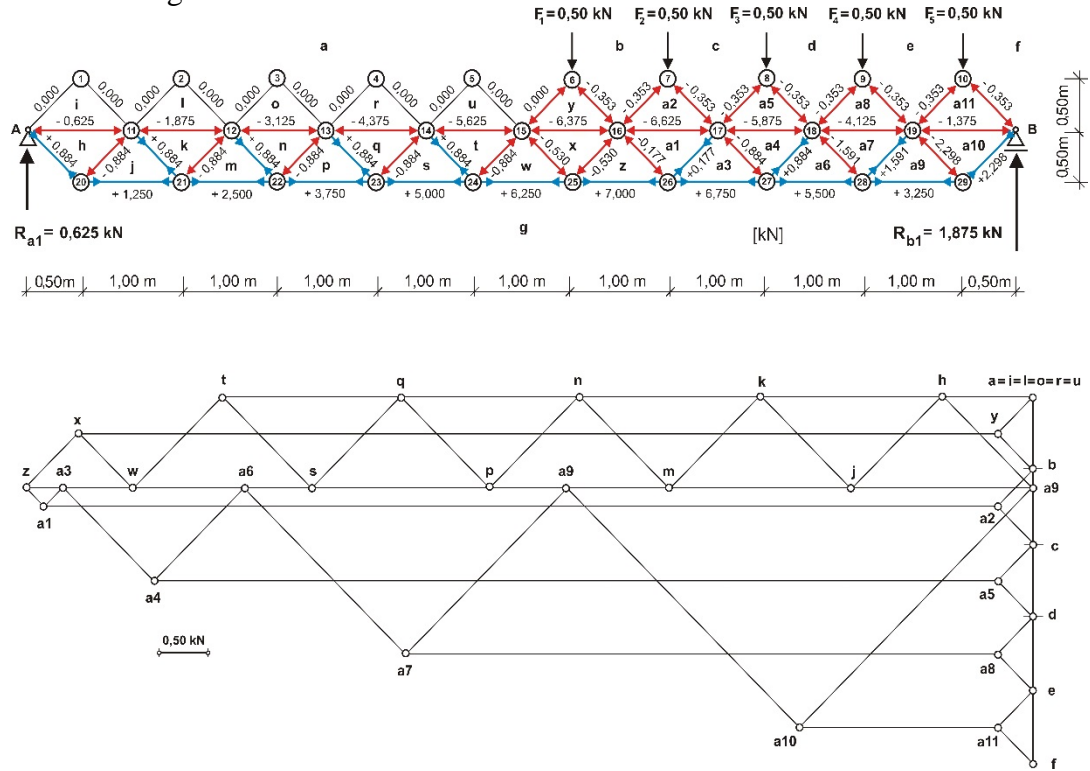


Figure 6. Results calculated in the first stage for the asymmetrically loaded truss together with Cremona's polygon of forces

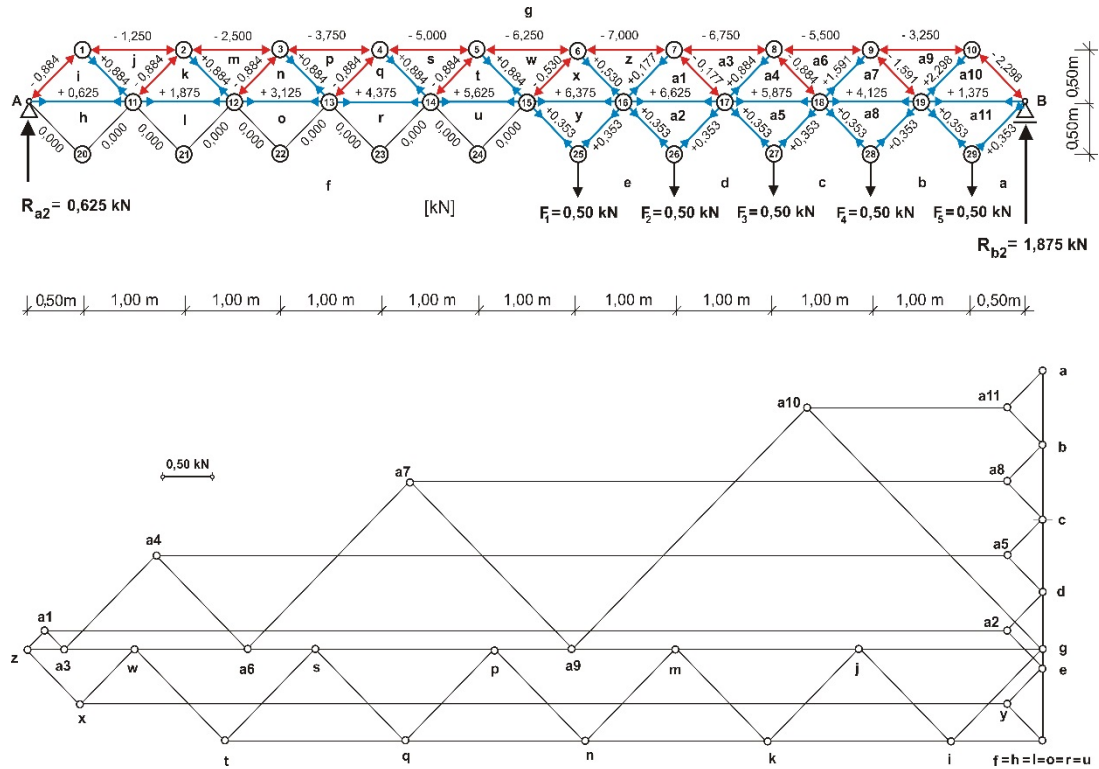


Figure 7. Results calculated in the second stage for the asymmetrically loaded truss together with Cremona's polygon of forces

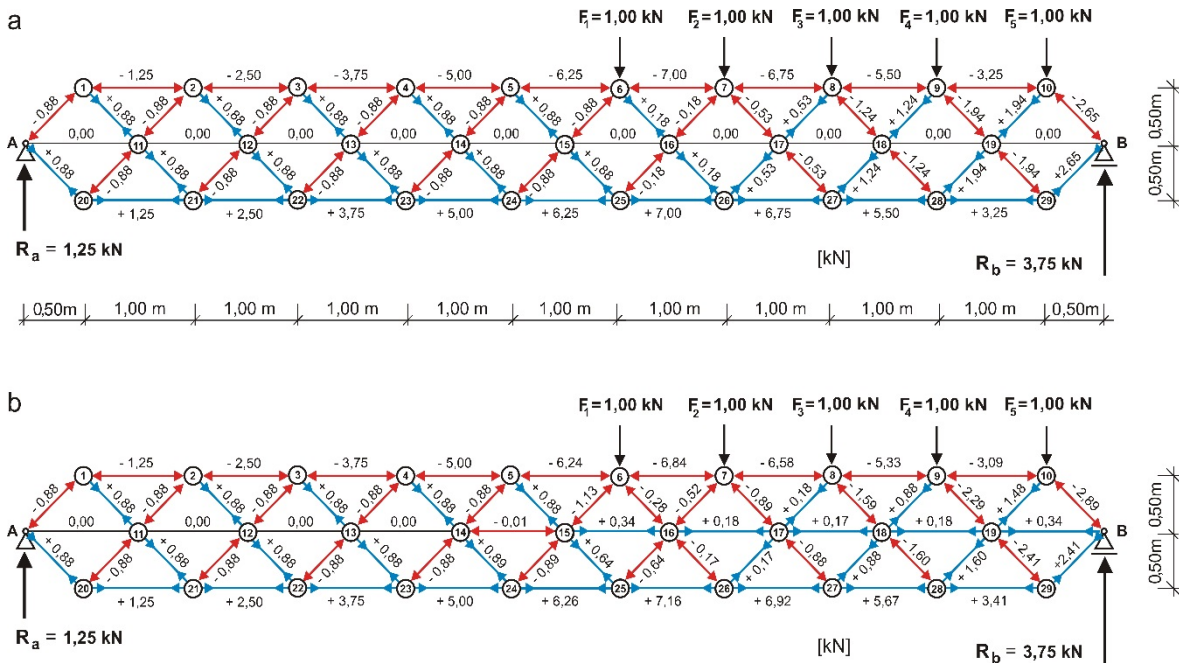


Figure 8. Final values of forces calculated for the asymmetrically loaded truss, a) in the two-stage method, b) by application of computer software

Differentiation of force values in members of the outer chords in the considered truss is rather small. For instance the force value calculated in the two-stage method in compression member of upper chord located between nodes No 6 and No 7 equals -7.00 kN, while by application of the computer software it is equal to -6.84 kN. The difference in amount of 0.16 kN constitutes towards the smaller value only 2.33 % of its basic dimension. One can notice some differences between the calculated force values acting in members of the middle chord spaced between support nodes A and B. Forces calculated in both of the compared methods in such members placed closely to support A are of the same value equals 0.00 kN. One should notice that these members are located quite far from the area, where the applied outer load forces are located. One can observe substantial differences in the members of the middle chord located closely to the support node B in the sector, where there are only the applied load forces. For example the force value calculated in the two stage method in the member placed between nodes No 15 and No 16 equals zero, while by application of the computer software it is estimated that the tensile force has a value of +0.34 kN. The differences can be considered as significantly big, but one has to be aware that they refer to very small values of forces. A similar force differentiation can be noticed in values calculated in cross braces. In the area located in the vicinity of the support A, forces defined in both the compared methods are of identical values. In the zone located close to support B differences of the estimated force values are evident. The tensile force calculated in the member placed between node No 17 and No 8, by means of the two-stage method, is of the value +0.53 kN, while by application of the computer software Autodesk Robot Structural Analysis Professional 2017 it is defined to equal +0.18 kN. However the difference is only equal to 0.35 kN of the absolute value, but the relative difference is of 194% in relation to the smallest force value. In spite of this observation one should be aware that the biggest relative differentiation appears only for the smallest force, close to the zero value.

Conclusions

This paper presents examples of the application of the two-stage method of calculation of statically indeterminate trusses together with a comparison of the results gained by the

application of a commonly used computer software package. The investigated plane truss has two outer chords and a middle chord consisting of horizontal members, while its clear span is ten times bigger than its construction depth. The basic truss is subjected to symmetrical and asymmetrical way of load. The two-stage method is an approximate method of calculation of statically indeterminate trusses because it incorporates processes of calculations, which among others do not take into consideration different stiffness of members connecting in particular nodes. From the comparison of results obtained from both of the compared methods follows that in general the results obtained by means of the two-stage method are very compatible with results gained by application of the exact calculations made by means of Autodesk Robot Structural Analysis Professional 2017. The differentiation of the force values calculated in both methods in the same members are subjected to the biggest forces is relatively small, it is around couple of percent. The differences are considerably big for members, where values of calculated forces are very small, being close to the zero value. One should point out, that members subjected to such small forces are designed by applying suitable rules or requirements of building codes, which implies that areas of their cross-sections are sometimes many times bigger than it directly follows from the calculated values and forces and strength of a material. Due to its simplicity and due the sufficient approximation of obtained results the two-stage method can be applied not only in the initial structural design of the statically indeterminate trusses. It can be adapted for an appropriate and new type computer software what will be not complex because of the simplicity of elementary calculation procedures needed for this purpose. Accuracy of the two-stage method can be significantly improved by application of a set of suitable coefficients, which can define ways of the force distribution according to differentiation of stiffness of members connecting in the same node. One should expect that principles of the method can be relatively easily adapted for the calculation procedures of space trusses.

References

- [1] Allen E., Zalewski W. and Boston Structures Group (2010) *Form and forces. Designing efficient, expressive structures*, John Wiley & Sons, Hoboken, New Jersey.
- [2] Kolendowicz, T. (1993) *Theory of structures for architects*, Arkady, Warszawa, - in Polish
- [3] Hibbeler, R.C. (1994) *Mechanics of materials*, Prentice Hall.
- [4] Hibbeler, R.C. (1995) *Structural analysis*, Prentice Hall.
- [5] Timoshenko, S.P. (1966) *History of strength of materials*, Arkady, Warszawa, - in Polish
- [6] Niezgodziński, M.E. and Niezgodziński, T. (1979) *Strength of materials*, State Scientific Publishing House, Warszawa, - in Polish
- [7] Zienkiewicz, O.C. and Taylor R.L. (2000) *The finite element method*, Oxford Press, UK.
- [8] Przewłocki, J. and Górski, J. (2006) *Basis of theory of structures*, Arkady, Warszawa, - in Polish
- [9] Dyląg, Z. and Krzemińska-Niemiec, E. and Filip, F. (1989) *Theory of structures*, State Scientific Publishing House, Warszawa, - in Polish
- [10] Cywiński, Z. (1976) *Theory of structures in problems. Vol.II. The rudiments of statically indeterminate systems*, State Scientific Publishing House, Warszawa - Poznań, - in Polish
- [11] Tan, Z-Q., Jiang, X-D., He, Y-S., Ban, S-H., Xu, R. and Xi R-Q. (2018) Generalized variational principles for solutions of statically indeterminate trusses under mechanical-thermal-assembly loadings, *Journal of Engineering Mechanics*, Volume 144, Issue 1.
- [12] Makowski, Z.S. (1981) *Analysis, design and construction of double-layer grids*, Applied Science Publishers, London.
- [13] Rębielak, J. (2014) A two-stage method for an approximate calculation of statically indeterminate trusses, *Journal of Civil Engineering and Architecture*, **78**, 567-572.
- [14] Rębielak, J. (2016) Simple method of approximate calculation of statically indeterminate trusses, *Proceedings of The 7th International Conference on Computational Methods (ICCM2016)*, August, 1-4, 2016, University of California, Berkeley, USA, 748-753.
- [15] Rębielak, J. (2017) Examples of applications of two-stage method in calculations of statically indeterminate trusses, *Proceedings of The 8th International Conference on Computational Methods (ICCM2017)*, July, 25-29, 2017, Guilin, China, 644-649.

Implementing Axisymmetric Smoothed Finite Method (S-FEM) Element in ABAQUS to Analyze defective Pressure Piping

X. Cui¹, Y.H. Qie¹, S.Y. Duan^{1,2}, and S. H. Huo¹

¹ State Key Laboratory of Reliability and Intelligence of Electrical Equipment,
HeBei University of Technology, Tianjin City, 300401 China

² State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body,
Hunan University, Changsha, Hunan 410082, China

Abstract

Pressure piping is the most productive way for large volume compressed natural gas (CNG) transportation. In pipeline constructions, the thickness of between two pipes joining together is often not consistent due to the miss-match in dimensions, and thus stress concentrations can often occur at the pipe joints, arising safety concerns. Therefore, it is of importance to accurately analyze the key influencing factors of dimensional miss-matching defects, providing theoretical basis for the preliminary design and post-repair of pipelines. This work uses the smoothed finite element method (S-FEM) that has been proven accurate in stress analysis compared with the traditional FEM. Since the geometry and the load of the pressure piping are both axisymmetric, a novel axisymmetric S-FEM element is firstly developed, coded and integrated in ABAQUS using the user-element-library (UEL). Intensive studies are then carried out to examine the effects of different level of miss-match in the thicknesses of two joined pipes and the effects of the radius of the transitional fillet used to bridge the miss-matches. It is found that the maximum hoop stress reduces as the radius of the transitional fillet increases. For the thinner section of the pipe, the maximum hoop stress is only affected by the thickness miss-match.

Keywords: Pressure piping; Finite element analysis (FEA); Smoothed finite element method (S-FEM); Hoop stress; Transitional fillet

Introduction

As demand of compressed natural gas (CNG) is increasing, pressure piping is widely used for long distance CNG transportation. However, leaks and explosions of pressure piping could cause huge property losses even threaten people's life safety. There are many possible sources

of defect such as production defect, installation error, fatigue, etc. [1-3]. But there is very little research has been done about defect of different level of miss-match in the thicknesses of two joined pipes caused by production process. Since it is an important kind of reason for pressure piping's failure, it's necessary to analyze the defective piping under working condition providing theoretical basis for the preliminary design and post-repair.

In general, finite element analysis (FEA) [4-5] is the most common way for model simulation as its high efficiency and low cost. For this study, as geometry and load of the pressure piping are both axisymmetric, two-dimension (2D) axisymmetric elements is selected to discrete the problem domain. Relate large mesh density should be applied at the defective section to compute more accurate results. Nevertheless, for current FEA software packages, it's hard to control the most suitable mesh density at defective sections. If mesh density is too small, it's impossible to compute accurate maximum stress for stress concentration issue. On the contrary, overlarge mesh density will cause computing inefficiency and heavily distorted elements will occur between dense and sparse mesh regions. Therefore, smoothed finite element method (S-FEM)[6-[8] is processed to analyze models instead of the traditional FEM. It has been proved that S-FEM computing more accuracy results and quicker convergence speed compared with traditional FEM. In addition, models with distorted mesh condition can be calculated by S-FEM with accuracy results.

For pressure piping in this study, the defect of different level of miss-match in the thicknesses of two joined pipes is analyzed by S-FEM in ABAQUS. In general, at the stagger sections of interconnected piping, transitional fillet is applied to reduce the stress concentration. Examining the effects of different level of miss-match in the thicknesses of two joined pipes and the effects of the radius of the transitional fillet for the pressure piping.

Effect of thickness difference and transition fillet size on hoop stress

In this section, Pressure piping with the defect of thickness difference and transition fillet size have been studied. And the model of thickness difference at both inner surface and outer surface are all analyzed with S-FEM. Finally, the effect of the thickness difference and the size of the transition fillet on the maximum hoop stress is obtained.

We firstly analyzed a 500mm pressure piping with thickness difference at middle of the inner surface as shown in Fig. 1. The pressure piping can be divided into two parts. For the lower part, inner and outer radius are fixed to 230mm and 250mm, respectively. But for the upper part, the outer radius is fixed to 250mm, the inner radius is changed from 230.5mm to 234mm,

and the increment is 0.5mm. Meanwhile, the radius of transition fillet is also changed from 0.5mm to the maximum thickness difference, and the increment also is 0.5mm.

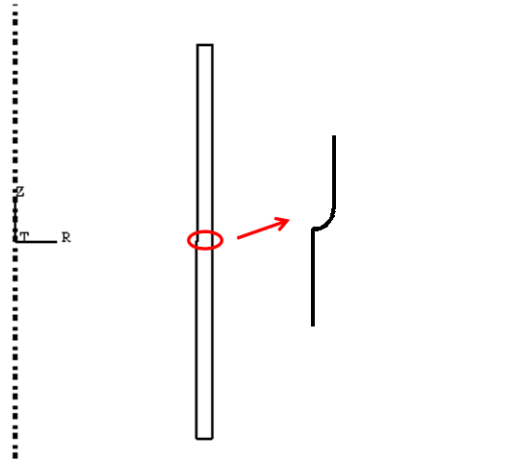


Fig. 1 Pressure piping with defect of thickness difference at inner surface

All the models are analyzed by the S-FEM in ABAQUS. Since the hoop stress is the largest among the stress components, the hoop stress is compared for each model. The maximum hoop stress is at the transition fillet, and there is also relatively larger hoop stress at the upper part because of the thin wall thickness. Therefore, the maximum hoop stress is compared at the transition fillet and upper part respectively for all the models. And the effect of thickness difference and transition fillet size on hoop stress can be obtained. The purpose of making convenient for results analyzing, the value of the maximum hoop stress of the defective piping is divided by that of the piping without defect. For the thin-walled part, the maximum hoop stress is only effected by the thickness difference rather than the size of the transition fillet as shown in Fig. 2 (a) and (b). And for the transition fillet part, the maximum hoop stress has linear relationship with the thickness difference as shown in Fig. 2 (c). In addition, keeping the thickness difference unchanged, the maximum hoop stress has quadratic function relation with the radius of the transition fillet as shown in Fig. 2 (d). For the case of thickness difference at the outer surface, the change rule of the maximum hoop stress for different size of thickness difference and radius of the transition fillet is the same as that of the defect at inner case.

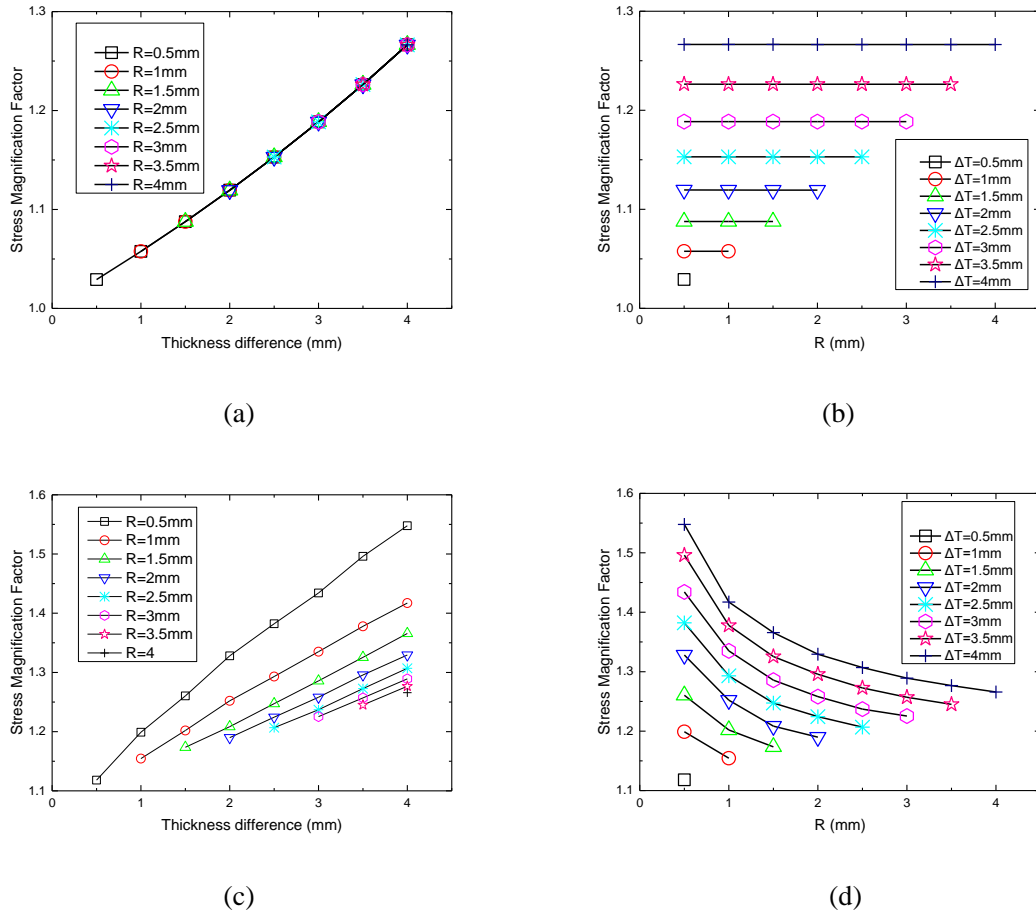


Fig. 2 The maximum hoop stress: (a) for different thickness difference at the thin-walled part; (b) for different radius of the transition fillet at the thin-walled part; (c) for different thickness difference at the transition fillet part; (d) for different radius of the transition fillet at the transition fillet part.

References

- [1] M. Law, G. Bowie. (2007) Prediction of failure strain and burst pressure in high yield-to-tensile strength ratio linepipe, *International Journal of Pressure Vessels and Piping* 84: 487–492
- [2] T. Aseer Brabin, T. Christopher, B. Nageswara Rao. (2010) Finite element analysis of cylindrical pressure vessels having a misalignment in a circumferential joint, *International Journal of Pressure Vessels and Piping* 87: 197-201
- [3] J.E. Christopher, F.M. Timothy. (2015) Failure prediction of pressure vessels using finite element analysis, *Journal of Pressure Vessel Technology*: 137
- [4] G.R. Liu, S.S. Quek. (2010) *The Finite Element Method: A Practical Course* (second edition), Butterorth-Heinemann. Oxford.
- [5] K.J. Bathe. (1996) *Finite Element Procedures*. Prentice Hall NJ.
- [6] G.R. Liu, N.T. Trung. (2010) *Smoothed Finite Element Methods*. CRC Press. FL .
- [7] Liu GR. (2009) *Meshfree methods: moving beyond the finite element method*. CRC Press, Boca Raton, 1st Edition, 2002; 2nd Edition.
- [8] G.R. Liu, Y.T. Gu. (2005) *An introduction to meshfree methods and their programming*. Springer.

An inverse method for identification of continuously varying material properties in post-manufactured structures through neural networks

† S.Y. Duan^{1,2}, Z.M. Zhang¹

¹ School of Mechanical Engineering, Hebei University of Technology, Tianjin, China

² State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha, Hunan 410082, China

†Corresponding author, dsy605106@163.com

Abstract

Determination of material properties is essential for structural design, optimization and prediction of the mechanical behavior of the structures. Currently, the material properties are measured after the material is produced using standard specimen. However, the material properties at some parts of the structure can change non-uniformly as the results of plastic deformation and hardening during the manufacturing processes, such as metal forming processes. Therefore, the material properties of the manufactured structure components become nonhomogeneous, can vary continuously, and can be very much different than those evaluated based on the specimen measurements. This paper presents an inverse method to systematically determine the variation of the material properties in post-manufactured structure components via, for example, a metal forming process. Our inverse technique uses ABAQUS to simulate the mechanical behavior of the structure, and then the responses are used to train a radial basis function (RBF) neural network (NN). The material properties at any point of the structure can then be obtained rapidly and accurately by measuring the structure responses. Our technique is validated using a hat-shaped plate formed by stamping. It is found from a simulated measurement data that the average error of the inversely identified is less than 2%, which is sufficient for most of the engineering applications.

Keywords: Material properties; Inverse problem; Inverse identification; RBF neural network; Metal forming

Introduction

In structural design, optimization or safety analyses, the accuracy of material properties directly determines the reliability of the results [1,2]. Currently, most of the material parameter identification methods are applied to materials without considering the effects of manufacturing processes, and hence the obtained results are effectively the equivalent parameters that omits the spatial inhomogeneity. In actual fact, however, due to the non-uniform loading and plastic deformation in the structure during manufacturing processes such as forming processes, there are significant discrepancy in the material properties at different parts of the structure [3,4,5]. To overcome this shortcoming, an inverse method is proposed in this paper which allows the material properties changes continuously in space manufactured structures.

Description of the problem

The material properties are assumed vary continuously spatially. In the space over which the material properties change, it can be divided into a number of key regions using discrete points. Then, the material properties of the structure's region are obtained by interpolating values at these discrete points. Finally, only the material properties at these key points are

needed to be identified, thus it can reduce the number of inversion parameters. Radial basis function neural networks and finite element methods are used in the present method for solving the inverse problem and collecting training data needed for neural networks. The task of this work is to establish such a method that can systematically and rapidly identify continuously distributed material properties. In order to develop an effective and practical procedure for continuously distributed material properties, we employ the following strategies:

- Use the well-established FEM (in particular ABAQUS®) as the forward model. This is because the material properties in any element in a FEM model can be different from the others. In addition, ABAQUS is mature and reliable, and can compute structural responses very fast for a given distribution of materials for a structure.
- The ABAQUS models are then used to train a neural network (NN), so that during the inverse analysis, we can infer the material parameters in real-time without calling for time-consuming ABAQUS.
- In training of the neural networks, we use radial basis functions (RBF). This is because the moment matrices using RBF is always invertible and hence the training process becomes reliable and effective. Our RBF-NN is expected offering a real-time inverse solver for identifying the continuous distribution of material properties of post-manufactured structures.

Finite element model

This method is applied to the material parameter identification of hat-shaped structure. Due to the symmetry of the plate, only the one-half model is selected for modeling shown in Fig. 3.1. The left end of the one-half model is completely fixed and the load in z direction is applied to the right end of the structure.

The material used for the model is stainless steel SUS201, which has an initial Young's modulus of 207000 MPa. In fact, after stamping process, the Young's modulus of material is continuously variable throughout the entire structure, even the Young's modulus in some local areas of the structure may reach twice of the original Young's modulus [6,7]. Therefore, the model is firstly divided into four blocks by five measuring points 1, 5, 8, 9, 12 as shown in Fig. 3.1 and each block is then divided into several equal-sized areas. The whole model is divided into 254 areas with different Young's modulus, and the Young's modulus of each area is obtained by interpolating from that of the measurement points at both ends of the block. As long as the Young's modulus of the measuring points (5, 8, 9, 12) is determined, the Young's modulus at other positions is obtained. To avoid stress singularity, appropriate triangular meshes is applied. The total number of elements in the model is 101600 and the number of nodes is 51255.

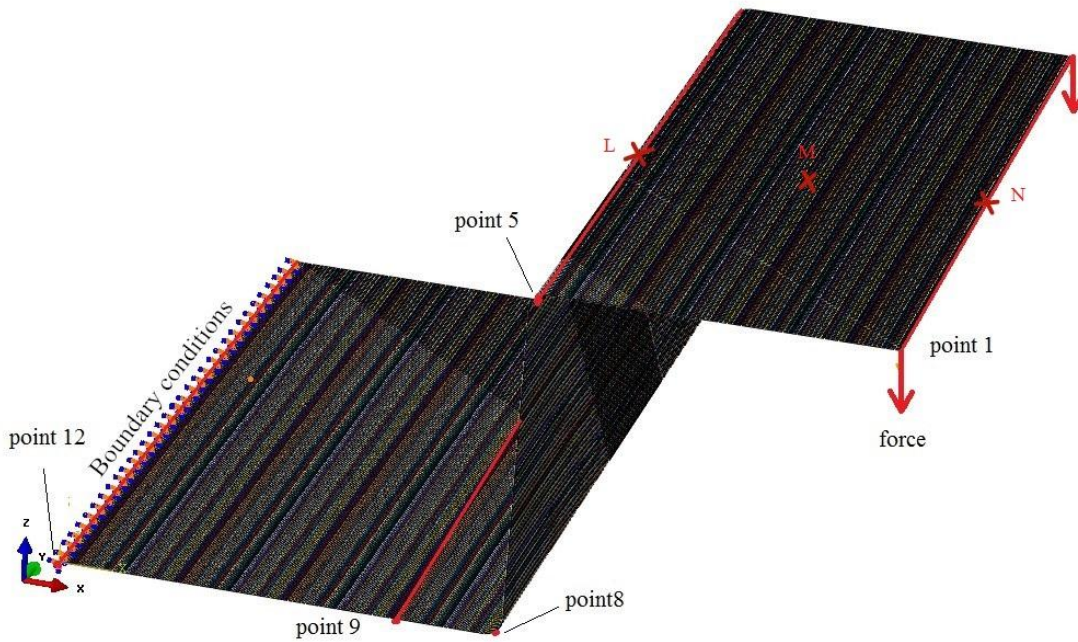


Fig. 0.1 One-half model of hat-shaped structure

Sensitivity analysis and establishment of neural networks

In the process of material properties inversion, there should be a strong sensitivity between the parameters to be reversed and the measured response of the load, in order to guarantee the existence and solvability of the inverse solution [8,9,10,11,12]. After sensitivity analysis, it is found that the z-direction and the x-direction displacement of positioning points L, M, and N are sensitive to the change in material parameters as shown in Fig. 3.1 and Table 0.1. The displacement responses are represented by U_{Lx} , U_{Lz} , U_{Mx} , U_{Mz} , U_{Nx} , U_{Nz} , where U represents the displacement response, L, M, and N are the names of the positioning points, and x and z represent the direction of the displacement. It can be seen in Table 0.2 that the displacement responses of the three positioning points (L, M, N) are less sensitive to the change of Young's modulus at point 5 while they are highly sensitive to the change of Young's modulus at points 9 and 12. Because the deformation of the structure always occurs firstly in the position with a lower Young's modulus when the structure withstands external forces. The Young's modulus at point 9 and point 12 are smaller than that of point 5 and point 8, thus, the displacement is likely to be more sensitive to here. Based on the results of sensitivity analysis, with these six displacement responses of the points (L, M, N) as inputs, the material parameters of the five measuring points are outputs, the inverse problem neural network is established, and appropriate training and error analysis are performed.

Table 0.3 The sensitivity coefficient of displacement responses of points L, M, and N to the changes of Young's modulus of the measuring points

point	U_{Lx}	U_{Lz}	U_{Mx}	U_{Mz}	U_{Nx}	U_{Nz}
5	4.44%	1.04%	4.44%	8.27%	4.44%	12.02%
8	21.37%	4.00%	21.37%	13.33%	21.37%	16.44%
9	40.07%	38.61%	40.07%	35.96%	40.07%	34.42%
12	39.09%	58.52%	39.09%	46.30%	39.09%	41.10%

Conclusions

In this work, a novel method of continuous inverse identification of material properties was proposed to accurately obtained the material properties at arbitrary position of structure. It is found that the average error of the inversely identified result is less than 2%, compared with a simulated measurement data. The radial basis function neural network is used to solve the inverse of the positive problem because of its advantages in solving complex nonlinear problems. The method of approximating the material properties at any position of the structure by interpolating material properties of key positions of the structure reduces the number of outputs and simplifies the solution of inverse problem. In addition, this method separates the simulation from the solution and it is more convenient for engineering applications.

References

- [1] Cao Yinfeng, Li Guangyao, Zhong Zhihua. (2004). An inverse method for material parameters in metal forming. *Chinese Journal of Computational Mechanics*, 21(3) :291-297
- [2] Liu, G. R., & Quek, S. S. (2003). The finite element method: A practical course. The finite element method: a practical course. Butterworth-Heinemann.
- [3] Chao Li. (2012). An inverse method of material parameters about vehicle body panels based on region division. Master's dissertation, Hunan University.
- [4] Gao Hui, Li Guangyao, Han Xu. (2008). Study on Direct Identification Method of Material Parameters. *CHINA MECHANICAL ENGINEERING*, 19(17), 2081-2084.
- [5] Chao Li. (2012). An inverse method of material parameters about vehicle body panels based on region division. Master's dissertation, Hunan University.
- [6] Wang Jinlun, Lan Fengchong, Chen Jiqing. (2012). The Effects of Stamping Forming-caused Changes in Material Property on Crashworthiness. *Automotive Engineering*, 36(4), 367-372.
- [7] Han Xu. (2015). Numerical simulation-based Design: theory and method. Science Press.
- [8] Hossain, M. S., Zhi, C. O., Ismail, Z., Noroozi, S., & Khoo, S. Y. (2016). Artificial neural networks for vibration based inverse parametric identifications. *Applied Soft Computing*, 52(C), 203-219.
- [9] Lifan HAN. (2005) The Research on Inverse Problems in Sheet Metal Forming Processes Based on Neural Network. Doctoral dissertation, Hunan University.
- [10] Liu, G. R, Darling, T. (2014). Computational inverse techniques in nondestructive evaluation. *Well Logging & Formation Evaluation*, ix.
- [11] Muñoz-Sánchez, A., González-Farías, I. M., Soldani, X., & Miguélez, M. H. (2011). Hybrid fe/ann and lpr approach for the inverse identification of material parameters from cutting tests. *International Journal of Advanced Manufacturing Technology*, 54(1-4), 21-33.
- [12] Pabisek, E., & Waszczyszyn, Z. (2015). Identification of thin elastic isotropic plate parameters applying guided wave measurement and artificial neural networks. *Mechanical Systems & Signal Processing*, 64-65, 403-412.

An oscillation-free finite volume method with staggered grids for solving problems of poroelasticity

***Clovis R. Maliska¹ and Hermínio T. Honório¹**

¹Department of Mechanical Engineering
Federal University of Santa Catarina
88040-900 – Florianópolis – SC - Brazil

*Presenting and corresponding author: maliska@sinmec.ufsc.br

Abstract

This paper advances a new finite volume alternative for solving poroelasticity problems employing a staggered arrangement for pressure and displacements in an unstructured grid framework. By staggering these variables, an improvement is obtained for the pressure-displacement coupling, which is claimed by the authors to prevent the numerical solution from instabilities in the pressure field. The two-dimensional formulation is still under development, but preliminary one-dimensional results are presented to demonstrate this capability. It is shown that the staggered formulation keeps second order accuracy for both pressure and displacement, even for highly non-uniform grids. In addition, the formulation does not present any spurious pressure oscillations, a key issue when solving poroelasticity problems under undrained conditions. The results suggest that staggering the rock displacements related to the pore pressure is a very promising approach to confer robustness to the numerical scheme. The novel method is founded on the analogy among pressure-velocity coupling for the Navier-Stokes equations and pressure-displacement for the poroelasticity problems. It is worth to mention that both physics are treated with the same conservative method.

Keywords: Staggered grids, unstructured grids, numerical stability, poroelasticity

Introduction

Several engineering problems are modeled by systems of coupled partial differential equations, many of them involving different physics. In geomechanics, in which compacting porous media is coupled with the fluid flow, is one example. In this case, a delicate coupling between pore-pressure and rock displacement is present, since under certain conditions, as in the very beginning of the transient, or at the interface of two materials with different permeability, pressure wiggles appear in the numerical solution. Those situations, which resemble an undrained condition, impose an almost zero compressibility, which creates the condition for this pathology to appear. In the class of Finite Element methods, extensively used for solving the rock mechanics in porous media, several remedies for this pathology is available, being mixed finite element [1] and discontinuous Galerkin some of the possibilities. However, those remedies are at a cost of considerably increasing in computer time. Alternatively, some authors [2,3] have proposed stabilization techniques that do not increase the computational cost and still eliminate the instabilities, but at a cost of introducing numerical diffusion to the solution. Recently, in the context of finite volumes, Honório and Maliska [4] have proposed a strategy for avoiding such instabilities, which can be also regarded as a stabilization technique. In spite of all these alternatives, a numerical scheme that efficiently eliminates the pressure wiggles without increasing computational cost, while keeping the same order of accuracy for both pressure and displacements, is still pursued.

An analysis of the coupling between pressure and displacement for poroelasticity, and pressure and velocity for Navier-Stokes flows, reveals that they are of the same nature, so it is expected that the remedies employed in one class of problems can be applied to the other one with success. It should be recalled that the oscillatory pressure fields arising when solving incompressible Navier-Stokes flows, and its remedies, is known for more than four decades, and can be fully mitigated if a staggered grid approach is employed [5]. This remedy was abandoned when unstructured grids were required for solving fluid flows in complex geometries, due to the alleged complexity of implementation. This paper addresses this issue, advancing a finite volume method using unstructured grids with staggered variables, avoiding the oscillatory pressure field that appears in poroelasticity. Another important feature of the methodology herein presented is the solution of the both physics, fluid flow and rock mechanics, with the same finite volume technique [6].

Firstly, it is presented the mathematical model for the coupled poroelasticity problem. Then, the fundamentals of the staggered arrangement of variables are discussed, and a brief analogy is established between pressure-displacement in poroelasticity and pressure-velocity for Navier-Stokes flows. The model equations are discretized for two-dimensions and results are presented for 1D problem for both staggered and collocated arrangements. Finally, a few remarks close the work.

Mathematical Model

The mechanical behavior of saturated porous media, besides being affected by its mechanical properties, it is also influenced by the pressure of the fluid filling its pores. Terzaghi [7] introduced the concept of effective stress into the equations of stress equilibrium in order to take the pore pressure into account, yielding

$$\nabla \cdot \boldsymbol{\sigma} - \alpha \nabla p = \mathbf{b} \quad (1)$$

in which ∇ is the nabla operator, $\boldsymbol{\sigma}$ is the effective stress tensor, α is the Biot coefficient, p is the pore pressure and \mathbf{b} is a source term. Moreover, considering small strains and a stress-strain relationship represented by the constitutive matrix \mathbf{E} (Voigt notation), the effective stress tensor can be written in terms of the displacement vector \mathbf{u} by the expression

$$\boldsymbol{\sigma} = \mathbf{E} \nabla_s \mathbf{u} \quad (2)$$

with ∇_s being the symmetric nabla operator. The closure of the model is ensured by the mass conservation equation for deformed porous media, given by

$$\frac{1}{M} \frac{\partial p}{\partial t} + \nabla \cdot (\mathbf{v}^f + \mathbf{v}^s) = q \quad (3)$$

in which $1/M$ is the Biot module and q is a source term. Equation (3) is conveniently written here in terms of the fluid velocity, \mathbf{v}^f , and the solid grains velocity, \mathbf{v}^s , which are respectively given by

$$\mathbf{v}^f = -\frac{\mathbf{k}}{\mu} \cdot \nabla p \quad (4)$$

$$\mathbf{v}^s = \frac{\partial \mathbf{u}}{\partial t} \quad (5)$$

with \mathbf{k} being the absolute permeability tensor and μ the fluid viscosity. The gravitational term in equation (4) has been neglected with no loss of generality.

Staggered Grid Arrangement

One of the major challenges faced by the numerical schemes developed to solve equations (1) and (3) is how to avoid pressure wiggles that can appear under undrained consolidation. In this situation the consolidation process takes place in a much smaller time scale than the fluid motion ($\mathbf{v}^s \gg \mathbf{v}^f$), which yields the following mass conservation equation

$$\frac{1}{M} \frac{\partial p}{\partial t} + \nabla \cdot \mathbf{v}^s = q \quad (6)$$

Equation (6) is very similar to the mass conservation equation that appears when solving the Navier-Stokes equations. It is well known that satisfying this equation is of utmost importance to avoid the well-known checkerboard pressure problem [8]. The key issue in poroelasticity resides on how to determine a displacement field that satisfy both mass and momentum equations. For the Navier-Stokes equations the problem is exactly the same, except that the unknown variable is the velocity field instead of displacement.

Ensuring mass and momentum conservation is not a trivial task to be accomplished. The pioneering work to address this issue is due to Harlow and Welch [5], in the context of finite differences. They staggered the positions of pressure and velocities, such that momentum and mass conservation are satisfied for different control volumes but for the same set of variables. In this manner, pressure and velocities are directly available where they are required when integrating the corresponding partial differential equation in a control volume for mass or momentum. This technique is recognized to completely mitigate pressure wiggles for the Navier-Stokes equations. Due to the similarity of equation (6) with its counterpart in Navier-Stokes flows, a staggered grid arrangement between pressure and displacement might have strong chances to completely eliminate the pressure wiggles appearing in poroelasticity problems.

Finite Volume Formulation

The traditional procedure for obtaining the approximate equations via a finite volume method starts by integrating the differential equations over each control volume. Since it is being proposed a staggered arrangement for \mathbf{u} and p the control volumes for pressure and displacement must be clearly identified. In this work, we follow the methodology presented by Peters and Maliska [9] for building the staggered control volumes on unstructured grids.

In figure (1a) it is shown the base mesh (the elements) provided by the grid generator. As depicted in figure (2b) the control volumes for mass conservation, Ω^p , coincides with the elements of the base mesh. This control volume is for pressure. For the momentum equilibrium, however, the control volumes, Ω^u , are built around the edges of the elements by connecting the vertices of the edge with the centroids of the two adjacent elements. A control volume Ω^u is represented in figure (1c) and the position of \mathbf{u}_j is at the midpoint of the edge

of the element. The key point of this configuration is that the displacements are located at the faces of the control volume Ω^p (figure (1b)), which will have a direct impact on the volumetric strain computation over Ω^p , as will be shown later.

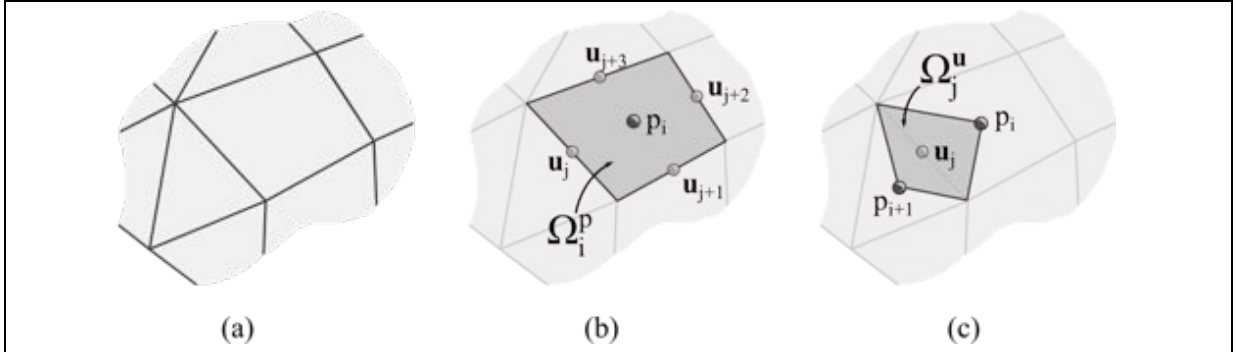


Figure 1 Geometrical entities: (a) mesh composed by triangular and quadrilateral elements; (b) control volume for mass conservation and (c) control volume for momentum conservation.

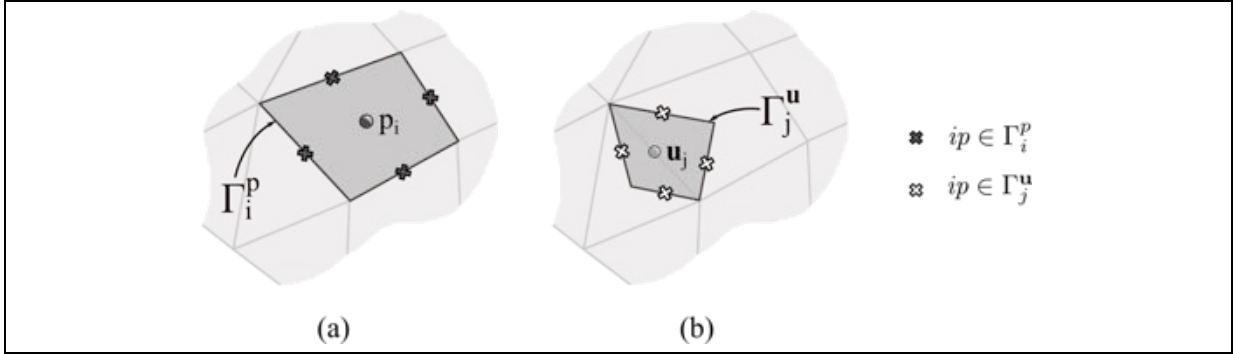


Figure 2 Integration points (a) for mass and (b) for momentum

Mass Conservation Equation

Equation (3) is integrated over a time step, Δt , along with an implicit first-order backward Euler scheme and over the control volume Ω_i^p . The divergence theorem is applied to obtain the surface integrals. By the midpoint rule, the semi-discretized form of equation (3) is

$$\frac{\Delta \Omega_i^p}{M} \frac{p_i}{\Delta t} + \sum_{ip \in \Gamma_i^p} \left[(\mathbf{v}^f + \mathbf{v}^s) \cdot \mathbf{s} \right]_{ip} = q_i \Delta \Omega_i^p + \frac{\Delta \Omega_i^p}{M} \frac{p_i^o}{\Delta t} \quad (7)$$

in which the variables evaluated at the previous time level carries the superscript o , and no superscript refers the current time level. Each control volume Ω^p is bounded by a set of faces (or edges) and at the midpoint of each face is located an integration point ip . The set of integration points surrounding Ω_i^p is denoted by Γ_i^p , as highlighted in figure (2). Each integration point has an area vector, \mathbf{s}_{ip} , pointing outwards the control volume. In addition, the volume of Ω^p is represented by $\Delta \Omega^p$. Recalling equation (5), the mass fluxes crossing the faces of Ω^p due to the rock deformation and fluid motion are, respectively, given by

$$w_{ip}^s \approx (\mathbf{v}^s \cdot \mathbf{s})_{ip} = \frac{(\mathbf{u}_{ip} - \mathbf{u}_{ip}^o)}{\Delta t} \cdot \mathbf{s}_{ip} \quad (8)$$

$$w_{ip}^p \approx (\mathbf{v}^f \cdot \mathbf{s})_{ip} = -\frac{1}{\mu} (\mathbf{k} \cdot \nabla p)_{ip} \cdot \mathbf{s}_{ip} \quad (9)$$

The main advantage of staggering Ω^p and Ω^u becomes clear by inspecting equation (8), noting that the displacement vectors \mathbf{u}_{ip} and \mathbf{u}_{ip}^o are directly available at the integration points of Γ_i^p (see figure (2)), avoiding any kind of interpolation. The benefits of this feature are of particular importance during undrained consolidation (equation (6)), where the mass fluxes through the control volume's faces is entirely given by w^s . This is precisely the point one is claiming to be the key point for avoiding the pressure instabilities.

The next step is to choose how to reconstruct the pressure gradient of equation (9) at the integration points belonging to Γ^p . The literature is abundant on these kind of methods, and Cerbato et al. [10] present an extensive analysis of several techniques for gradient reconstruction specifically applied to unstructured grids, which could be readily applied here to approximate equation (9). The reconstruction could also be done by a Multi-Point Flux Approximation (MPFA), as proposed by Aavastmark et al. [11].

Equilibrium Equations

Equation (1) is integrated over the control volume Ω_j^u , as depicted in figure (1c), and the divergence theorem is applied to the divergent operator yielding

$$\sum_{ip \in \Gamma_j^u} (\boldsymbol{\sigma} \cdot \bar{\mathbf{s}})_{ip} - \alpha \int_{\Omega_j^u} \nabla p \, d\Omega_j^u = \mathbf{b}_j \Delta \Omega_j^u \quad (10)$$

in which Γ_j^u is the set of integration points surrounding Ω_j^u , as shown in figure (2b), and $\bar{\mathbf{s}}$ is an appropriate arrangement of the area vector components, which for the two-dimensional case is

$$\bar{\mathbf{s}} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \\ s_y & s_x \end{bmatrix} \quad (11)$$

The volumetric integral of the pressure gradient in equation (10) is approximated by the Green Gauss theorem

$$\int_{\Omega_j^u} \nabla p \, d\Omega_j^u \approx \nabla p_j \Delta \Omega_j^u \quad , \quad (12)$$

Now, it is important to notice that ∇p_j is exactly the same as the pressure gradient required by equation (9), since a displacement position j always coincide with an integration point belonging to Γ_i^p , as it can be seen in figure (2). Therefore, the methodology chosen to evaluate equation (9) can be the same used to compute equation (12).

The remaining term to be evaluated in equation (10) is the stress tensor, σ_{ip} , at the integration point belonging to Γ_j^u . This is performed using equation (2). The procedure to compute the displacement derivatives $\nabla_s \mathbf{u}$ at the integration points of Γ_j^u follows the approach presented in [9] for computing the velocity derivatives.

One Dimensional Formulation

The corresponding 1D formulation of the methodology just described is now considered. The 1D formulation simplifies considerably the geometry, but still carries all the ingredients to evaluate the ability of the scheme to avoid pressure instabilities. Therefore, some preliminary results of the above formulation using the grid shown in figure (3) are presented. The results are compared with the traditional collocated arrangement of variables, as depicted in figure (4). It is worth to mention that the grids can be unequally spaced.

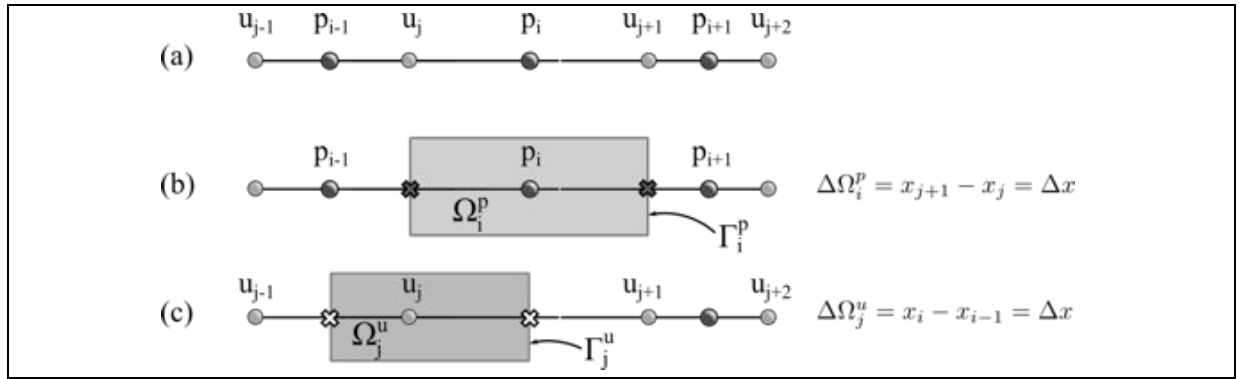


Figure 3 (a) 1D grid, (b) control volume for pressure and (c) control volume for displacement.

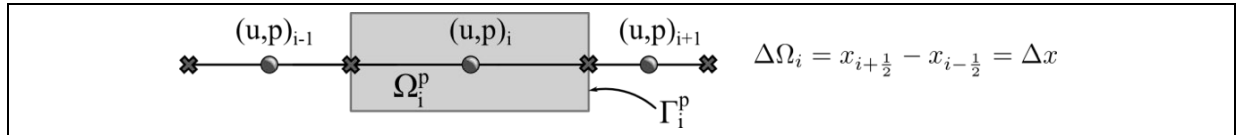


Figure 4: Collocated (coincident) control volumes associated to a 1D grid

1D Staggered Grid Formulation

Integrating equations (3) and (1) over Ω_i^p and Ω_j^u of figure (3), respectively, results in

$$\frac{\Delta\Omega_i^p}{M} \frac{p_i}{\Delta t} - \frac{k}{\mu} \left(\left. \frac{\partial p}{\partial x} \right|_{j+1} - \left. \frac{\partial p}{\partial x} \right|_j \right) + \frac{\alpha}{\Delta t} (u_{j+1} - u_j) = \frac{\Delta\Omega_i^p}{M} \frac{p_i^o}{\Delta t} + \frac{\alpha}{\Delta t} (u_{j+1}^o - u_j^o) \quad (13)$$

$$\sigma_i - \sigma_{i-1} - \alpha(p_i - p_{i-1}) = 0 \quad (14)$$

with the following approximations at the integration points

$$\left. \frac{\partial p}{\partial x} \right|_{j+1} \approx \frac{p_{i+1} - p_i}{x_{i+1} - x_i} = \frac{p_{i+1} - p_i}{\Delta x_{j+1}} \quad (15)$$

$$\left. \frac{\partial p}{\partial x} \right|_j \approx \frac{p_i - p_{i-1}}{x_i - x_{i-1}} = \frac{p_i - p_{i-1}}{\Delta x_j} \quad (16)$$

$$\sigma_i = (\lambda + 2G) \left. \frac{\partial u}{\partial x} \right|_i \approx (\lambda + 2G) \frac{u_{j+1} - u_j}{x_{j+1} - x_j} = (\lambda + 2G) \frac{u_{j+1} - u_j}{\Delta x_i} \quad (17)$$

$$\sigma_{i-1} = (\lambda + 2G) \left. \frac{\partial u}{\partial x} \right|_{i-1} \approx (\lambda + 2G) \frac{u_j - u_{j-1}}{x_j - x_{j-1}} = (\lambda + 2G) \frac{u_j - u_{j-1}}{\Delta x_{i-1}} \quad (18)$$

The set of equations (13) and (14), along with equations (15-18), composes a linear system of equations for pressure and displacement that is solved in a simultaneous (monolithic) fashion.

Co-located Formulation

The discretization of the equations for a collocated grid is basically the same as the previous case, except that the integration for both variables is performed over the same control volume Ω_i . In this manner, the resulting equations are

$$\frac{\Delta \Omega_i}{M} \frac{p_i}{\Delta t} - \frac{k}{\mu} \left(\left. \frac{\partial p}{\partial x} \right|_{i+\frac{1}{2}} - \left. \frac{\partial p}{\partial x} \right|_{i-\frac{1}{2}} \right) + \frac{\alpha}{\Delta t} (u_{i+\frac{1}{2}} - u_{i-\frac{1}{2}}) = \frac{\Delta \Omega_i}{M} \frac{p_i^o}{\Delta t} + \frac{\alpha}{\Delta t} (u_{i+\frac{1}{2}}^o - u_{i-\frac{1}{2}}^o) \quad (19)$$

$$\sigma_{i+\frac{1}{2}} - \sigma_{i-\frac{1}{2}} - \alpha (p_{i+\frac{1}{2}} - p_{i-\frac{1}{2}}) = 0 \quad (20)$$

The approximations at the integration points still hold

$$\left. \frac{\partial p}{\partial x} \right|_{i+\frac{1}{2}} \approx \frac{p_{i+1} - p_i}{x_{i+1} - x_i} = \frac{p_{i+1} - p_i}{\Delta x_{i+\frac{1}{2}}} \quad (21)$$

$$\left. \frac{\partial p}{\partial x} \right|_{i-\frac{1}{2}} \approx \frac{p_i - p_{i-1}}{x_i - x_{i-1}} = \frac{p_i - p_{i-1}}{\Delta x_{i-\frac{1}{2}}} \quad (22)$$

$$\sigma_{i+\frac{1}{2}} \approx (\lambda + 2G) \frac{u_{i+1} - u_i}{x_{i+1} - x_i} = (\lambda + 2G) \frac{u_{i+1} - u_i}{\Delta x_{i+\frac{1}{2}}} \quad (23)$$

$$\sigma_{i-\frac{1}{2}} \approx (\lambda + 2G) \frac{u_i - u_{i-1}}{x_i - x_{i-1}} = (\lambda + 2G) \frac{u_i - u_{i-1}}{\Delta x_{i-\frac{1}{2}}} \quad (24)$$

It can be seen now that equations (19) and (20) require the evaluation of u and p at the integration points, where they are not available, due to the co-located arrangement. In this case, we use a linear interpolation of these variables, thus the following relationships are employed

$$p_{i+\frac{1}{2}} \approx \frac{\Delta x_{i+\frac{2}{3}} p_i + \Delta x_{i+\frac{1}{3}} p_{i+1}}{\Delta x_{i+\frac{1}{2}}} \quad (25)$$

$$p_{i-\frac{1}{2}} \approx \frac{\Delta x_{i-\frac{2}{3}} p_i + \Delta x_{i-\frac{1}{3}} p_{i-1}}{\Delta x_{i-\frac{1}{2}}} \quad (26)$$

$$u_{i+\frac{1}{2}} \approx \frac{\Delta x_{i+\frac{2}{3}} u_i + \Delta x_{i+\frac{1}{3}} u_{i+1}}{\Delta x_{i+\frac{1}{2}}} \quad (27)$$

$$u_{i-\frac{1}{2}} \approx \frac{\Delta x_{i-\frac{2}{3}} u_i + \Delta x_{i-\frac{1}{3}} u_{i-1}}{\Delta x_{i-\frac{1}{2}}} \quad (28)$$

in which $\Delta x_{i+\frac{2}{3}} = x_{i+1} - x_{i+\frac{1}{2}}$, $\Delta x_{i+\frac{1}{3}} = x_{i+\frac{1}{2}} - x_i$, $\Delta x_{i-\frac{1}{3}} = x_i - x_{i-\frac{1}{2}}$ and $\Delta x_{i-\frac{2}{3}} = x_{i-\frac{1}{2}} - x_{i-1}$.

These two formulations will be used for solving the classical problem of poroelasticity, the Terzaghi's column, shown in Figure 5, with two main goals, to obtain the order of approximation of the staggered scheme, and demonstrate that the co-located grid arrangement without a stabilizing scheme is unable to dampen the pressure oscillations, while the staggered arrangement fully mitigate the pressure oscillations without any stabilizing scheme.

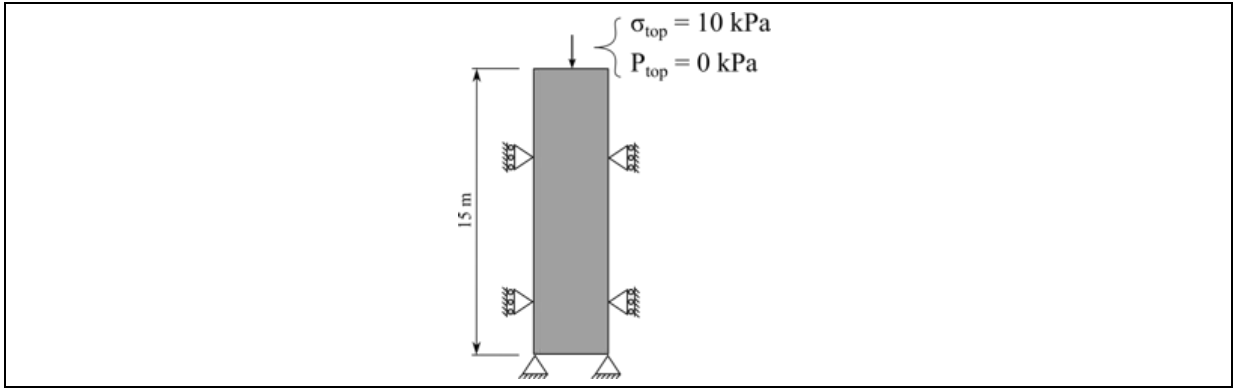


Figure 5 Geometry and boundary conditions for the one-dimensional consolidation problem (Terzaghi's column).

Numerical Results

As depicted in figure (5), the domain has its bottom boundary fixed and impermeable, and the top boundary is fully-permeable ($p_{top} = 0$ kPa) and subjected to a compressive load of $\sigma_{top} = 10$ kPa. The structure is initially not deformed and the initial pore pressure equals to zero. The fluid phase properties are: $\rho = 998,2$ kg/m³, $\mu = 1,002 \times 10^{-3}$ Pa.s and $c_f = 1,0 \times 10^{-4}$ MPa⁻¹. The solid phase properties are: $G = 1,732$ MPa, $\lambda = 2,597$ MPa, $\phi = 0,3$, $\alpha = 1,0$ and $K = 1,0 \times 10^{-4}$ m/s, where K represents the hydraulic conductivity.

The numerical solutions are now checked against analytical solutions for validation purposes, followed by an analysis of the order of approximation of the staggered scheme, concluding demonstrating that the staggered scheme is efficient in damping the pressure oscillations. Slightly and highly non-uniform spaced grids are employed. With a fixed time step size of 0,1 seconds, the pressure and vertical displacement profiles along the vertical direction are plotted against the analytical solution for specified time levels.

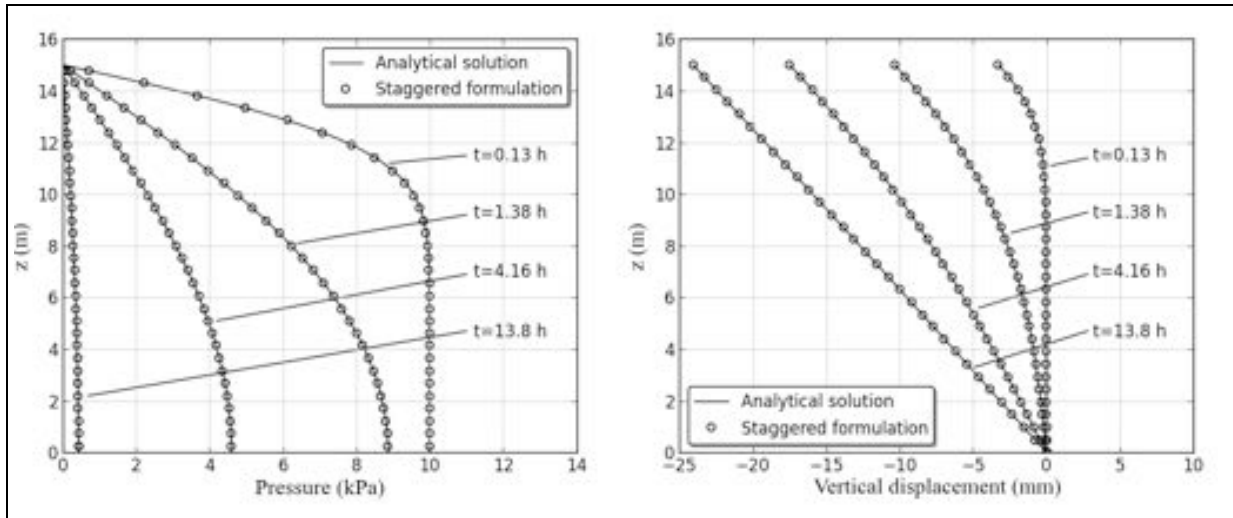


Figure 6 Pressure and displacement fields. Slightly non-uniform grid

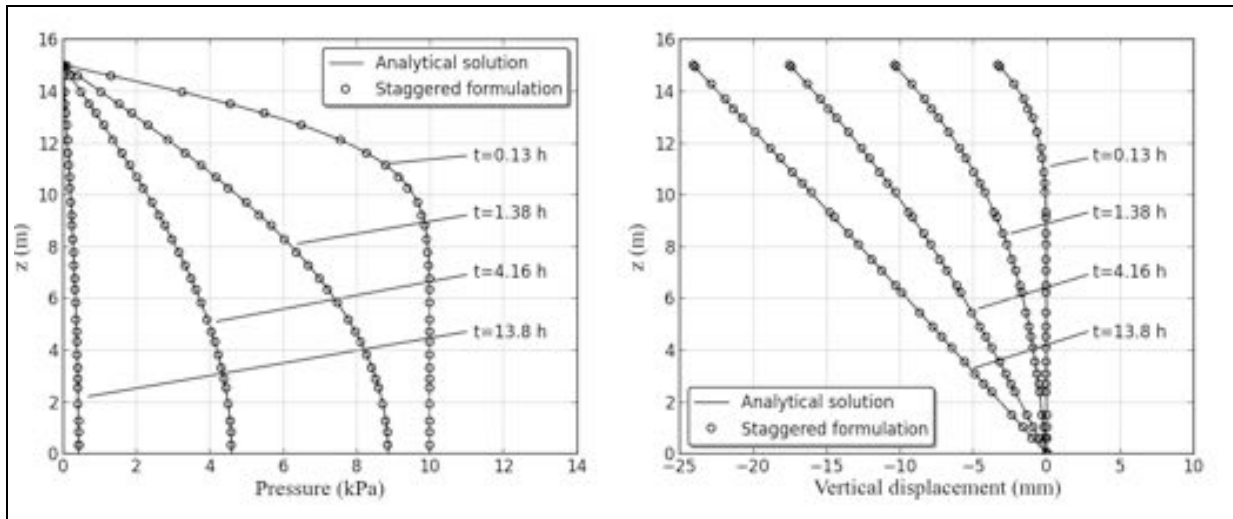


Figure 7 Pressure and displacement fields – Highly non-uniform grid

Figures (6) and (7) show the numerical and analytical solutions for the pressure and displacement profiles for slightly and highly non-uniform grids, indicating that the solutions are not affected by the non-uniformity of the grid.

Convergence Analysis

The assessment of the convergence characteristics of the staggered scheme is performed considering two sets of progressively refined grids. The sets have slightly and highly non-uniform spaced grids randomly generated. For each set of grids, pressure and displacement profiles are taken at $t = 500$ seconds. These profiles are compared with the analytical solutions and the Euclidean norm (L_2 -norm) of the error vector is computed. Four different time step sizes are considered: 0,1, 1, 10 and 100 seconds.

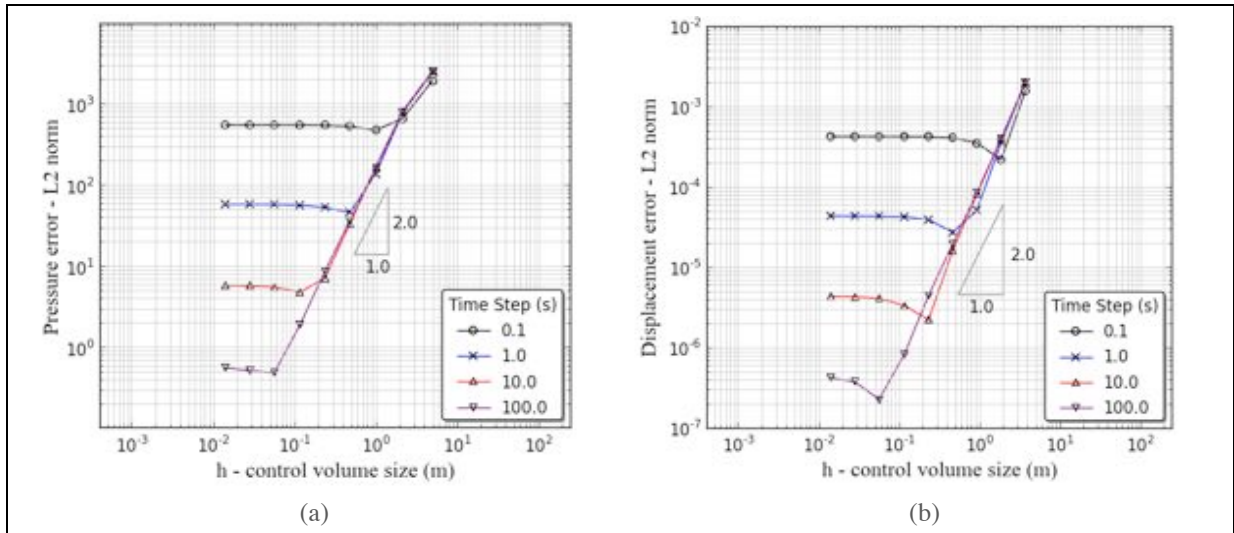


Figure 8 Convergence analysis of the staggered scheme for slightly non-uniform grid. (a) Pressure and (b) displacement.

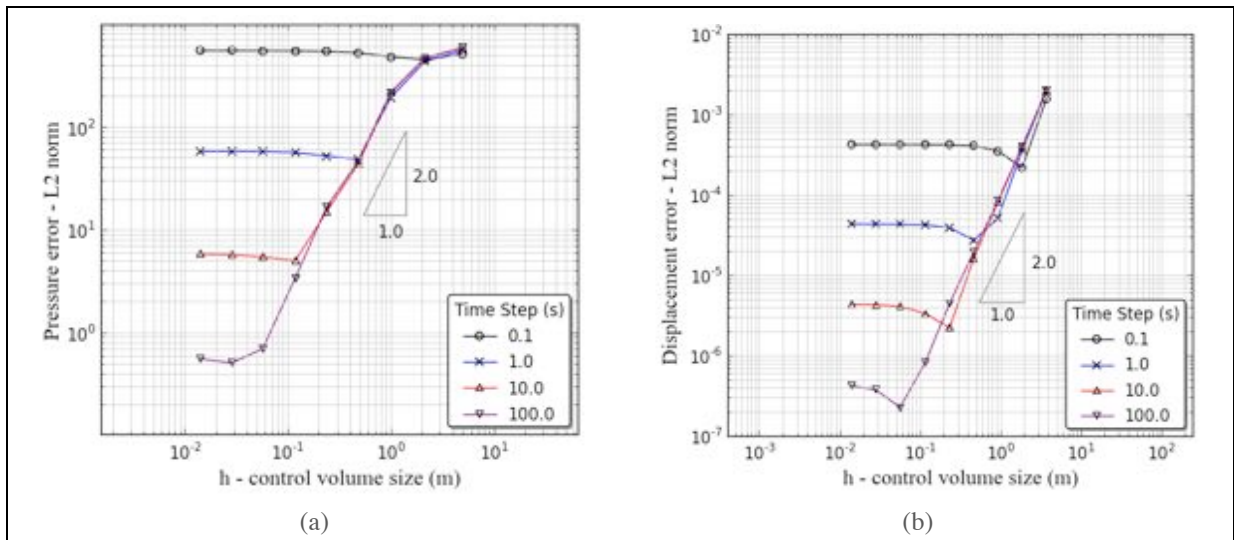


Figure 9 Convergence analysis of the staggered scheme for highly non-uniform grids. (a) Pressure and (b) displacement.

The behavior of the pressure and displacement error as the grid is refined is presented in figures (8) for slightly non-uniform grids. As can be seen, a second order decay of the error is obtained for both pressure and displacement. For highly non-uniform grid Figure (9b) shows second order of accuracy for displacement, while Figure (9a) suggests that the pressure is somehow affected by the grid non-uniformity, but it still can be regarded as a second order approximation.

Numerical Instabilities in the Pressure Field

As shown in the previous section, the staggered formulation is second-order accurate for pressure and displacement, even for highly non-uniform grids. It is well known that equal-order approximations for both pressure and displacement can cause numerical instabilities during undrained consolidation, where the fluid velocity is nearly zero. In this section, the same problem (Terzaghi's column) is solved with a time step size of 0,1 seconds and the

solution taken at $t = 1,0$ second. In this situation the fluid does not have enough time to move as the solid matrix deforms, inducing an undrained consolidation. In fact, pressure instabilities are expected for equal-order formulations when the time step size is small enough to violate the minimum time step criteria postulated in [12]. This problem is solved with four different grids summarized as below

Grid 1	16 nodes	Slightly non-uniform
Grid 2	16 nodes	Highly non-uniform
Grid 3	32 nodes	Slightly non-uniform
Grid 4	32 nodes	Highly non-uniform

Table 1: Types of grids.

The numerical and analytical profiles are depicted in the figures below. The numerical solutions of the pressure and displacement profiles are obtained by both staggered and collocated arrangement of variables. Both numerical solutions are compared with the analytical one.

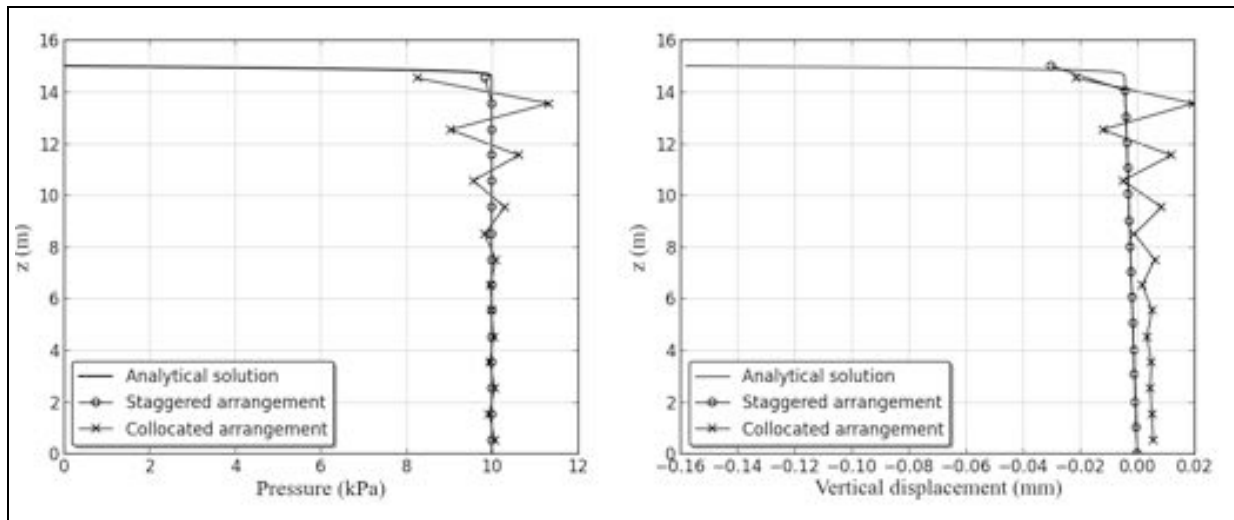


Figure 10 Pressure and displacement profiles for grid 1 (16 nodes).

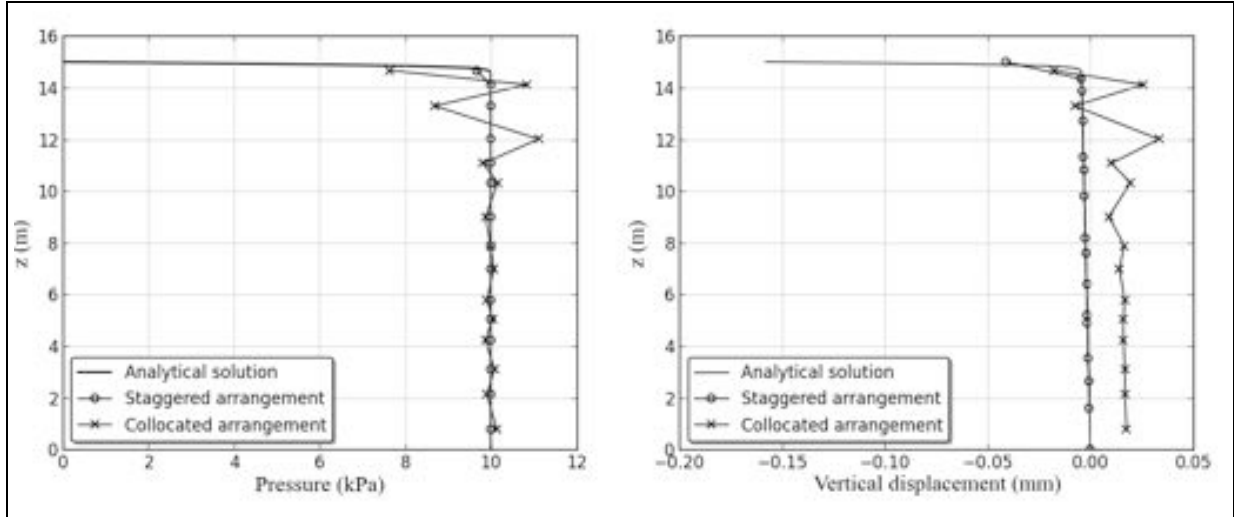


Figure 11 Pressure and displacement profiles for grid 2 (16 nodes).

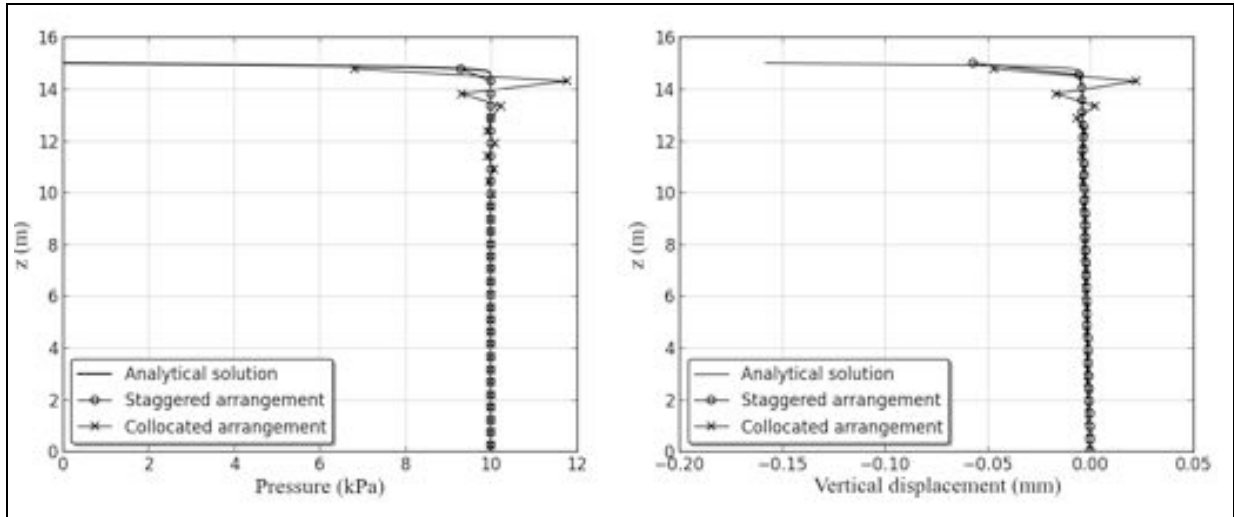


Figure 12 Pressure and displacement profiles for grid 3 (32 nodes).

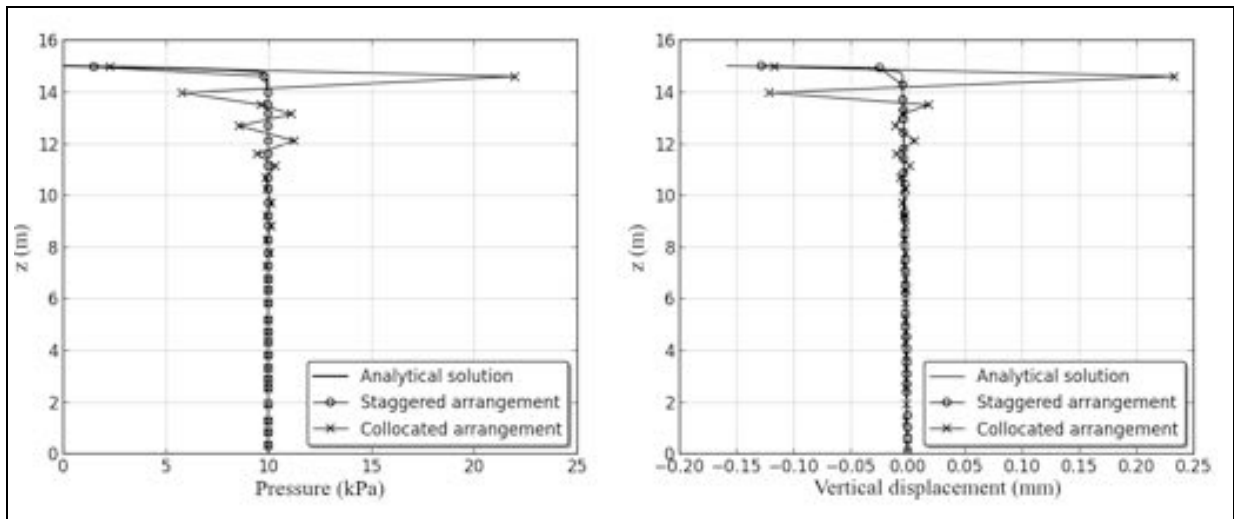


Figure 13 Pressure and displacement profiles for grid 4 (32 nodes).

As it can be seen in figures (10-13), the collocated formulation shows numerical instabilities also for the displacement field, what would be expected, since it directly depends on the

pressure gradient. For the grids with 16 nodes, figures (10) and (11) show an unacceptable solution obtained by the collocate formulation for the vertical displacement, which deteriorates even more with the grid non-uniformity. For the more refined grids (32 nodes), the numerical instabilities of the collocated arrangement concentrate along the upper-middle of the domain and they also get worse with the grid non-uniformity.

The staggered formulation, by its turn, despite presenting the same order of approximation for both pressure and displacement, does not show numerical instabilities at all. Instead, an excellent agreement is verified with the analytical solution. In addition, the staggered formulation for randomly spaced grids does not introduce any numerical diffusion, which is a common drawback of stabilization techniques commonly employed to equal-order formulations.

Conclusions

In this work, a two-dimensional finite volume formulation has been presented for modeling the coupled fluid flow/geomechanics using staggered arrangement of variables for unstructured grids. An equivalent one-dimensional formulation for randomly spaced grids has been implemented and tested. The results show an overall second order accuracy for both pressure and displacement, even for highly non-uniform grids. Despite this fact, it has been shown that the staggered formulation does not present any numerical instability during critical situations and does not introduce any numerical diffusion to the solution. If this conclusion persists for 2D and 3D cases, and we believe it will since the 1D formulations carries the key reasons for the stability, the formulation advanced will generate poroelasticity solutions free from spurious oscillations, free from numerical diffusion and keeping second order of accuracy for both pressure and displacements, all embodied in a conservative scheme for mass, momentum and equilibrium equations.

References

- [1] Ferronato, M., Castelletto, N. and Gambolati, G.: "A fully-coupled 3-D mixed finite element model of Biot consolidation", *Journal of Computational Physics*, vol. 229, pp. 4813-4830, 2010.
- [2] Choo, J. and Borja, R.I.: "A stabilized mixed finite elements for deformable porous media with double porosity", *Computer Methods in Applied Mechanics and Engineering*, vol. 293, pp. 131-154, 2015.
- [3] Preisig, M. and Prévost, J.H.: "Stabilization procedures in coupled poromechanics problems: A critical assessment", *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 35, pp. 1207-1225, 2011.
- [4] Honório, H.T. and Maliska, C.R.: "A stabilized finite volume method for solving one-dimensional poroelastic problems", *Rio Oil and Gas, Expo and Conference 2016 Proceedings*, Rio de Janeiro, 2016.
- [5] Harlow, F.H. and Welch, J.E.: "Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface", *Physics of Fluids*, vol. 8, pp. 2182-2189, 1965.
- [6] dal Pizzol, A. and Maliska, C.R.: "A finite volume method for the solution of fluid flows coupled with the mechanical behavior of compacting porous media", *Porous Media and its Applications in Science, Engineering and Industry AIP Proc.*, 1453, pp. 205-210, 2012.
- [7] Terzaghi, K.: "Die berechnung der durchlässigkeitsziffer des tones aus dem verlauf der hydrodynamischen spannungsercheinungen", *Akademie der Wissenschaften, Wien Mathematisch-Naturwissenschaftliche Klasse*, 1923.
- [8] Patankar, S.V.: "Numerical Heat Transfer and Fluid Flow", Hemisphere Publishing Corporation, 1980.
- [9] Peters, R. and Maliska, C.R.: "A staggered grid arrangement for solving incompressible flows with hybrid unstructured meshes", *Numerical Heat Transfer, Part B Fundamentals*, vol. 71, pp. 50-65, 2017.
- [10] Cerbato, G., Hurtado, F.S.V., Silva, A.F.C. and Maliska, C.R.: "Analysis of gradient reconstruction methods on polygonal grids applied to petroleum reservoir simulation", *15th Brazilian Congress of Thermal Science and Engineering, ENCIT Proceedings*, Belém, 2014.

- [11] Aavastmark, I., Barkve, T., Bøe, O. and Mannseth, T.: " Discretization on unstructured grids for inhomogeneous, anisotropic media. Part I, Derivation of the methods", SIAM Journal on Scientific Computing, vol. 19, No. 5, pp. 1700-1716, 1998.
- [12] Vermeer, P.A. and Verruijt, A.: " An accuracy condition for consolidation by finite elements", International Journal of Numerical and Analytical Methods in Geomechanics, vol. 5, pp. 1-14, 1981.

SPH modelling of consolidation problem based on two-phase mixture theory

Hao Wu¹, Jinjian Chen¹, Chencong Liao¹, Jianhua Wang¹, Jian Wang²

¹Department of civil engineering,
Shanghai Jiao Tong University, P.R.China

²Department of Water Conservancy and Hydropower Engineering,
Hohai University, P.R.China
Corresponding author: chulizi_hhu@hotmail.com

Abstract

The SPH based two-phase mixture model was improved by taking the soil porosity into account. The soil porosity was treated as a spatial variable but not a constant. A servo-control method was developed to model the stress boundary condition based on the frictional sliding contact algorithm. Then a 2-D consolidation numerical analysis was conducted to validate the ability of the SPH based two-phase model to predict the pore water pressure. Comparison with previous research proved that the SPH based two-phase mixture could capture the pore water pressure satisfactorily. In addition, the servo-control method could model the stress boundary condition well.

Keywords: SPH; Porosity; Two-phase mixture; Pore water pressure; Servo-control

1. Introduction

The well-known Biot-Zienkiewicz consolidation theory has been widely applied to consolidation and seepage problems in geotechnical engineering field [1]. Generally speaking, the Biot-Zienkiewicz theory focus on the soil deformation or the stationary state. The acceleration of the fluid phase is ignored and the soil-water interaction is taken into account implicitly. Few attention was paid to the flow process or the interaction between pure water region and the mixture. Therefore, the Biot-Zienkiewicz theory was rarely reported on studying the fast flows through high permeable porous media such as piping and scouring. By contrast, the two-phase mixture theory [2], in which soil and water are assumed to occupy part of the macroscopic mixture, is suitable for dealing with these problems. In the two-phase mixture theory, the soil and water satisfy their own governing equations and the interaction force is composed of pore water pressure and viscous drag force. Hence, not only the soil-water interaction within the mixture but also the interaction between the pure fluid region and the mixture could be investigated. Detailed theoretically comparison between the mixture theory and the Biot-Zienkiewicz theory could be found in Coussy [3].

Although extremely large soil deformation has been encountered when solving the abovementioned problems, smoothed particle dynamics (SPH) has been utilized to avoid the mesh distortion [4]. Recently, an amount of numerical studies has been reported on problems involving fast flow through porous media [5] and large soil deformation using SPH. Typical examples include saturated soil excavation by water jet [6, 7] and liquefaction problems [8–10]. However, the applicability of the SPH based two-phase mixture model has not been sufficiently validated. The most troublesome one is the ability to predict the pore water pressure. Different from the Biot-Zienkiewicz theory, in the two-phase mixture theory, the conservation equations of fluid phase are solved separately but not combined to the soil phase. The pore water pressure is in fact calculated through density variation. In this study, the ability of the SPH based two-phase mixture to capture the pore water pressure is validated through a classical 2-D

consolidation problem. In addition, the original SPH mixture model was improved by considering the effect of the porosity. Besides, a servo-control algorithm has been developed to model the stress boundary condition. It is proved that the SPH based two-phase model could also predict the pore water pressure satisfactorily and the servo-control method could model the stress boundary condition well.

2. SPH background

In SPH, the computation domain is discretized by a finite number of particles, which carry field variables and material properties [11–13]. All the field variables and functions are interpolated on the particles and governing equations could be solved. The final particle approximation form of function and its derivative are given by

$$\langle f(\mathbf{x}_i) \rangle = \sum_{j=1}^N \frac{m_j}{\rho_j} f(\mathbf{x}_j) W(\mathbf{x}_i - \mathbf{x}_j, h) \quad (1)$$

$$\langle \nabla \cdot f(\mathbf{x}_i) \rangle = - \sum_{j=1}^N \frac{m_j}{\rho_j} f(\mathbf{x}_j) \cdot \nabla_j W(\mathbf{x}_i - \mathbf{x}_j, h) \quad (2)$$

where i and j denote particles; N is the total number of neighbor particles; m is mass and ρ is the density; m_j/ρ_j actually gives the finite volume ΔV_j that originates in the infinitesimal volume $d\mathbf{x}'$. W is the kernel or smoothing function; h is the smoothing length defining the influence domain of W . Of the several proposed kernels, we apply here the Wendland type [14] for its accuracy and efficiency,

$$W(q, h) = \alpha_d \times \begin{cases} (1 - q/2)^4 (2q + 1), & 0 \leq q \leq 2, \\ 0, & q \geq 2 \end{cases} \quad (3)$$

where α_d is the normalizing factor, $\alpha_d = 7/4\pi h^2$ for 2-D problems; q is relative distance, $q = |\mathbf{x} - \mathbf{x}'|/h$.

3. Governing equations and SPH formulations

The two-phase mixture theory is based on the assumption that each constitute occupies part of the macroscopic mixture [2]. The mass conservations are given in the following Lagrangian forms

$$\frac{d\rho_s}{dt} = -\rho_s \nabla \cdot \mathbf{v}_s, \quad \rho_s = (1-n)\tilde{\rho}_s \quad (4)$$

$$\frac{d\rho_f}{dt} = -\rho_f \nabla \cdot \mathbf{v}_f, \quad \rho_f = n\tilde{\rho}_f \quad (5)$$

where n is the soil porosity; $\tilde{\rho}_s$ is the particle density of soil and $\tilde{\rho}_f$ is the intrinsic density of the water; ρ_s and ρ_f are the apparent density of soil and water, respectively; \mathbf{v}_s and \mathbf{v}_f are the spatially averaged velocity of soil and water, respectively. Assuming that the particle density of soil keeps unchanged, the governing equation for the soil porosity could be obtained by

$$\frac{dn}{dt} = (1-n) \nabla \cdot \mathbf{v}_s \quad (6)$$

It could be seen from equation (6) that the soil porosity was treated as a spatial and temporal field variable but not a constant. Hence, the effect of soil porosity on the mixture behavior could

be considered. In contrast, the soil porosity was either neglected or treated as a constant in previous studies, which was not in accord with the reality.

The conservation equation of momentum are given as

$$\rho_s \frac{d\mathbf{v}_s}{dt} = \nabla \cdot \boldsymbol{\sigma}' - (1-n) \nabla \tilde{p}_f + \mathbf{f}_d + \rho_s \mathbf{g} \quad (7)$$

$$\rho_f \frac{d\mathbf{v}_f}{dt} = -\nabla(n\tilde{p}_f) + \tilde{p}_f \nabla n + \nabla \cdot (n\tilde{\boldsymbol{\tau}}_f) - \mathbf{f}_d + \rho_f \mathbf{g} \quad (8)$$

$$\mathbf{f}_d = \frac{n^2 \tilde{\rho}_f \mathbf{g} (\mathbf{v}_f - \mathbf{v}_s)}{k} \quad (9)$$

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}' - (1-n) \tilde{p}_f \mathbf{I} + n\tilde{\boldsymbol{\sigma}}_f \quad (10)$$

$$\tilde{\boldsymbol{\sigma}}_f = -\tilde{p}_f \mathbf{I} + \tilde{\boldsymbol{\tau}}_f \quad (11)$$

where $\boldsymbol{\sigma}$ is the total stress tensor decomposed based on the Terzaghi's concept of effect stress. $\boldsymbol{\sigma}'$ is the effective stress relating to the strain rate in the constitutive model for soil. \tilde{p}_f and $\tilde{\boldsymbol{\tau}}_f$ are the pore water pressure and shear stress of water, respectively. \mathbf{g} is the gravity acceleration. \mathbf{f}_d is the viscous drag force calculated by Darcy's law. k is the hydraulic conductivity.

By applying the particle approximation formulation, equation (5)-(9) could be rewritten in the following SPH form,

$$\frac{d\rho_a}{dt} = \sum_{b=1}^M m_b \mathbf{v}_{ab} \cdot \nabla_a W_{ab} + \delta_f h_a c_f \sum_{b=1}^M \frac{m_b}{\rho_b} \Psi_{ab} \cdot \nabla_a W_{ab} \quad (12)$$

$$\frac{dn_i}{dt} = (1-n_i) \sum_{j=1}^N \frac{m_j}{\rho_j} \mathbf{v}_{ji} \cdot \nabla_i W_{ij} \quad (13)$$

$$\frac{d\mathbf{v}_i}{dt} = \sum_{j=1}^N m_j \left(\frac{\boldsymbol{\sigma}'_i}{\rho_i^2} + \frac{\boldsymbol{\sigma}'_j}{\rho_j^2} + \Pi_{ij} \right) \cdot \nabla_i W_{ij} - (1-n_i) \sum_{a=1}^M \frac{m_a}{\rho_a} \frac{\tilde{p}_a}{\rho_i} \nabla_i W_{ai} + \sum_{a=1}^M \frac{m_a}{\rho_a} \frac{\mathbf{f}_{ia}}{\rho_i} W_{ai} + \mathbf{g}_i \quad (14)$$

$$\frac{d\mathbf{v}_a}{dt} = -\sum_{b=1}^M m_b \left(\frac{\tilde{p}_a n_a}{\rho_a^2} + \frac{\tilde{p}_b n_b}{\rho_b^2} + \Pi_{ab} + f_{ab}^g (R_a + R_b) \right) \nabla_a W_{ab} + \sum_{b=1}^M m_b \left(\frac{\boldsymbol{\tau}_a n_a}{\rho_a^2} + \frac{\boldsymbol{\tau}_b n_b}{\rho_b^2} \right) \cdot \nabla_a W_{ab} \quad (15)$$

$$+ \sum_{i=1}^N \frac{m_i}{\rho_i} \frac{\tilde{p}_a}{\rho_a} n_i \nabla_a W_{ai} - \sum_{i=1}^N \frac{m_i}{\rho_i} \frac{\mathbf{f}_{ia}}{\rho_a} W_{ai} + \mathbf{g}_a$$

$$\mathbf{f}_{ia} = \frac{n_a^2 \tilde{\rho}_f \mathbf{g} (\mathbf{v}_a - \mathbf{v}_i)}{k} \quad (16)$$

where i, j denote soil particles and a, b denote water particles. $\mathbf{v}_{ab} = \mathbf{v}_a - \mathbf{v}_b$, $\mathbf{v}_{ji} = \mathbf{v}_j - \mathbf{v}_i$. The second term on the right side of equation (12) is added to avoid density fluctuation and to obtain accurate pore water pressure, which is based on the δ -SPH method [15].

$\Psi_{ab} = 2(\rho_a - \rho_b) \frac{\mathbf{x}_{ab}}{|\mathbf{x}_{ab}|^2}$. δ_f is a constant normally set to 0.1, c_f is the sound speed of water. Π_{ij}

and Π_{ab} are Monaghan-type artificial viscosity [16] used to remove unphysical penetration, defined as

$$\Pi_{ij} = \begin{cases} \frac{-\alpha_{\Pi} c_{ij} \Phi_{ij} + \beta_{\Pi} \Phi_{ij}^2}{\rho_{ij}}, & \mathbf{v}_{ij} \cdot \mathbf{x}_{ij} < 0 \\ 0, & \mathbf{v}_{ij} \cdot \mathbf{x}_{ij} \geq 0 \end{cases} \quad (17)$$

$$\Phi_{ij} = \frac{h_{ij} \mathbf{v}_{ij} \cdot \mathbf{x}_{ij}}{|\mathbf{x}_{ij}|^2 + 0.01 h_{ij}^2} \quad (18)$$

$$c_{ij} = \frac{c_i + c_j}{2}, \rho_{ij} = \frac{\rho_i + \rho_j}{2}, h_{ij} = \frac{h_i + h_j}{2} \quad (19)$$

Π_{ab} could be obtained by simply replace i, j with a, b . α_{Π} and β_{Π} are respectively set to 0.1 and 1.0 for soil, while for water take values of 0.01 and 1.0.

In this research, serious tensile instability was observed in low permeable soil. The artificial pressure method proposed by Monaghan [17] has been adopted throughout this study, i.e. the term $f_{ab}^g(R_a + R_b)$. f_{ab} is the repulsive term and specified by $f_{ab} = \frac{W_{ab}}{W(\Delta d, h)}$. Δd denotes the initial particle spacing, g is usually taken as $W(0, h)/W(\Delta d, h)$. For Wendland kernel, n has the value about 3.24 with h equals to $1.2\Delta d$. The factor R_a and R_b are determined in terms of pressure,

$$R_a = \begin{cases} \frac{\ell |\tilde{p}_a|}{\rho_a^2}, & \text{if } \tilde{p}_a < 0 \\ 0, & \text{otherwise} \end{cases} \quad (20)$$

where ℓ is a small constant and typically taken as 0.2; R_b is calculated analogously.

To close the above equations, constitutive models are needed to determine σ' , \tilde{p}_f , $\tilde{\tau}_f$. In order to keep consistent with the results to be compared in Boer et al. [18] and Breuer [19], an elastic constitutive relationship is adopted. The water is considered as weakly compressible Newtonian fluid. The final SPH discretized constitutive model for soil and water are given as follows.

For soil,

$$\dot{\sigma}_i^{\alpha\beta} = 2G\dot{\varepsilon}_i^{\alpha\beta} + K\dot{\varepsilon}_i^{\gamma\gamma}\delta_i^{\alpha\beta} \quad (21)$$

$$\dot{\varepsilon}_i^{\alpha\beta} = \frac{1}{2} \left(\sum_{j=1}^N \frac{m_j}{\rho_j} v_{ji}^\alpha \frac{\partial W_{ij}}{\partial x_i^\beta} + \sum_{j=1}^N \frac{m_j}{\rho_j} v_{ji}^\beta \frac{\partial W_{ij}}{\partial x_i^\alpha} \right) \quad (22)$$

For water,

$$\tilde{p}_a = B \left[\left(\frac{\tilde{\rho}_a}{\tilde{\rho}_{f0}} \right)^\chi - 1 \right] \quad (23)$$

$$\tau_a^{\alpha\beta} = \mu \varepsilon_a^{\alpha\beta} \quad (24)$$

$$\varepsilon_a^{\alpha\beta} = \sum_{b=1}^M \frac{m_b}{\rho_b} v_{ba}^\alpha \frac{\partial W_{ab}}{\partial x_a^\beta} + \sum_{b=1}^M \frac{m_b}{\rho_b} v_{ba}^\beta \frac{\partial W_{ab}}{\partial x_a^\alpha} - \frac{2}{3} \left(\sum_{b=1}^M \frac{m_b}{\rho_b} \mathbf{v}_{ba} \cdot \nabla_a \mathbf{W}_{ab} \right) \delta^{\alpha\beta} \quad (25)$$

where α and β denote the Cartesian components x , y , or z . $\delta^{\alpha\beta}$ is the Kronecker delta symbol. G is the shear modulus and K is the bulk modulus, respectively given by $G = \frac{E}{2(1+\nu)}$, $K = \frac{E}{3(1-2\nu)}$. E is Young's modulus and ν is the Poisson's ratio. $\varepsilon_i^{\alpha\beta}$ is the deviatoric strain rate tensor, $\varepsilon_i^{\alpha\beta} = \varepsilon_i^{\alpha\beta} - \frac{1}{3} \varepsilon_i^{\gamma\gamma} \delta^{\alpha\beta}$. B is a problem dependent parameter that sets

a limit to maximum density variation. χ is a constant normally set to 7. $\tilde{\rho}_{f0}$ is the reference intrinsic density of water. μ is the dynamic viscosity of water.

4. Boundary contact and servo-control method

Boundary deficiency is an inherent drawback of SPH. For particles moving near or on the boundary, the support domain is incomplete and the calculated acceleration is not accurate. Several attempts have been tried in previous studies. The boundaries of rigid have been modeled using a) ghost particles, b) fluid particles, c) normalizing conditions, d) boundary particle force and e) particle-to-particle or particle-to-surface contact based on momentum equations [13]. Regrettably, the above methods can only be applied to completely smooth or rough boundary conditions. Here we adopt the frictional sliding contact algorithm proposed by Wang et al [20, 21] to simulate the contact between the mixture and boundary. The final form of the contact force is given by

$$\mathbf{F}_n = (1 - \varsigma) \left[\frac{2m_i}{(\Delta t)^2} (d_0 + \mathbf{G} \cdot \mathbf{n}) \right] \mathbf{n} \quad (26)$$

$$\mathbf{F}_\tau = \begin{cases} \frac{\xi |\mathbf{F}_n|}{|\mathbf{F}'_\tau|} \mathbf{F}'_\tau, & \text{if } |\mathbf{F}'_\tau| > \xi |\mathbf{F}_n| \\ \mathbf{F}'_\tau, & \text{otherwise} \end{cases} \quad (27)$$

$$\mathbf{F}'_\tau = \frac{2m_i}{(\Delta t)^2} (\Delta \mathbf{u} - \Delta \mathbf{u} \cdot \mathbf{n}) \quad (28)$$

where \mathbf{F}_n , \mathbf{F}_τ are the normal, tangent component of the contact force, respectively. ς defines the extent of penetration allowed and was taken as 0.01-0.1. ξ is the frictional coefficient. \mathbf{G} is the vector from the particle to its perpendicular foot on the boundary. \mathbf{n} is the outward normal vector of the contact surface.

Different from the single phase or quasi-single phase theory, stress boundary condition can not be applied to the mixture. It is because that the portion of the external load carried by each phase is uncertain. As shown in Fig.1, a novel method base on the above contact algorithm is developed here to simulate the stress boundary.

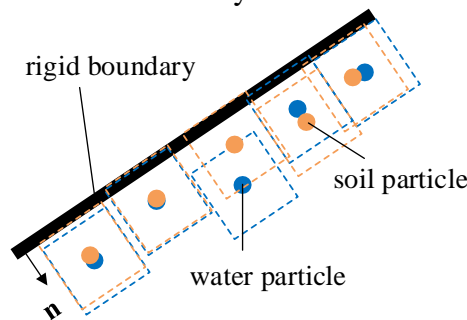


Fig.1 Schematic diagram of servo-control method

The rigid boundary is assigned velocity along the outward normal vector through

$$v_R^n = \frac{\alpha_{cs} A_{cs} \Delta t}{2 \left(\sum_{i=1}^{N_s} m_i + \sum_{a=1}^{N_f} m_a \right)} (\sigma_g - \sigma_m) \quad (29)$$

$$\sigma_m = \frac{\sum_{i=1}^{N_s} F_i + \sum_{a=1}^{N_f} F_a}{A_{cs}} \quad (30)$$

where σ_g is the target stress, σ_m is the measured contact stress in the present time step; m_i and m_a represent the mass of soil and water, respectively. N_s and N_f are respectively the total number of soil and water particles contacting with the rigid. A_{cs} is the contact area. α_{cs} is a scaling factor used to weaken the oscillation, taken as 0.01. F_i and F_a are calculated through equation (26).

5. 2-D consolidation modelling and analysis

A 2-D consolidation problem was simulated to validate the two-phase mixture SPH model and the servo-control algorithm. Boer *et al.* [18] and Breuer [19] studied the same problem by FEM. The parameters used here were taken the same for comparison. The geometry of the model is shown in Fig.2. The soil-water mixture was 20.0 m long and 10.0 m wide, with 15 KPa uniformly distributed load at the top. The left, right and the bottom were fixed and undrained, whereas the top was drained. The soil parameters are: $\lambda_s = 5583$ KPa, $\mu_s = 8375$ KPa, $\tilde{\rho}_s = 2000$ kg/m³, $n = 0.33$, $\nu = 0.2$, $k = 0.01$ m/s. The water parameters are: $\tilde{\rho}_f = 1000$ kg/m³, $c_f = 97.5$ m/s.

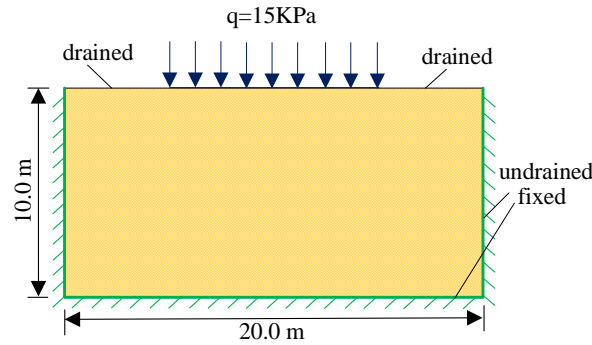


Fig.2 Geometry of the 2-D consolidation model

The initial SPH model is shown in Fig.3. Totally 800 particles were used with initial resolution $\Delta d = 0.5$ m. The left, right and bottom were all modeled as non-slip using ghost particles. The velocity of the rigid boundary at the top was assigned velocity through equation (29) to model the stress boundary condition. Soil and water particles were initially superimposed and then moved separately according to their own governing equations.

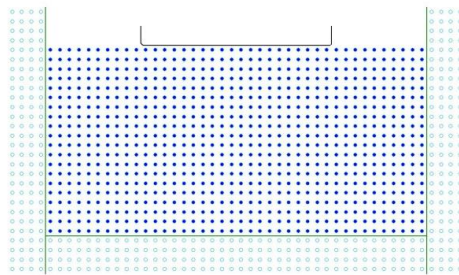


Fig.3 Initial SPH model of the 2-D consolidation problem

The evaluation of the excess pore water pressure at different intervals of time is shown in Fig.4. FEM results by Breuer [19] are also included for comparison (on the left, in KPa). It is shown that the excess pore water pressure predicted by SPH corresponds to FEM results well. In the beginning, i.e. $t = 0.01$ s, the excess pore water pressure increased to around 14000 Pa quickly. The reason is that the deformation of soil lagged behind the water. Accordingly, the whole external load was mainly carried by the water. With the passage of time, the pore water flow out of the void and the soil skeleton carried more and more external load. As the result, the excess pore water pressure decreased gradually. After 10 s the excess pore water pressure was about zero.

It was proved that the excess pore water pressure could be captured satisfactorily by means of the proposed method. Besides, the servo-control algorithm method could be used to simulate the stress boundary condition.

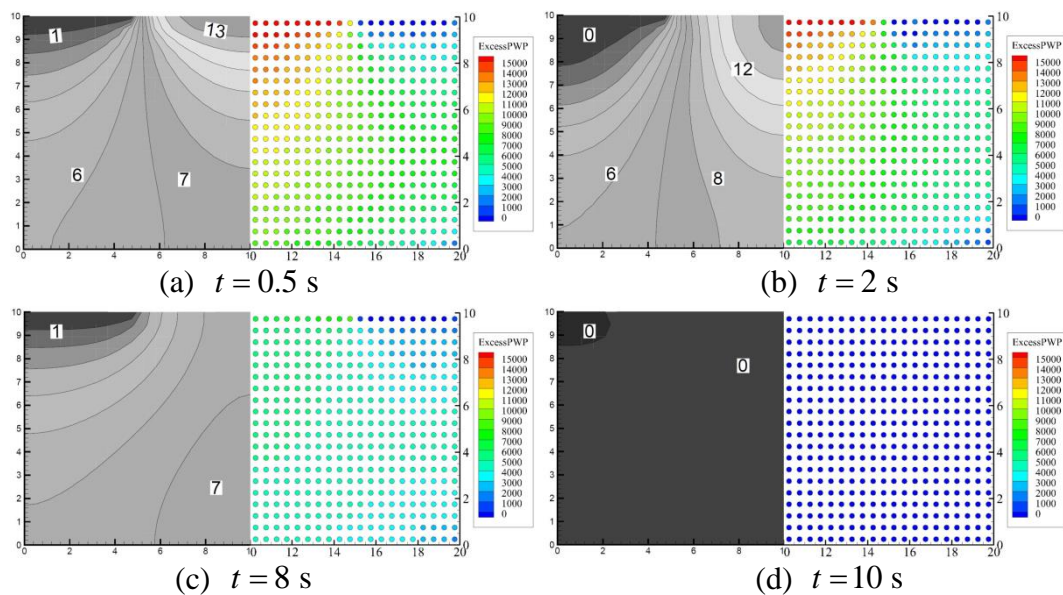


Fig 4 Comparison of excess pore water pressure between FEM (left: in KPa) and SPH (right: in Pa) at different intervals of time

6. Conclusions

The SPH based two-phase mixture model has been recently applied to geotechnical problems involving fast flow through porous media and large soil deformation. However, the applicability of the SPH based two-phase mixture model to the evaluation of pore water pressure has not been validated.

In this study, the SPH based two-phase mixture model was first improved by taking the soil porosity into account. The soil porosity was treated as a spatial variable but not a constant and was interpolated and integrated at all particles. Soil and water particles were superimposed and then moved separately according to their own governing equations. The interaction force of the two phases was composed of pore water pressure and viscous drag force. Tensile instability was properly handled by using the Monaghan's artificial pressure method. Then a servo-control method was proposed based on the frictional sliding contact algorithm in order to model the stress boundary condition. Finally, a 2-D consolidation numerical test was conducted and

compared to previous research. It was proved that the SPH based two-phase mixture model could satisfactorily predict the pore water pressure. Besides, the servo-control method could model the stress boundary condition well.

Acknowledgement

This research is supported by the National Natural Science Foundation of China (Grant Nos. 51678360, 41727802 and 51779084).

References

- [1] Biot MA (1956) Theory of Propagation of Elastic Waves in a Fluid-Saturated Porous Solid. I. Low-Frequency Range. *J Acoust Soc Am* 28:168–178 .
- [2] Drew DA (1983) Mathematical Modeling of Two-Phase Flow. *Annu Rev Fluid Mech* 15:261–291
- [3] Coussy O, Dormieux L, Detournay E (1998) From mixture theory to biot's approach for porous media. *Int J Solids Struct* 35:4619–4635 .
- [4] Wu H, Wang J, Wang JH, Liao CC (2017) Asymmetric Adaptive Particle Refinement in SPH and Its Application in Soil Cutting Problems. *Int J Comput Methods* 1850052
- [5] Peng C, Xu G, Wu W, et al (2017) Multiphase SPH modeling of free surface flow in porous media with variable porosity. *Comput Geotech* 81:239–248 .
- [6] Bui HH, Sako K, Fukagawa R (2007) Numerical simulation of soil-water interaction using smoothed particle hydrodynamics (SPH) method. *J Terramechanics* 44:339–346 .
- [7] Wang C, Wang Y, Peng C, Meng X (2016) Smoothed Particle Hydrodynamics Simulation of Water-Soil Mixture Flows. *J Hydraul Eng* 142:4016032 . doi: 10.1061/(ASCE)HY.1943-7900.0001163
- [8] Huang Y, Zhang W, Dai Z, Xu Q (2013) Numerical simulation of flow processes in liquefied soils using a soil–water-coupled smoothed particle hydrodynamics method. *Nat Hazards* 69:809–827 .
- [9] Maeda K, Sakai H, Sakai M (2006) Development of Seepage Failure Analysis Method of Ground With Smoothed Particle Hydrodynamics. *Struct Eng / Earthq Eng* 23:307s–319s .
- [10] Nakamura K, Satomi T, Takahashi H (2014) Improved Model for Soil as a Two-Phase Mixture Based on Smoothed Particle Hydrodynamics (SPH). 1053–1060
- [11] Liu MB, Liu GR, Zong Z (2008) An overview of smoothed particle hydrodynamics. 1:135–188
- [12] Liu MB, Shao J, Chang J (2012) On the treatment of solid boundary in smoothed particle hydrodynamics. *Sci China Technol Sci* 55:244–254 .
- [13] Monaghan JJ, Kajtar JB (2009) SPH particle boundary forces for arbitrary boundaries. *Comput Phys Commun* 180:1811–1820 .
- [14] Wendland H (1995) Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv Comput Math* 4:389–396 .
- [15] Antuono M, Colagrossi A, Marrone S, Molteni D (2010) Free-surface flows solved by means of SPH schemes with numerical diffusive terms. *Comput Phys Commun* 181:532–549 .
- [16] Monaghan JJ (1995) SPH simulation of multi- phase flow. *Comput Phys Commun* 87:1–2
- [17] Monaghan JJ (2000) SPH without a Tensile Instability. *J Comput Phys* 159:290–311 .
- [18] Boer R, Ehlers W, Liu Z (1993) One-dimensional transient wave propagation in fluid-saturated incompressible porous media. *Arch Appl Mech* 63:59–72 . doi: 10.1007/BF00787910
- [19] Breuer S (1999) Quasi-Static and Dynamic Behavior of Saturated Porous Media with Incompressible Constituents. *Transp Porous Media* 34:285–303 .
- [20] Wang J, Wu H, Gu C, Hua H (2013) Simulating frictional contact in smoothed particle hydrodynamics. *Sci China Technol Sci* 56:1779–1789 .
- [21] Wang J, Chan D (2014) Frictional contact algorithms in SPH for the simulation of soil-structure interaction. *Int J Numer Anal Methods Geomech* 38:747–770 .

CFD simulation of chemical gas dispersion under atmospheric boundary conditions

†*George XU¹, Arthur LIM¹, Harish GOPALAN¹, Jing LOU¹ and Hee Joo POH¹

¹Fluid Dynamics Department, Institute of High Performance Computing, A*STAR, Singapore.

*Presenting author: xu_xiangguo@ihpc.a-star.edu.sg

†Corresponding author: xu_xiangguo@ihpc.a-star.edu.sg

Abstract

Pollutant control is one of the key concerns in the design of buildings, for the sake of occupational health, safety and environment sustainability. In particular, risk analyses related to emergency leakage of chemicals from storage tanks or chemical processes have aroused increasing attentions in recent days, as well as the effectiveness of mitigation measures in order to eliminate, reduce and control the risks. In this paper, a CFD methodology with non-reactive chemical gases treated as passive scalars has been developed to simulate the gas dispersion across urban environments, subject to atmospheric wind conditions. Special treatments to maintain the consistency in atmospheric flow profiles, turbulence modeling and boundary conditions have also been accounted for. The developed model for gas dispersion has been implemented in the open source CFD code - OpenFOAM. The proposed methodology has been validated by modeling the gas dispersions for two urban-related test cases: the street canyon measured in a laboratory wind tunnel and the Mock Urban Setup Test (MUST) field experiment conducted in the desert area of Utah State. Effects of turbulent Schmidt number have been primarily addressed in this study. Statistical analyses about the discrepancies between predicted and experimental data have been carried out to quantify the accuracy of the proposed methodology. Simulation results from passive scalar transport equation demonstrate good agreement with experimental data, though tracer gases heavier than the atmospheric air were used in the both measurements. Furthermore, sensitivity tests also indicate that the accuracy of the simulation results is sensitive to the value of turbulent Schmidt number.

Keywords: CFD, gas dispersion, OpenFOAM, urban environment modeling

Introduction

In the design of residential, commercial, industrial or infra-structure buildings, pollutant control has become an important design feature to be addressed for the sake of occupational health, safety and environmental sustainability. The same concern has been extended for chemical plant design where emergency leakage of chemicals from storage tanks or chemical processes may take place. Risk analysis related to such emergency cases should be carefully investigated during the planning and design stages, and then effective mitigation measures should be proposed and evaluated to eliminate, reduce and control the risks. All these concerns are related to modeling of the gas dispersions.

Chemical gas dispersion in space is conventionally evaluated using empirical methods, such as Gaussian plume modeling [1]. Such empirical methods are very efficient, usually with the consideration of dynamic changes in atmospheric wind conditions. They have been widely adopted to study the impacts of plumes out of chimneys or vent shafts upon the environment within a large space. However, the obstructions due to buildings is one of the drawbacks, in particular near the ground area, cannot be accurately resolved with these methods. Recently, due to the rapid advance in computers, computational fluid dynamics (CFD) method has

become more attractive in assessing the environmental pollution, because its higher accuracy and richer field information. Comparatively, CFD method can be used to accurately predict the pollutant dispersion due to the obstructions of buildings, thus it is more suitable tool to simulate the pollutant dispersion in an urban environment that is highly occupied by high-density of buildings.

Multi-species modeling and passive scalar modeling are the two CFD methods that are usually adopted to simulate gas dispersions. The former can result in high-accuracy prediction, due to the well consideration of most physics relevant. However, it depends on the accuracy and full set of thermodynamic properties of chemical species. Besides, it is very time consuming in simulation. Comparatively, the passive scalar modeling is much more cost effective and less demand in input parameters, which is suitable for quick solutions for engineering problems.

As a part of the objectives for a government-funded project, a passive scalar modeling methodology for non-reactive chemical gases dispersion across urban environments has been developed, subject to atmospheric boundary conditions. The main objective in this study is to characterize the accuracy of passive scalar modeling methodology, when it is applied to simulate chemical gases dispersed in urban environments. Research efforts are limited to the development of efficient solver for dilute dispersion of chemical gases in spaces relatively away from the leakage site. In the proposed methodology, the leaked gas is represented by well-mixed volumetric clouds and the transport equation for passive scalar is adopted to trace the concentration of chemical gases. Special treatments to maintain the consistency in atmospheric flow profiles, turbulence modeling and boundary conditions have also been accounted for. All the development work has been implemented in OpenFOAM [2] – an open source CFD code.

In this paper, the proposed CFD methodology is to be first introduced briefly, followed by the in-depth discussion about the simulation results for two testing cases. Some remarks upon the proposed CFD methodology for gas dispersion modeling are finally summarized in the conclusion section.

CFD Methodology

In this study, steady-state incompressible flows under isothermal conditions are primarily focused. In addition to the continuity and momentum equations, the passive scalar transport equation has been chosen to model the concentration changes. The two-equation k- ϵ models are adopted to address the turbulence effects.

The governing equation for passive scalar transportation takes the following form:

$$\frac{\partial}{\partial t}(c) + \nabla \cdot (c\mathbf{u}) = \nabla \cdot [(D_c + \nu_t / Sc_t) \nabla c] + S_c \quad (1)$$

where c , \mathbf{u} and t denote the concentration of pollutant, air velocity and time, respectively; D_c and ν_t represent the molecular diffusivity of pollutant of concern and the turbulence kinematic viscosity; S_c is the source term; and Sc_t denote the turbulence Schmidt number.

In the proposed method, the chemical gas released from leakage region is modeled with the source term, S_c . Linearization of the source term, as proposed by Patankar [3] has been adopted during the solution procedure for the sake of improved numerical stability.

Second-order discretization schemes have been adopted to approximate the partial differential terms in governing equations. The resultant algebraic equations are solved with the well-known SIMPLE solution procedure.

Results and Discussion

The accuracy of the proposed methodology has been demonstrated with the solutions to two test cases: gas dispersions across street canyon in a laboratory scale and gas dispersion around a mock up setting test (MUST) under a realistic atmospheric scale. Details about the flow behaviors will be discussed in the following section.

Test Case 1: Street Canyon

The first test case is based on the wind tunnel test for a street canyon configuration that was designed and measured by the Laboratory of Building and Environmental Aerodynamics at the Institute for Hydromechanics (IfH) in the University of Karlsruhe, Germany [4]. As shown in Figure 1, the computational domain consists of two parallel buildings which form the street canyon of concern. Four sub cell zones on the street are chosen to represent the line sources for tracer gas (SF_6), emulating the chemicals emitted from traffic vehicles. The total release rate is 10g/s. Compliant with the test conditions, the power-law wind profile is adopted to represent the incoming wind perpendicular to the street.

Consistent with the measured data, the normalized concentration of tracer gas is calculated as

$$c^+ = \frac{cu_H H}{Q/l} \quad (2)$$

where u_H is the reference velocity at the building height, H ; Q and l denote the gas release rate and the length of the line sources, respectively. It should be noted that the values of c^+ measured on the leeward and windward surfaces are available for validation.

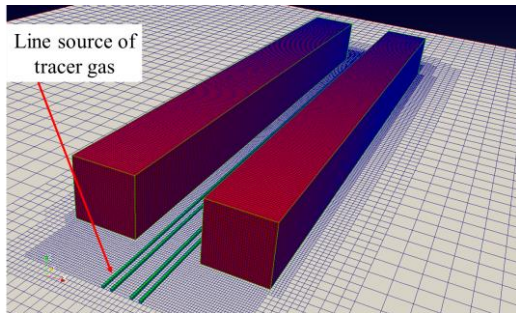


Figure 1. Computational domain and grid.

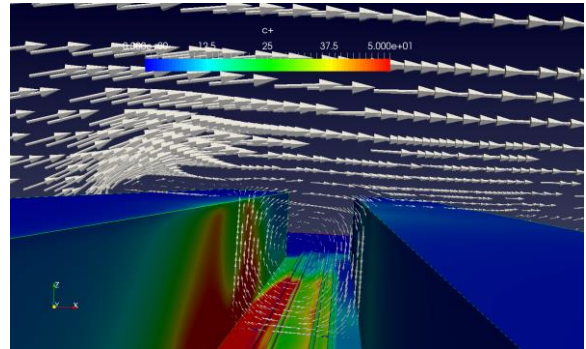


Figure 2. Airflow across the street canyon and contours of gas concentration on wall surfaces.

As shown in Figure 2, the incoming atmospheric wind, perpendicular to the street canyon, induces the recirculated air flow across the canyon, similar to the lid-driven cavity flow. The pollutant released from the sources on the ground turns to follow the airflow travel in space. The downward movement of the airflow along the windward surface turns to clear the pollutant from the respective surface. On the other hand, the recirculated air eventually carry the pollutant to travel towards the opposite building. It results in the higher concentration of the pollutant gas on the leeward surface.

The predicted gas concentration on the leeward and windward surfaces are presented in Figures 3 and 4, where the measured values are also included for comparison. It is apparent that good agreement between the numerical results and the measured values is achieved.

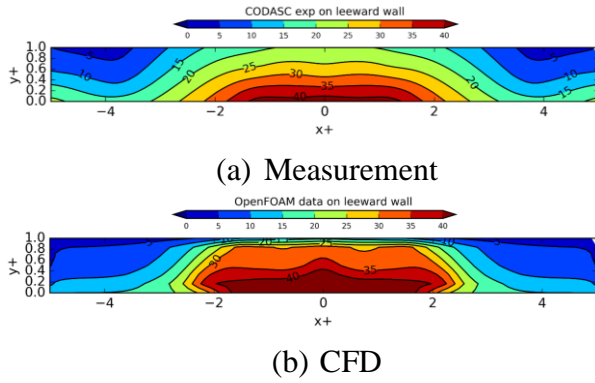


Figure 3. Contours of c^+ on the leeward surface of the street canyon.

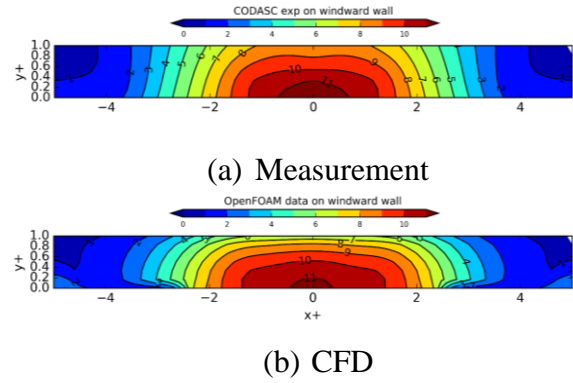


Figure 4. Contours of c^+ on the windward surface of the street canyon.

Sensitivity studies about different forms of $k-\varepsilon$ models, i.e. standard, realizable and RNG $k-\varepsilon$ models, have been investigated for the street canyon test case. It is found that standard $k-\varepsilon$ model yields the most accurate prediction. Besides, subject to the standard $k-\varepsilon$ turbulence model, additional studies about the impact of Sc_t value have also been performed.

Statistical analyses [5] about the predicted and measured values of c^+ on the two building surfaces of concern have been conducted. Fractional Bias (FB), Normalized Mean Square Error (NMSE) and Fraction of Predictions within a factor of two of the observations (FAC2) are selected to quantify the accuracy of the proposed CFD model. As summarized in Table 1, all the values for Sc_t that have been tested produce reasonable results. In addition, the most agreeable prediction with the measurement can be obtained when $Sc_t = 0.3$.

Table 1. Impacts of Sc_t on the gas dispersion.

Name	Definition	Model Perfects	Acceptable Model	Sc_t		
				0.7	0.3	0.2
FB	$2 \times (\bar{C}_0 - \bar{C}_p) / (\bar{C}_0 + \bar{C}_p)$	0	$-0.3 < FB < 0.3$	-0.39	-0.02	0.17
NMSE	$(\bar{C}_0 - \bar{C}_p)^2 / (\bar{C}_0 \times \bar{C}_p)$	0	$NMSE < 4$	0.06	0.09	0.11
FAC2	\bar{C}_0 / \bar{C}_p	1	$FAC2 > 0.5$	1.49	1.02	0.84

Test Case 2: MUST

The second case that has been simulated correspond to the airflow across the array of containers geometry of the Mock Urban Setting Test (MUST) field experiment. In the field test, 12 by 10 array of containers (each 12.2m long, 2.42m wide and 2.54m high) were deployed in desert of western Utah to mimic the urban environment. Details about the experimental studies are clearly addressed in DPG Document No. WDTC-FR-01-121 [6].

Test Scenario #2681829, as studied by Bekka *et. al.* [7], is chosen to be simulated with the proposed methodology. In the MUST field test, Propylene (CH_2CHCH_3) was used as tracer gas and released at different locations of concern. In the test scenario simulated here, the release rate of propylene is 225 l/min. Concentration of the released gas at 48 locations, as indicated by elliptic dots in Figure 5, are monitored and compared with the measured values. It should be noted that 40 receptors are located at 1.8m height and additional eight receptors are installed at various vertical heights on the tower located at the centre of the test field.

The neutral-state log-law wind and turbulence profiles, consistent with the atmospheric boundary conditions during the measurement, are implemented as the inflow conditions. To maintain the horizontal homogeneity, as recommended by Blocken [8], changes have been

made to the constants for k-ε turbulence model and wall functions and care has been taken in the grid generation in particular for the first-layer of grid on the ground. A total of 9.6-million hexa-dominated elements have been generated and used in simulations.

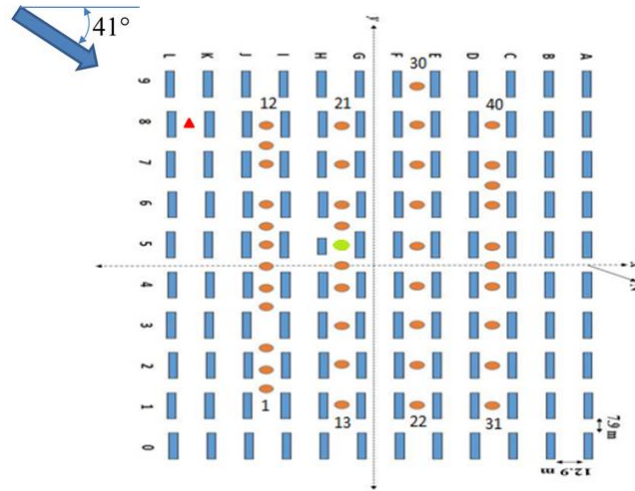


Figure 5. MUST site and locations of receptors for Test Scenario #2681829.
(Courtesy to Bekka *et. al.* [7] for the reuse of their image).

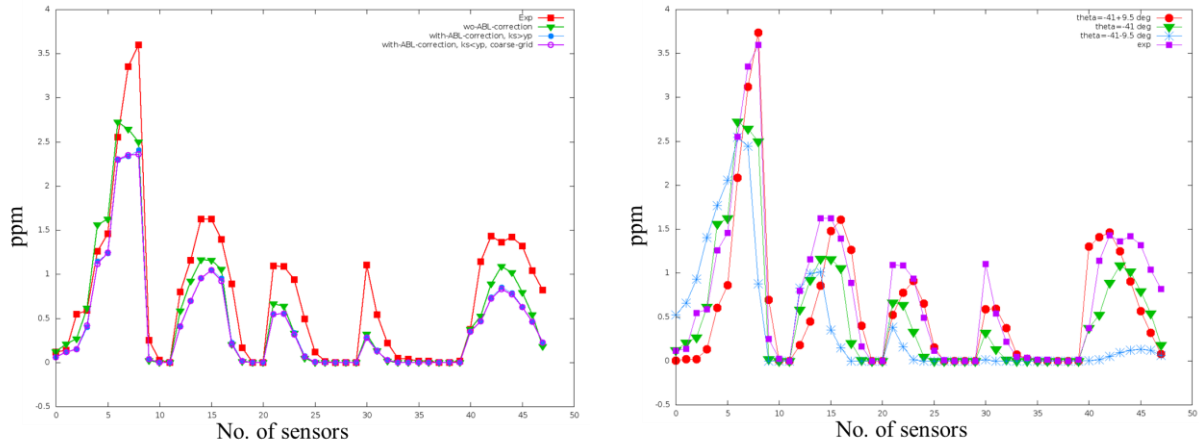
Sensitivity tests about the impacts of Sc_t have been carried out. As summarized in Table 2, predicted results show poor agreement with measured data when Sc_t takes the smaller value, and the most accurate prediction can be achieved when Sc_t becomes as high as 0.9. Thus most of simulations subsequently conducted are based on $Sc_t = 0.9$.

Table 2. Impacts of Sc_t on the gas dispersion.

Name	Definition	Model Perfects	Acceptable Model	Sc_t		
				0.3	0.7	0.9
FB	$2 \times (\bar{C}_0 - \bar{C}_p) / (\bar{C}_0 + \bar{C}_p)$	0	$-0.3 < FB < 0.3$	0.60	0.23	0.15
NMSE	$(\bar{C}_0 - \bar{C}_p)^2 / (\bar{C}_0 \times \bar{C}_p)$	0	$NMSE < 4$	0.87	0.60	0.73
FAC2	\bar{C}_0 / \bar{C}_p	1	$FAC2 > 0.5$	1.85	1.26	1.16

It is interesting to find that the without atmospheric boundary layer (ABL) correction as proposed by Blocken *et. al.* [5], predicted gas concentration show better agreement with experimental results, as shown in Figure 6(a). Besides, the simulated results, subject to ABL correction, seem insensitive to the relationship between sand grain roughness and the height of first-layer ground elements. All these findings are not consistent with the ABL treatment techniques [8]. Details for the reasons behind will be further investigated in future.

Sensitivity studies about the variation in wind direction for the test scenario have been conducted. As indicated in the measurement report [6], for the test scenario of concern here, the change in wind direction is $\pm 9.5^\circ$. Two additional simulations have been carried out according to the two extreme wind directions. As shown in Figure 6 (b), it is manifest that the predicted concentrations are sensitive to the change in the wind direction. Better agreement with measured data can be obtained when $\theta = (-41 + 9.5)^\circ$. Such sensitivity studies demonstrate the complexity in the uncertainties qualification of prediction models, when the real-life measurement data are used for validation.



(a) Effect of ABL correction (b) Effect of wind direction
Figure 6. Comparison of predicted gas concentration with measured at various locations.

Conclusions

A CFD approach for gas dispersion modeling in urban-alike environment has been proposed, subject to atmospheric boundary layer conditions. The chemical gas is represented by passive scalar and the impact of the species transport upon the airflow is neglected.

Numerical tests for two different scales of urban-alike test cases show that the simulation results are agreeable well with the measured data, though tracer gases heavier than air were used in measurement. Sensitivity tests about Sc_t yield different optimal values for the two test cases. This implies that care should be taken in the choice of Sc_t value when Reynolds-averaged turbulence models are adopted to study the gas dispersion.

Subsequent research efforts will be put further to address the heavy gas dispersion using multi-species modeling methodology.

References

- [1] *User's Guide for the AMS/EPA Regulatory Model – AERMOD*. EPA-454/B-03-001, Environmental Protection Agency, United States, September 2004.
- [2] *User Guide of OpenFOAM – The Open Source CFD Toolbox*. version v1612+, OpenCFD Limited.
- [3] Patankar S. V., *Numerical Heat Transfer and Fluid Flow*. McGRAW-HILL BOOK COMPANY, ISBN 0-07-048740-5, 1980.
- [4] Gromke C., Buccolieri R., Sabatino S. D. and Ruck B., Dispersion study in a street canyon with tree planting by means of wind tunnel and numerical investigations – Evaluation of CFD data with experimental data. *Atmospheric Environment* 42(37): 8640-8650, 2008.
- [5] Chang J. and Hanna S., Air quality model performance evaluation. *Meteorology and Atmospheric Physics* 87 (2004):167-196.
- [6] Bilitoft C. A., *Customer Report for Mock Urban Setting Test*. DPG Document No. WDTC-FR-01-121, Dec 2001.
- [7] Bekka N., Kumar P., Feiz A., Singh S., Sellam M., Barbosa E., Ngae P., Turbelin G. and Chpoun A., A CFD modeling approach for a contaminant released in a city. *17th Int. Conf. on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*, May 2016, Budapest, Hungary.
- [8] Blocken B., Stathopoulos T. and Carmeliet J., CFD simulation of the atmospheric boundary layer: wall function problems. *Atmospheric Environment* 41(2): 238-252.

Numerical simulations of coupling effects in FGM plates by meshfree methods

***V. Sladek¹, L. Sator¹, M. Repka¹, J. Sladek¹**

¹Institute of Construction and Architecture, Slovak Academy of Sciences, Bratislava, Slovakia

* Corresponding and Presenting author: vladimir.sladek@savba.sk

Abstract

The paper deals with derivation of 2D formulation as well as numerical implementation and study of coupling effects in elastic functionally graded material (FGM) plates within the theory of stationary thermo-elasticity. Unified formulation is developed with involving the assumptions used in the classical Kirchhoff-Love theory for bending of thin elastic plates as well as the assumptions used in the shear deformation plate theory of the 1st and 3rd order. The governing equations and the boundary conditions for deformations are derived from the variational principle, while the formulation for thermal problem is derived by averaging the 3D heat conduction formulation with respect to the transversal direction. The strong formulation and meshless approximation are developed for the derived formulation. The coupling effects are studied by numerical simulations in FGM plates with possible variable thickness and subject to three kinds of stationary loading: (i) uniform transversal loading; (ii) simple tension in plane of the plate; (iii) prescribed different temperatures on the bottom and top surfaces of the plate.

Keywords: Stationary thermo-elasticity, continuous inhomogeneity, plate bending, 2D formulation, coupling effects, numerical study

Introduction

Plate structures are attracting attention of engineers, designer and researchers for a long time because of their superior properties and new features appearing with development of new materials. Due to the small aspect ratio of thickness to in-plane dimensions, in the plate theories the 3D formulation of elasticity problem is assumed in semi-integral form with integration across the plate thickness, and resulting into simplified 2D problems. In stationary thermo-elasticity [1], the temperature field is independent of elastic fields, though it is not valid in reverse. Thus, the thermal problem can be solved separately in advance and one can utilize the obtained temperature field in evaluation of the semi-integral fields occurring in the governing equations for bending problem. For this purpose, it is necessary to replace the Hooke law by the Duhamel-Neumann constitutive law known from the theory of thermo-elasticity [1]. In general, however, we don't know the temperature field in terms of integrable functions and the 2D formulation for bending of plates with including thermal effects cannot be derived in closed form. Therefore the development of 2D formulation for thermal problems in plates is desired. Functional gradation of material coefficients and/or variable thickness of the plate represent another reason why the correct formulation for plate problems must be derived for FGM plates by performing the integration with respect to the transversal coordinate in the variational formulation of the original 3D thermo-elasticity problem. The FGM composites [11-15] have significant utilization in design of structural elements not only because of superior properties of micro-constituents but also for elimination of interface discontinuities occurring in laminated composite structures [16]. The most frequently used modeling of functional variation of material coefficients is the rule of mixture where the material coefficients of multiphase

materials are related directly to the volume fractions and individual coefficients of the constituents. Besides several rather simple models for spatial gradation of volume fractions by analytical functions, there have been developed also more sophisticated models (see e.g. [17-19]) for spatial distribution of volume fractions with including some microstructural aspects of constituents of micro/nano-composites. In this paper, we confine to simple power-law gradations of material coefficients in two-constituent composites, in order to demonstrate some new coupling effects due to gradation of material coefficients and/or plate thickness.

In the most simplified theory, the Kirchhoff-Love theory (KLT), the shear deformations are omitted. There have been developed also generalized shear deformation theories including the 1st order shear deformation theory (FSDPT) [16, 20] and higher order shear deformation theories (HSDPT) [16, 20-23], which account for transverse shear strains and stresses in contrast to the KLT. In this paper, starting from the principle of virtual work and assuming the power-law gradation in the transversal direction, the dependence of all elasticity fields on the transversal coordinate is known a priori and the integrations along the direction of gradation can be accomplished analytically in closed form. Thus, the original 3D problems is converted to 2D problem with correctly derived governing equations and boundary conditions. In stationary thermoelasticity, the temperature field is not influenced by the elasticity fields and the thermal problem can be solved separately. In order to get the 2D formulation also for thermal problem in plates, we consider the power series expansion of the temperature field with respect the transversal coordinate up to the 2nd power, which is physically meaningful as long as the plate thickness is significantly smaller than its characteristic length in the mid-plane of the plate. Having known the dependence of thermal fields on the transversal coordinate, we can consider the 3D heat conduction equation in the averaged sense. This 2D governing equation together with the boundary conditions on the bottom and top surfaces of the plate play the role of governing equations for primary thermal fields. The complete 2D formulation for plate bending in stationary thermo-elasticity is discretized by using the meshless Moving Least Square approximation (MLS) [5-8] for spatial variations of all 2D field variables. Since the governing equations are represented by the partial differential equations (PDE) with variable coefficients and the accuracy of approximation of derivatives is decreasing with increasing their order, we decomposed the original PDE of the 4th order into the system of PDE with 2nd order derivatives by introducing new field variables, in order to eliminate high order derivatives. To facilitate the numerical solution as much as possible for the considered system of the PDE with variable coefficient, we propose to use strong formulation, which is free of any integrations and reduces the amount of evaluations of shape functions, since the evaluations are localized to nodal points. The numerical simulations are employed for study of coupling effects in FGM plates with possible variable thickness and subject to three kinds of stationary loading: (i) uniform transversal loading; (ii) simple tension in plane of the plate; (iii) prescribed different temperatures on the bottom and top surfaces of the plate.

2D formulation of bending problems for FGM plates in stationary thermo-elasticity

It is well known that in stationary thermo-elasticity [1], the temperature field is not affected by mechanical fields, while in linear theory the thermal strains are proportional to the deviation of temperature from its value at the reference state $e_{kl}^{\theta}(\mathbf{x}, z) = \alpha(\theta - \theta_0)\delta_{kl}$, with α being the linear thermal expansion coefficient. Therefore the thermal problem can be solved separately in advance and subsequently the elasticity problem can be solved with bearing in mind thermal strains known from the solution of thermal problem. Evidently, the thermo-elasticity problem is reduced to a pure elasticity problem, if either the temperature is kept on the reference θ_0 value or $\alpha = 0$.

It is well known that the original 3D elasticity problem for plate structures can be reduced to 2D problem because of significantly smaller thickness than the other in-plane length dimensions. According to assumptions adopted for deformation of plates, several theories have been developed for bending of elastic plates. Among the most frequently applied theories to bending of elastic plates, one can name the Kirchhoff-Love theory for bending of thin elastic plates (KLT), and the shear deformation theories of the 1st and 3rd order (FSDPT, TSDPT). Recall that a unified formulation can be developed for bending of elastic plates with possibility to switch between three above mentioned theories by selecting proper values for two key factors [2],[3]. Without going into details, we outline the derivation of the unified formulation for bending of FGM plates within stationary thermo-elasticity. The three components of displacements $v_i(\mathbf{x}, x_3)$ can be expressed in terms of the in-plane displacements $u_\alpha(\mathbf{x})$, transversal displacements (deflections) $w(\mathbf{x})$ and rotations of the normal to the mid-surface $\varphi_\alpha(\mathbf{x})$ as

$$v_i(\mathbf{x}, x_3, t) = \delta_{i\alpha} \left\{ u_\alpha(\mathbf{x}, t) + [c_1 \phi(x_3) - x_3] w_{,\alpha}(\mathbf{x}, t) + c_1 \phi(x_3) \varphi_\alpha(\mathbf{x}, t) \right\} + \delta_{i3} w(\mathbf{x}, t) \quad (1)$$

where $\phi(x_3) := x_3 - c_2 \psi(x_3)$, $\psi(x_3) := \frac{4}{3} \left(\frac{x_3}{h} \right)^2 x_3$, $x_3 \in [-h/2, h/2]$, $\mathbf{x} \in \Omega$

$$c_1 = \begin{cases} 0, & \text{KLT} \\ 1, & \text{SDPT} \end{cases}, \quad c_2 = \begin{cases} 0, & \text{FSDPT, KLT} \\ 1, & \text{TSDPT} \end{cases}.$$

The Latin subscripts can take values $\{1, 2, 3\}$, while the Greek ones $\{1, 2\}$. If we denote the characteristic length in the mid-surface Ω as L , the expression (1) is based on the assumption $h/L \ll 1$. Taking into account the thermal strains and the total strains $e_{ij} = (v_{i,j} + v_{j,i})/2$ together with Hooke's law, one can write the stress tensor components as

$$\sigma_{\alpha\beta}(\mathbf{x}, x_3) = \frac{E}{1-\nu^2} \frac{1-\nu}{H} \left[\tau_{\alpha\beta}^{(u)}(\mathbf{x}) + c_1 \phi(x_3) \tau_{\alpha\beta}^{(\varphi)}(\mathbf{x}) + (c_1 \phi(x_3) - x_3) \tau_{\alpha\beta}^{(w)}(\mathbf{x}) - \alpha \tau_{\alpha\beta}^{(\theta)}(\mathbf{x}, x_3) \right] \quad (2)$$

$$\sigma_{\alpha 3}(\mathbf{x}, x_3) = \frac{E}{1+\nu} \frac{c_1}{2} \phi'(x_3) [w_{,\alpha}(\mathbf{x}) + \varphi_\alpha(\mathbf{x})]$$

$$\sigma_{33}(\mathbf{x}, x_3) = \frac{Ev}{1-\nu^2} \frac{1-\nu}{H} \left[u_{\gamma,\gamma}(\mathbf{x}) + c_1 \phi(x_3) \varphi_{\gamma,\gamma}(\mathbf{x}) + (c_1 \phi(x_3) - x_3) w_{,\gamma\gamma}(\mathbf{x}) - \alpha \frac{1+\nu}{\nu} (\theta(\mathbf{x}, x_3) - \theta_0) \right]$$

where

$$\tau_{\alpha\beta}^{(u)}(\mathbf{x}) := H (u_{\alpha,\beta}(\mathbf{x}) + u_{\beta,\alpha}(\mathbf{x})) / 2 + \nu \delta_{\alpha\beta} u_{\gamma,\gamma}(\mathbf{x}) \quad (3)$$

$$\tau_{\alpha\beta}^{(\varphi)}(\mathbf{x}) := H (\varphi_{\alpha,\beta}(\mathbf{x}) + \varphi_{\beta,\alpha}(\mathbf{x})) / 2 + \nu \delta_{\alpha\beta} \varphi_{\gamma,\gamma}(\mathbf{x})$$

$$\tau_{\alpha\beta}^{(w)}(\mathbf{x}) := H w_{,\alpha\beta}(\mathbf{x}) + \nu \delta_{\alpha\beta} w_{,\gamma\gamma}(\mathbf{x}), \quad \tau_{\alpha\beta}^{(\theta)}(\mathbf{x}, x_3) := (1+\nu) (\theta(\mathbf{x}, x_3) - \theta_0) \delta_{\alpha\beta}$$

are strain contributions corresponding to in-plane displacements, rotations, deflections, and temperature, respectively. Furthermore, E and ν stand for the Young modulus and Poisson ratio, while

$$H = \begin{cases} 1-\nu, & \text{plane stress problems} \\ 1-2\nu, & \text{otherwise} \end{cases}.$$

The Young modulus and linear thermal expansion coefficient are allowed to be continuous functions of position with assuming the power-law gradation in the transversal direction as

$$E(\mathbf{x}, x_3) = E_0 E_H(\mathbf{x}) E_V(x_3), \quad E_V(x_3) = 1 + \zeta \left(\frac{1}{2} \pm \frac{x_3}{h} \right)^p \quad (4)$$

$$\alpha(\mathbf{x}, x_3) = \alpha_0 \alpha_H(\mathbf{x}) \alpha_V(x_3), \quad \alpha_V(x_3) = 1 + \xi \left(\frac{1}{2} \pm \frac{x_3}{h} \right)^r,$$

which result from utilization of rule of mixture for two-constituent micro-composite and the power-law gradation of volume fractions in the transversal direction. Moreover, the thickness of the plate is allowed to be variable on the in-plane coordinates, $h(\mathbf{x})$.

Since the dependence of mechanical fields on x_3 is known a priori, the pure elasticity 3D problem can be converted to 2D problem. In order to extend such a possibility to thermo-elastic problems, we should know also the dependence of temperature on the transversal coordinate prior to solving the 3D thermal boundary value problem. In thin structures ($h/L \ll 1$), it is physically reasonable to simulate the distribution of the temperature field by using the power series expansion

$$\theta(\mathbf{x}, x_3) \approx \theta_0 + \vartheta_0(\mathbf{x}) + z\vartheta_1(\mathbf{x}) + z^2\vartheta_2(\mathbf{x}), \quad z = \frac{x_3}{h} \in [-0.5, 0.5] \quad (5)$$

in which the new fields $\vartheta_s(\mathbf{x})$ for ($s=0,1,2$), are variable in the plate mid-plane. In view of (3) and (5), we obtain

$$\tau_{\alpha\beta}^{(\theta)}(\mathbf{x}, x_3) = \sum_{s=0}^2 z^s \tau_{\alpha\beta}^{(\vartheta_s)}(\mathbf{x}), \quad \tau_{\alpha\beta}^{(\vartheta_s)}(\mathbf{x}) := (1+\nu)\vartheta_s(\mathbf{x})\delta_{\alpha\beta}. \quad (6)$$

Now, the variational formulation of the mechanical part of the original 3D thermoelasticity problem is given by the principle of virtual work

$$\begin{aligned} \delta U - \delta W_e &= 0, \quad \delta U = \int_{\Omega} \left(\int_{-h/2}^{h/2} \sigma_{ij}(\mathbf{x}, x_3) \delta e_{ij}(\mathbf{x}, x_3) dx_3 \right) d\Omega \\ \delta W_e &= \int_{\Omega} \bar{t}_3(\mathbf{x}) \delta w(\mathbf{x}) d\Omega + h(\mathbf{x}) \int_{\partial\Omega} \bar{t}_{\alpha}(\mathbf{x}) \delta u_{\alpha}(\mathbf{x}) d\Gamma \end{aligned} \quad (7)$$

in which the work of external forces is represented by the transversal loading $\bar{t}_3(\mathbf{x})$ applied on the top/bottom surfaces, and $\bar{t}_{\alpha}(\mathbf{x})$ are in-plane tractions applied on the lateral surfaces of the plate. The integration with respect to x_3 can be performed analytically and we obtain the 2D formulation given by governing equations at $\mathbf{x} \in \Omega$:

$$T_{\alpha\beta,\beta}(\mathbf{x}) = 0 \quad (8)$$

$$M_{\alpha\beta,\alpha\beta}^{(w)}(\mathbf{x}) + T_{3\beta,\beta}^{(w\varphi)}(\mathbf{x}) = -t_3(\mathbf{x})$$

$$M_{\alpha\beta,\beta}^{(\varphi)}(\mathbf{x}) - T_{3\alpha}^{(w\varphi)}(\mathbf{x}) = 0$$

and boundary restrictions (possible boundary conditions) at $\mathbf{x} \in \partial\Omega$:

$$\left[n_{\beta}(\mathbf{x}) T_{\alpha\beta}(\mathbf{x}) - h \bar{t}_{\alpha}(\mathbf{x}) \right] \delta u_{\alpha}(\mathbf{x}) = 0 \Rightarrow n_{\beta}(\mathbf{x}) T_{\alpha\beta}(\mathbf{x}) - h \bar{t}_{\alpha}(\mathbf{x}) = 0 \quad \text{or} \quad u_{\alpha}(\mathbf{x}) = \bar{u}_{\alpha}(\mathbf{x}) \quad (9a)$$

$$n_{\alpha}(\mathbf{x}) n_{\beta}(\mathbf{x}) M_{\alpha\beta}^{(w)}(\mathbf{x}) \delta \left(\frac{\partial w}{\partial \mathbf{n}}(\mathbf{x}) \right) = 0 \Rightarrow n_{\alpha}(\mathbf{x}) n_{\beta}(\mathbf{x}) M_{\alpha\beta}^{(w)}(\mathbf{x}) = 0 \quad \text{or} \quad \frac{\partial w}{\partial \mathbf{n}}(\mathbf{x}) = 0 \quad (9b)$$

$$n_{\beta}(\mathbf{x}) M_{\alpha\beta}^{(\varphi)}(\mathbf{x}) \delta \varphi_{\alpha}(\mathbf{x}) = 0 \Rightarrow n_{\beta}(\mathbf{x}) M_{\alpha\beta}^{(\varphi)}(\mathbf{x}) = 0 \quad \text{or} \quad \varphi_{\alpha}(\mathbf{x}) = 0 \quad (9c)$$

$$V(\mathbf{x}) \delta w(\mathbf{x}) = 0 \Rightarrow V(\mathbf{x}) = 0 \quad \text{or} \quad w(\mathbf{x}) = 0 \quad (9d)$$

where

$$V(\mathbf{x}) := n_\alpha(\mathbf{x}) \left(M_{\alpha\beta,\beta}^{(w)}(\mathbf{x}) + T_{3\alpha}^{(w\varphi)}(\mathbf{x}) \right) + \frac{\partial}{\partial t} T^{(w)}(\mathbf{x}) - \sum_c \delta(\mathbf{x} - \mathbf{x}^c) \left\| T^{(w)}(\mathbf{x}^c) \right\|$$

is the generalized shear force and the following semi-integral fields have been introduced by the definitions

$$\begin{aligned} T_{\alpha\beta}(\mathbf{x}) &:= \int_{-h/2}^{h/2} \sigma_{\alpha\beta}(\mathbf{x}, x_3) dx_3, \\ M_{\alpha\beta}^{(\varphi)}(\mathbf{x}) &:= c_1 \int_{-h/2}^{h/2} \phi(x_3) \sigma_{\alpha\beta}(\mathbf{x}, x_3) dx_3, \quad M_{\alpha\beta}^{(w)}(\mathbf{x}) := \int_{-h/2}^{h/2} (x_3 - c_1 \phi(x_3)) \sigma_{\alpha\beta}(\mathbf{x}, x_3) dx_3, \\ T_{3\beta}^{(w\varphi)}(\mathbf{x}) &:= c_1 \int_{-h/2}^{h/2} \left\{ \left[(1 - c_2) \kappa + c_2 \right] - c_2 \psi'(x_3) \right\} \sigma_{3\beta}(\mathbf{x}, x_3) dx_3 \end{aligned} \quad (10)$$

where the shear correction factor κ has been introduced as the Reissner modification of the shear stresses in order to be predicted a correct amount of energy in the case of the FSDPT ($c_1 = 1 \wedge c_2 = 0$). Furthermore, the twisting moment has been introduced as

$$T^{(w)}(\mathbf{x}) := t_\alpha(\mathbf{x}) n_\beta(\mathbf{x}) M_{\alpha\beta}^{(w)}(\mathbf{x}) \quad (11)$$

and the jump at a corner point on the oriented boundary edge is defined as

$$\left\| A(\mathbf{x}^c) \right\| := A(\mathbf{x}^c + 0) - A(\mathbf{x}^c - 0).$$

The explicit expressions for semi-integral fields are given in Appendix, since the integrations prescribed in (10) can be performed in closed form. Substituting (A.2) into (8) and (9), one obtains the governing equations and the possible boundary conditions in terms of primary fields and their derivatives.

Up to now, we have supposed that the temperature is known from the solution of stationary thermal problem. Now, we need to derive the governing equations and boundary conditions for particular fields $\mathcal{G}_s(\mathbf{x})$ defined in Eq. (5) with starting from the 3D formulation, where the heat conduction equation is given by the PDE

$$\left(k(\mathbf{x}, x_3) \theta_{,j}(\mathbf{x}, x_3) \right)_{,j} = 0 \quad (12)$$

in which k is the heat conduction coefficient, which is prescribed by continuous functions in FGM. In accordance with above mentioned assumptions, we consider

$$k(\mathbf{x}, x_3) = k_0 k_H(\mathbf{x}) k_V(x_3), \quad k_V(x_3) = 1 + \omega \left(\frac{1}{2} \pm \frac{x_3}{h} \right)^s. \quad (13)$$

Substituting (5) into (12), we obtain the PDE, which is still dependent on the transversal coordinate x_3 ,

$$\left[k_H(\mathbf{x}) k_V(x_3) \left(\sum_{s=0}^2 \left(\frac{x_3}{h} \right)^s \mathcal{G}_s(\mathbf{x}, t) \right)_{,j} \right]_{,j} = 0. \quad (14)$$

In order to get the pure 2D formulation, we can consider Eq. (14) in averaged sense over the plate thickness, which is physically meaningful as long as $L/h \gg 1$. Performing the integration of Eq. (14) over the plate thickness, we obtain the averaged heat conduction equation

$$\sum_{s=1}^2 C^{(\theta g_s)}(\mathbf{x}) \mathcal{G}_s^*(\mathbf{x}, t) + \sum_{s=0}^2 G_{\beta}^{(\theta g_s)}(\mathbf{x}) \mathcal{G}_{s,\beta}^*(\mathbf{x}, t) + \sum_{s=0}^2 G^{(\theta g_s)}(\mathbf{x}) \mathcal{G}_{s,\beta\beta}^*(\mathbf{x}, t) = 0, \quad (15)$$

with

$$C^{(\theta g_a)}(\mathbf{x}) := \left(\frac{L}{h_0} \right)^2 \frac{k_H(\mathbf{x})}{\left(h^*(\mathbf{x}) \right)^2} \left[k_V(1/2) + (-1)^a k_V(-1/2) \right] \quad (16)$$

$$G_{\beta}^{(\theta g_a)}(\mathbf{x}) := L \left(d_{(0)a} + \omega d_{(s)a} \right) k_{H,\beta}(\mathbf{x}), \quad G^{(\theta g_a)}(\mathbf{x}) := L^2 \left(d_{(0)a} + \omega d_{(s)a} \right) k_H(\mathbf{x}).$$

Obviously, Eq. (15) is the PDE in 2D domain Ω and two additional equation result from the thermal boundary conditions on the bottom and top surfaces of the plate. Usually, we distinguish three kinds of thermal boundary conditions, which result into the following additional equations

$$(i) \text{ Dirichlet type: } \mathcal{G}_0(\mathbf{x}) \pm \frac{1}{2} \mathcal{G}_1(\mathbf{x}) + \frac{1}{4} \mathcal{G}_2(\mathbf{x}) = \bar{\theta}(\mathbf{x}, \pm h/2) - \theta_0 \quad (17a)$$

$$(ii) \text{ Neumann type: } \mp k_0 k_H(\mathbf{x}) k_V(\pm h/2) [\mathcal{G}_1(\mathbf{x}) \pm \mathcal{G}_2(\mathbf{x})] = h \bar{q}(\mathbf{x}, \pm h/2) \quad (17b)$$

$$(iii) \text{ Robin type: } A \left[1 + \mathcal{G}_0(\mathbf{x}) \pm \frac{1}{2} \mathcal{G}_1(\mathbf{x}) + \frac{1}{4} \mathcal{G}_2(\mathbf{x}) \right] \pm B h k_0 k_H(\mathbf{x}) k_V(\pm h/2) [\mathcal{G}_1(\mathbf{x}) \pm \mathcal{G}_2(\mathbf{x})] = 0 \quad (17c)$$

in which $\bar{\theta}(\mathbf{x}, \pm h/2)$ and $\bar{q}(\mathbf{x}, \pm h/2)$ stand for the prescribed temperature and heat flux on the top and bottom surfaces of the plate.

Finally, the PDE (15) is to be supplemented with the boundary conditions on the boundary edge of the plate. Since $\theta(\mathbf{x}, x_3 = 0) = \theta_0 + \mathcal{G}_0(\mathbf{x})$, the boundary conditions on $\partial\Omega$ can be given as

$$(i) \text{ Dirichlet type: } \mathcal{G}_0(\mathbf{x})|_{\partial\Omega} = \bar{\theta}(\mathbf{x}, x_3 = 0) - \theta_0 \quad (18a)$$

$$(ii) \text{ Neumann type: } -k_0 k_H(\mathbf{x}) k_V(0) n_{\beta}(\mathbf{x}) \mathcal{G}_{0,\beta}(\mathbf{x})|_{\partial\Omega} = \bar{q}(\mathbf{x}, 0) \quad (18b)$$

$$(iii) \text{ Robin type: } A \mathcal{G}_0(\mathbf{x})|_{\partial\Omega} + B k_0 k_H(\mathbf{x}) k_V(0) n_{\beta}(\mathbf{x}) \mathcal{G}_{0,\beta}(\mathbf{x})|_{\partial\Omega} = 0. \quad (18c)$$

in which $\bar{\theta}(\mathbf{x}, x_3 = 0)$ and $\bar{q}(\mathbf{x}, 0)$ are the prescribed values of the temperature and heat flux, respectively, on the boundary edge of the plate.

Note that the governing equations involve the 4th order derivatives of deflections and 3rd order derivatives of in-plane displacements and rotations. Since the accuracy of approximations of derivatives is decreasing with increasing the order of derivatives, we propose the decomposition of the derived system of the PDE into a set of PDE with derivatives not higher than the 2nd order by introducing new field variables [4] as

$$m(\mathbf{x}) := \nabla^2 w(\mathbf{x}), \quad s_{\alpha}(\mathbf{x}) := \nabla^2 u_{\alpha}(\mathbf{x}), \quad f_{\alpha}(\mathbf{x}) := \nabla^2 \varphi_{\alpha}(\mathbf{x}). \quad (19)$$

Summarizing, the governing equations for thermal problem are given by Eqs. (15) and (17) at $\mathbf{x} \in \Omega$ and the possible boundary conditions on $\partial\Omega$ are given by Eq. (18). The governing

equations for the mechanical part of the thermo-elastic problem are given by Eqs. (8) and (19) at $\mathbf{x} \in \Omega$, while the possible boundary conditions can be properly constructed from Eq. (9).

Numerical implementation

Although the proposed decomposition of the original system of high-order PDE of elliptic type into the system of 2nd order PDE increases the number of field variables and finally the size of the matrix of discretized equations, it brings the possibility to solve the system of the decomposed PDE using the strong formulation which accelerates the computation significantly as compared with the weak formulation especially in case of utilization of meshless approximations, because the evaluations of shape functions is localized to nodal points. The functional in-plane gradation of material coefficients leads to the PDE with variable coefficients and the classical element-based discretization methods are mostly disqualified for efficient treatment of such rather complex problems. In order to simplified the mathematical complexity as much as possibly with preserving the physical nature of the solved problems, we propose to utilize the strong formulation and the meshless approximation of spatial variations of field variables, which is in this paper, the Moving Least Square approximation [5]. The nodal points are freely distributed in the analyzed domain and on its boundary without creating any connectivity among the nodes. Without going into details [6, 7], the approximation of a field variable $g(\mathbf{x})$ and its derivatives around the central approximation node \mathbf{x}^q can be expressed by

$$g(\mathbf{x}) \approx \sum_{a=1}^{N^q} \hat{g}^{\bar{a}} \phi^{(q,a)}(\mathbf{x}), \quad g_{,\alpha}(\mathbf{x}) \approx \sum_{a=1}^{N^q} \hat{g}^{\bar{a}} \phi_{,\alpha}^{(q,a)}(\mathbf{x}), \quad g_{,\alpha\beta}(\mathbf{x}) \approx \sum_{a=1}^{N^q} \hat{g}^{\bar{a}} \phi_{,\alpha\beta}^{(q,a)}(\mathbf{x}), \quad (20)$$

where $\bar{a} = n(q,a)$ is the global number of the a -th node from the influence domain of \mathbf{x}^q , N^q is the number of nodal points in the influence domain, which is smaller than the total number of nodes, and $\phi^{(q,a)}(\mathbf{x})$ is the shape function associated with the node $n(q,a)$. This shape function is not known in a closed form [8], but it must be evaluated at each point \mathbf{x} . Recall that $\hat{g}^{\bar{a}}$ is the nodal unknown associated with the node \bar{a} and is different from the nodal value $g(\mathbf{x}^{\bar{a}})$. The central approximation node can be selected as the nearest node to the field point \mathbf{x} . For creation of shape functions, we have used cubic monomial basis and Gaussian weights.

In the strong formulation, the governing equations are collocated at interior nodes and the boundary conditions at boundary nodes.

Numerical examples

In numerical investigations, we consider a square plate $L \times L$ with $L=1$ and the results are presented for dimensionless quantities specified as:

$$x_\beta^* := \frac{x_\beta}{L}, \quad x_3^* := \frac{x_3}{h_0} = h^*(\mathbf{x})z, \quad u_\beta^*(\mathbf{x}^*) := \frac{u_\beta(\mathbf{x})}{h_0}, \quad \varphi_\beta^*(\mathbf{x}^*) := \varphi_\beta(\mathbf{x}), \quad w^*(\mathbf{x}^*) := \frac{w(\mathbf{x})}{h_0},$$

$$\vartheta_s^*(\mathbf{x}^*) := \frac{\vartheta_s(\mathbf{x})}{\theta_0}, \quad t_3^*(\mathbf{x}^*) := \frac{L^4}{D_0 h_0} t_3(\mathbf{x}), \quad \text{with } D_0 = \frac{E_0 (h_0)^3}{12(1-\nu^2)}.$$

In what follows, we shall omit the superscript $*$ in dimensionless Cartesian coordinates.

In all numerical computations, we have used a uniform distribution of nodal points (36 x36 nodes) with δ being the distance between two neighbour nodes. The other parameters in the MLS-approximation with Gaussian weights have been taken as: the radius of the influence domain $\rho^a = 3.001\delta$, shape function parameter $c^a = \delta$, and cubic monomial basis. Note that

the number of nodes falling into the influence domain of a global node varies from 11 to 27 depending on the position of the global node.

Elasto-static simulations

We start the study of coupling effects due to continuously variable: (i) Young's modulus in transversal direction $E_V(x_3) = 1 + \zeta(1/2 \pm x_3/h)^p$; (ii) in-plane gradations of Young's modulus $E_H(\mathbf{x}) = 1 + \kappa_0(x_1/L)^{b_0}$; (iii) variable plate thickness $h^*(\mathbf{x}) = 1 + \kappa(x_1/L)^s$. The Poisson ratio is assumed to be constant $\nu = 0.3$. The boundary edges of the plate are clamped and the plate is subjected to uniform transversal loading $t_3^*(\mathbf{x}) = 1$.

The influence of thickness of the plate on deflections is shown in Fig.1 for homogeneous plate. It is seen that the KLT is applicable only to thin plates $L/h_0 > 50$, when the deviation of the KLT results from those by the SDPT is less than 1%.

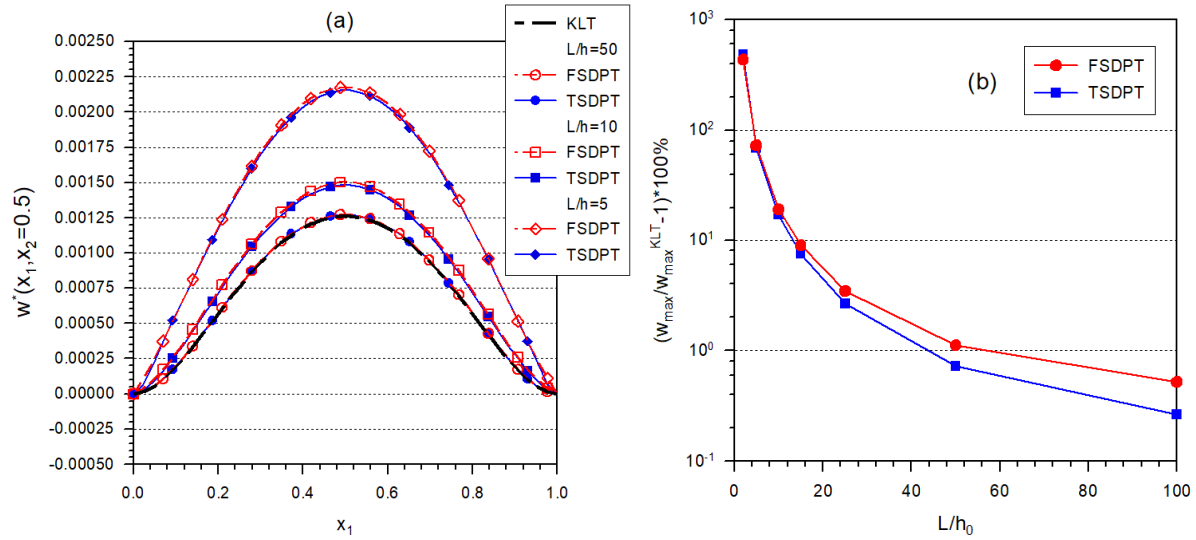


Figure 1. Comparison of deflections by KLT and SDPT: (a) deflections along x_1 -axis; (b) % deviations vs. ratio of length to thickness

Fig. 2 shows the response of three thin plates (homogeneous and FGM plates with 2 different levels of gradation of Young's modulus, $\zeta = 1$ and $\zeta = 3$) to uniform transversal loading. The evidence of coupling between the bending and in-plane deformation modes is clearly seen from Fig. 2(c), since finite in-plane displacements arise only in FGM plates.

The influence of the in-plane power-law gradation of Young's modulus on the response of thick plate is shown in Fig. 3. It is seen that deflections are affected more expressively by the level of gradation than by the exponent of gradation. This can be explained by the effect of lower bulk content of the material with higher value of the Young modulus. The shift of maximal deflections toward the softer side of the plate is evident.

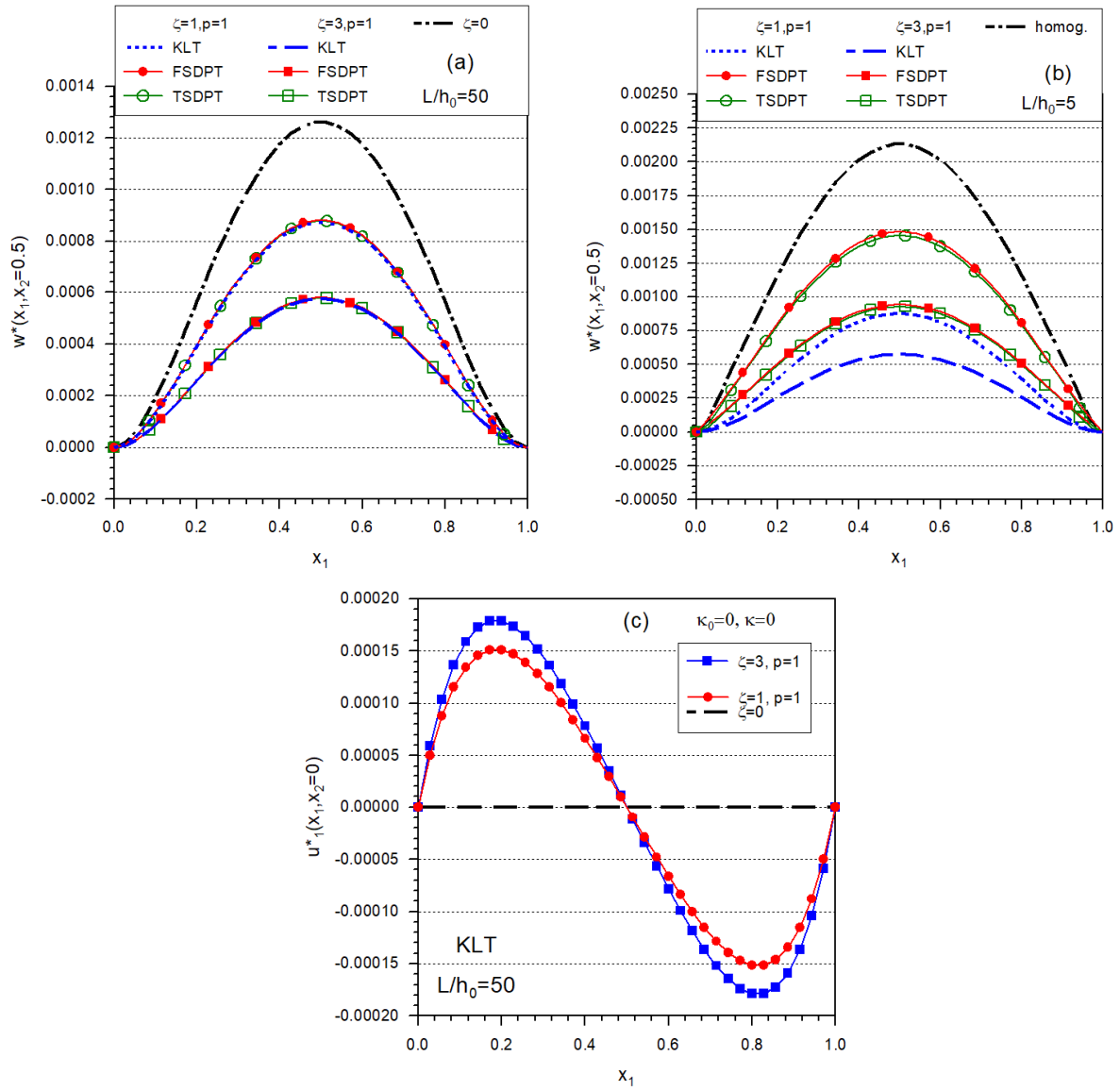


Figure 2. The response of three plates by various theories. In-plane distribution: (a) deflections of thin plate; (b) deflections of thick plate; (c) in-plane displacements

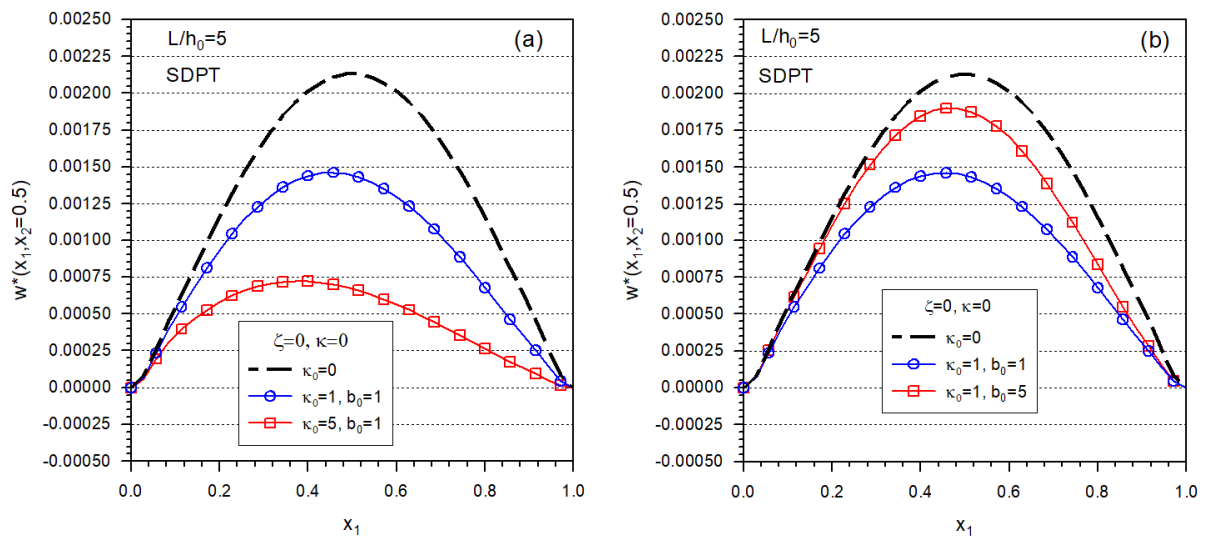


Figure 3. Influence of in-plane gradation of Young's modulus on the response of thick FGM plate. The effects by: (a) different levels of gradation; (b) different exponents of power-law gradation

Form Fig. 4(a), it can be seen that the influence of levels of linear in-plane gradations of plate thickness on deflections is more significant than the influence of gradation of Young's modulus. The negative effect of thinning the plate can be compensated or suppressed by parallel gradual increasing the Young modulus with resulting in overall reduction of deflection (Fig. 4(b)).

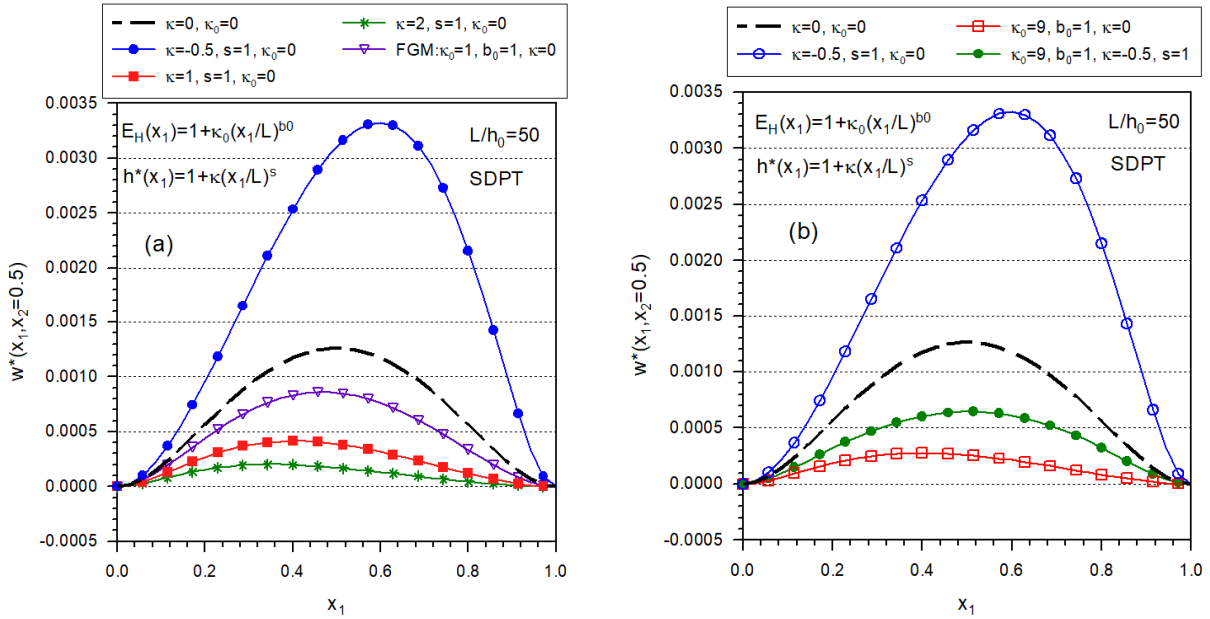


Figure 4. Response of plates with variable parameters of in-plane gradations: (a) κ and/or κ_0 ; (b) κ and/or κ_0 including combined gradations of Young's modulus and plate thickness

Now, we shall continue in numerical simulations and study of multi-gradation effects in square FGM elastic plates with transversal gradation of Young's modulus $E_V(x_3) = 1 + \zeta(1/2 \pm x_3/h)^p$ combined with: (i) in-plane gradation of Young's modulus $E_H(\mathbf{x}) = 1 + \kappa_0(x_1/L)^{b_0}$; (ii) in-plane continuous variation of plate thickness $h^*(\mathbf{x}) = 1 + \kappa(x_1/L)^s$; (iii) simultaneous in-plane gradations of Young's modulus and plate thickness. The plate is subject to uniform in-plane tension $T_{\alpha 1}^*(x_1=1, x_2) = \delta_{\alpha 1} h_0 h^*(x_1=1, x_2)$, and the boundary conditions on the other edges are given as: $u_1^*(x_1=0, x_2) = 0$, $T_{21}^*(x_1=0, x_2) = 0$, $T_{\alpha 2}^*(x_1, x_2=0) = T_{\alpha 2}^*(x_1, x_2=1) = 0$, with assuming two alternatives for bending modes:

- clamped boundary edges

$$w^*(\mathbf{x})|_{\partial\Omega} = 0, \quad \frac{\partial w^*(\mathbf{x})}{\partial \mathbf{n}}|_{\partial\Omega} = 0, \quad \varphi_\alpha^*(\mathbf{x})|_{\partial\Omega} = 0, \quad t_3^*(\mathbf{x}) = 0$$

- simply supported edges

$$w^*(\mathbf{x})|_{\partial\Omega} = 0, \quad n_\alpha(\mathbf{x}) n_\beta(\mathbf{x}) M_{\alpha\beta}^{(w)*}(\mathbf{x})|_{\partial\Omega} = 0, \quad n_\beta(\mathbf{x}) M_{\alpha\beta}^{(\varphi)*}(\mathbf{x})|_{\partial\Omega} = 0, \quad t_3^*(\mathbf{x}) = 0.$$

The in-plane loading doesn't yield finite deflections in homogeneous as well as FGM with only transversal gradation of Young's modulus.

Firstly consider the FGM plates with clamped sliding edges. Fig. 5 show the in-plane displacements and deflections in thin FGM plates with combined gradations of Young's modulus for various levels and/or exponents of in-plane gradation. The combined gradation of Young's modulus is sufficient for arising finite deflections in FGM plates subject to in-plane tension.

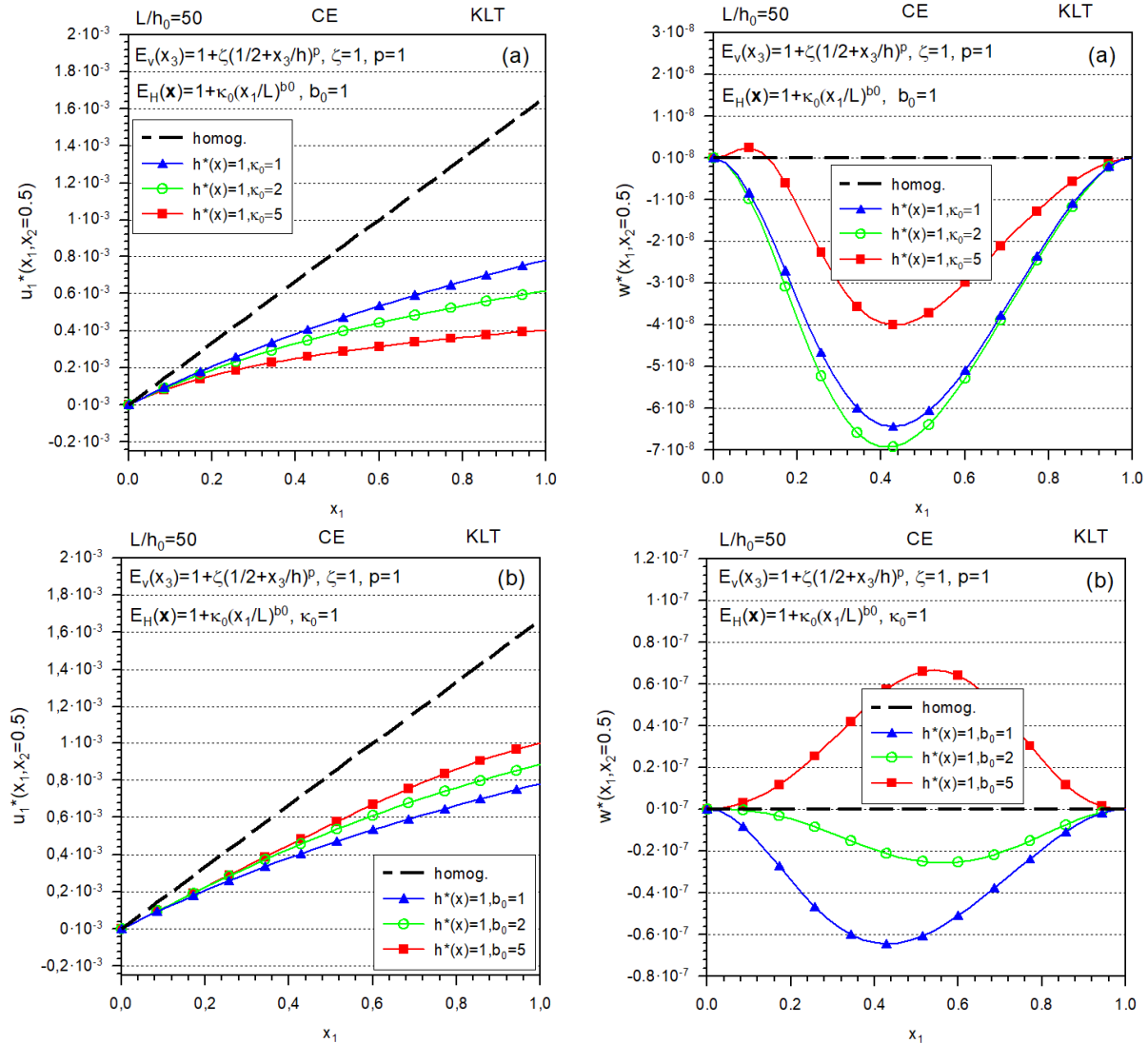


Figure 5. Influence of gradation parameters: (a) levels of gradations; (b) exponents of gradations on in-plane displacements and deflections in thin FGM plates with combined transversal and in-plane gradations of Young's modulus

Numerical simulations indicate that similar behavior is observed also in FGM plates with transversal gradation of Young's modulus and continuously variable thickness (Fig. 6). Note that the nonlinear gradation of the plate thickness leads to more significant deflection response than in the case of linear gradation.

Note that in the case of thin plates, the KLT results are almost identical with those by SDPT. However, in the case of thick plates, it is necessary to use the TSDPT. More remarkable deflection response to in-plane loading is observed in the case of thick plates if $s = 1$ (Fig. 7).

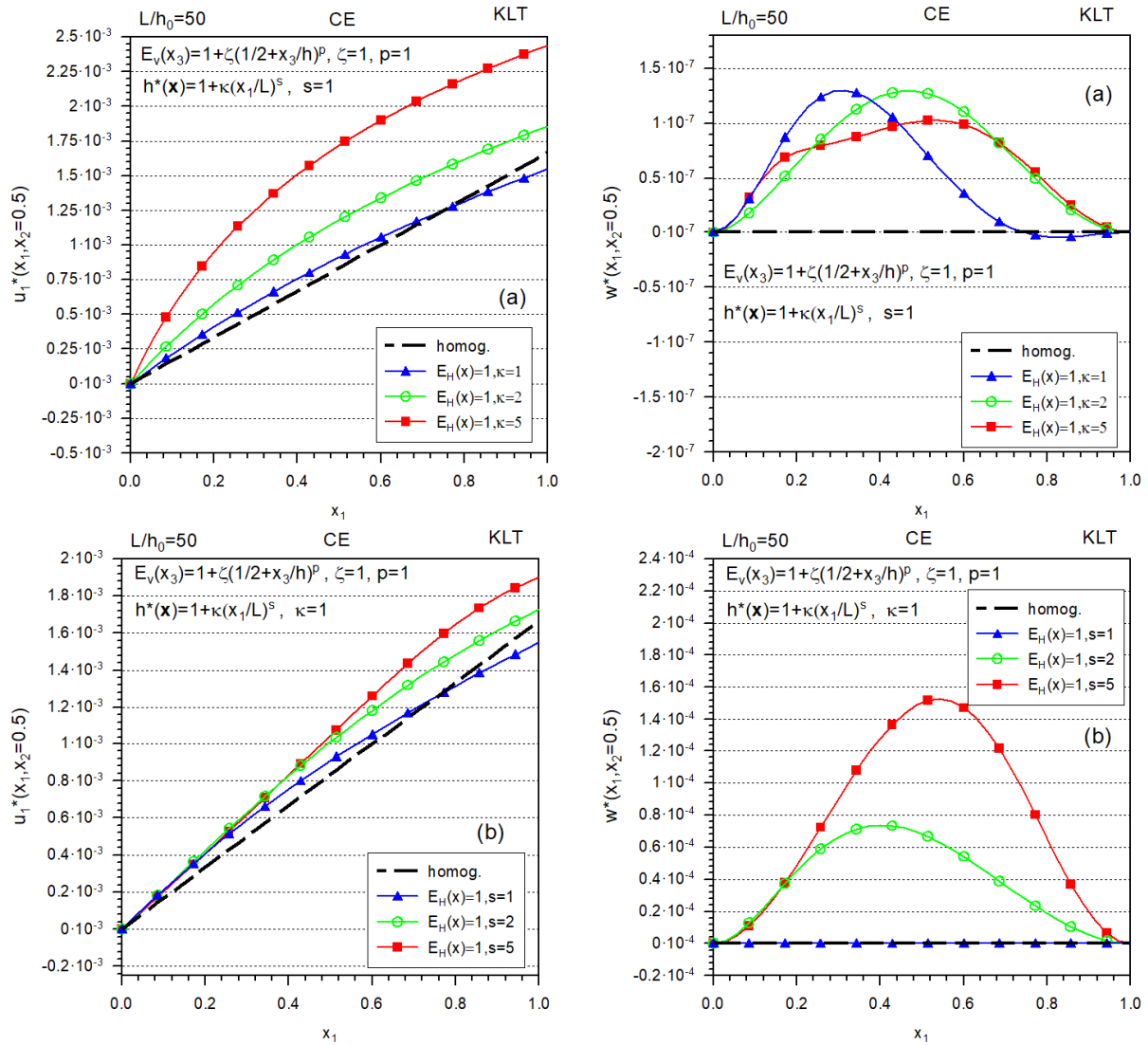


Figure 6. Influence of gradation parameters: (a) levels of gradations; (b) exponents of gradations on in-plane displacements and deflections in thin FGM plates with combined transversal gradation of Young's modulus and in-plane variation of plate thickness

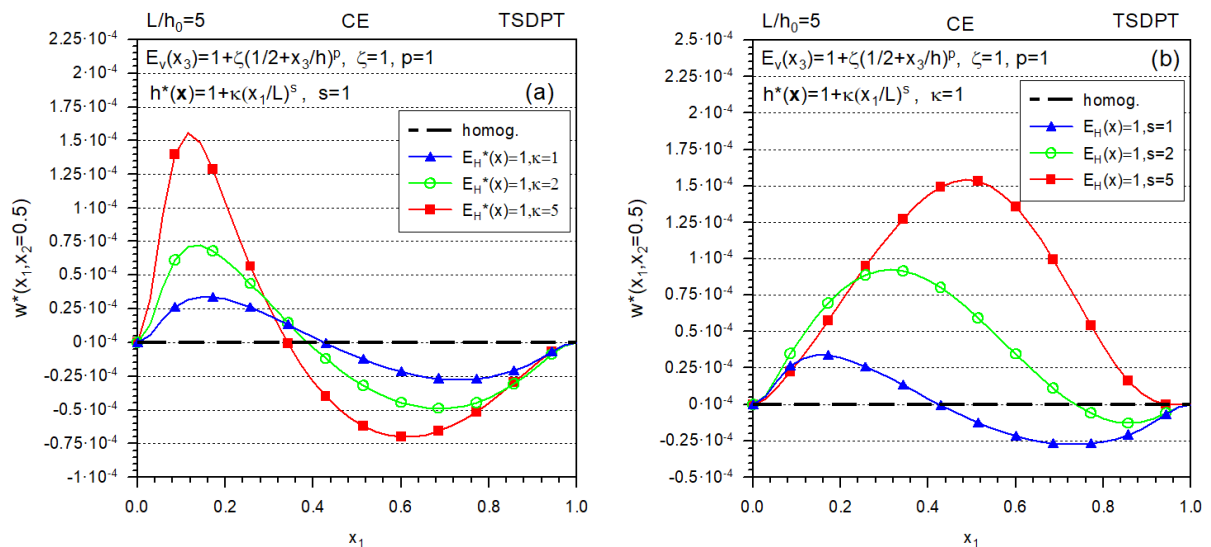


Figure 7. Influence of gradation parameters: (a) levels of gradations; (b) exponents of gradations on deflections of thick FGM plates with combined transversal gradation of Young's modulus and in-plane variation of plate thickness

In the rest of the elastostatical subsection, we shall consider FGM plates with simply supported sliding edges. Now, the transversal gradation of Young's modulus is sufficient for finite deflection response, in contrast to the plates with clamped sliding edges (Fig.8). For more details, we refer the reader to the work [9].

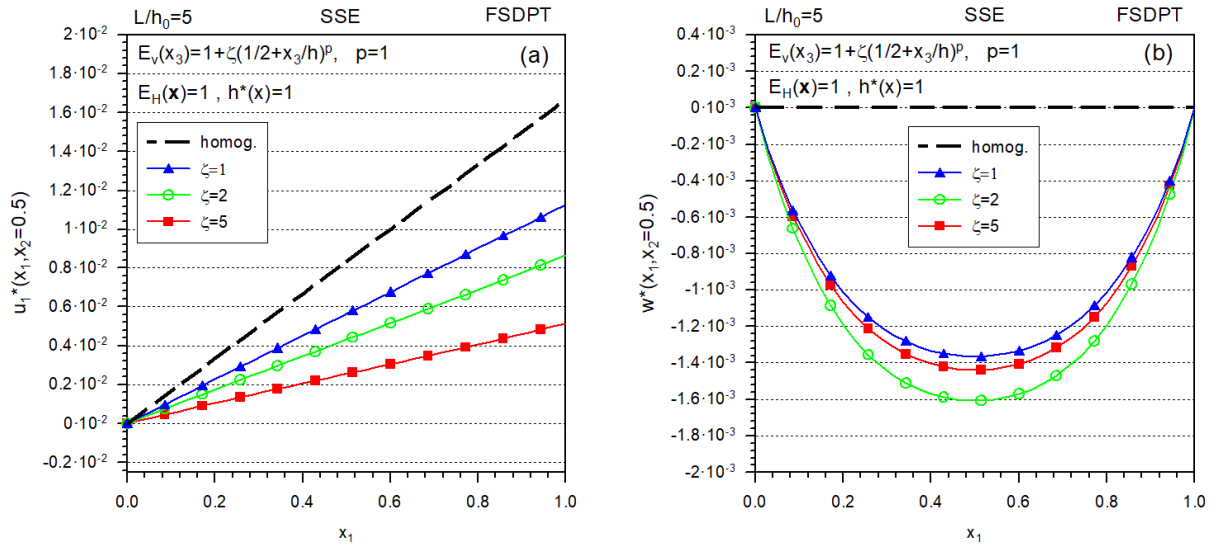


Figure 8. Influence of the level of linear transversal gradation of Young's modulus on the response of the FGM plates with SSE to in-plane tension: (a) in-plane displacements; (b) deflections

The numerical simulations in the FGM plates with combined gradations and simply supported edges (SSE) resembles qualitatively those in the FGM plates with multi-gradation and clamped edges (CE). However, the deflection response in FGM plates with SSE is much more expressive than that in the FGM plates with multi-gradation and CE. Finally, the results for the FGM plates with variable thickness and combined transversal and in-plane gradations of Young's modulus are illustrated in Fig.9.

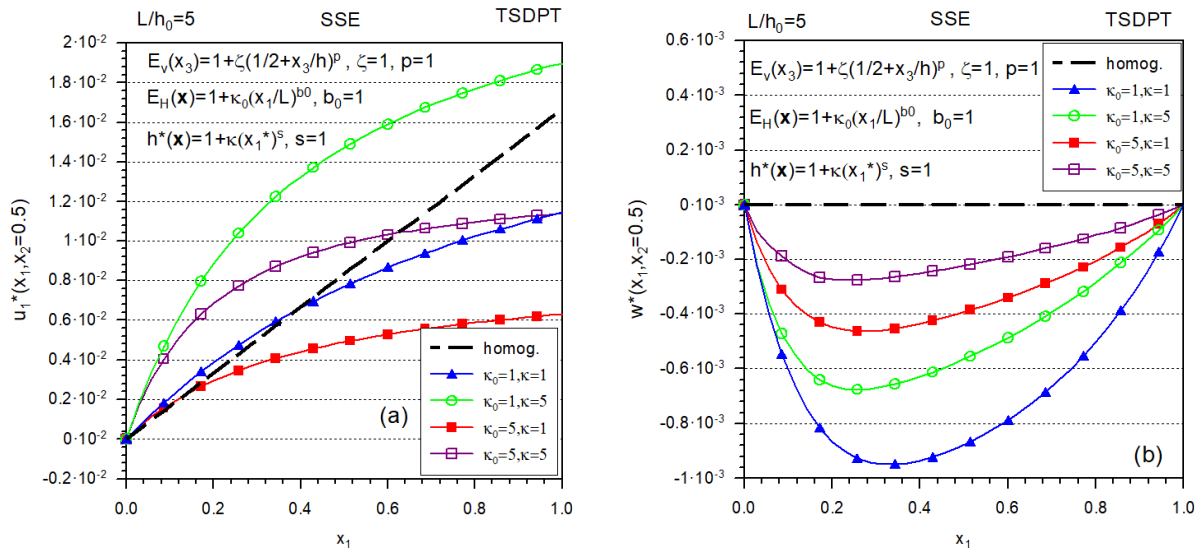


Figure 9. Influence of the multi-gradations of Young's modulus and plate thickness on the response of the FGM plates with SSE to in-plane tension: (a) in-plane displacements; (b) deflections

Thermo-elastic simulations

In addition to functional gradation of elasticity coefficients, we assume the transversal gradation of the linear thermal expansion and the heat conduction coefficients specified by Eqs. (4) and (13), and the following in-plane power-law gradations

$$\alpha_H(\mathbf{x}) = 1 + \kappa_1 (x_1 / L)^{b_1}, \quad k_H(\mathbf{x}) = 1 + \kappa_2 (x_1 / L)^{b_2}.$$

The natural thermal boundary conditions are assumed on the bottom and top surfaces of plate

$$\theta(\mathbf{x}, \pm h/2) = \theta^\pm \quad \text{with} \quad \theta_0 = 1, \quad \theta^+ = \theta_0 + 20, \quad \theta^- = \theta_0$$

and heat flux is assumed to be vanishing on the boundary edges of the plate

$$-k_0 k_H(\mathbf{x}) k_V(0) n_\beta(\mathbf{x}) \mathcal{G}_{0,\beta}(\mathbf{x}) \Big|_{\partial\Omega} = \bar{q}(\mathbf{x}, 0) = 0.$$

Then, the temperature field is distributed uniformly within the mid-plane Ω and the value of the temperature is affected only by the level of gradation of the heat conduction coefficient, ω . The numerical simulations presented in Fig. 10 show that the KLT and TSDPT give different response of the FGM plates with clamped edges (CE) to considered thermal loading even if the plates are thin. It can be seen from the analysis of the governing equations [10] that there is a coupling between deflections and thermal fields in the KLT only if $\nabla^2 (E_H(\mathbf{x}) \alpha_H(\mathbf{x})) \neq 0$ while in the case of the SDPT such a coupling appears even if $\nabla (E_H(\mathbf{x}) \alpha_H(\mathbf{x})) \neq 0$.

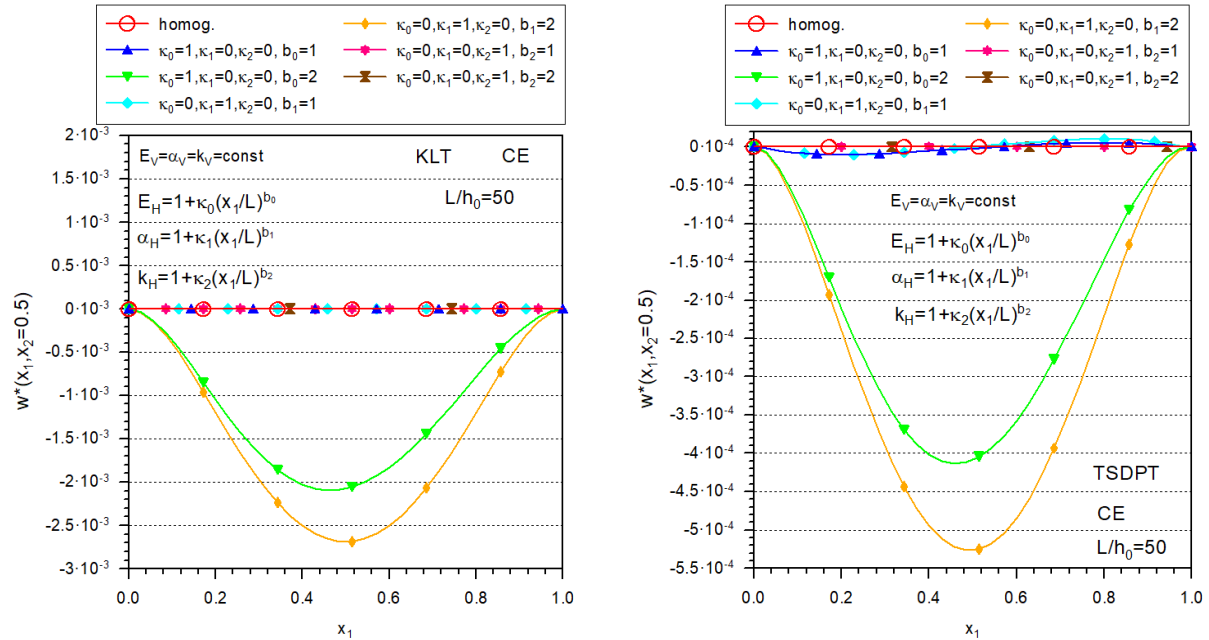


Figure 10. Deflection response to thermal loading in FGM plates with in-plane gradation of material coefficients and clamped edges

The plates with simply supported edges (SSE) exhibit quite different behavior as plates with clamped edges. One observes finite deflections also in homogeneous plates with SSE, because

the thermal contribution to the bending moment, $n_\alpha n_\beta M_{\alpha\beta}^{*(w\theta)} = n_\alpha n_\beta \sum_{a=0}^2 C^{(w\theta)} \tau_{\alpha\beta}^{*(\theta_a)}$, on the boundary edge is compensated by contribution associated with deflection strains, i.e. $n_\alpha n_\beta M_{\alpha\beta}^{*(w\theta)}$ is the boundary source for $n_\alpha n_\beta M_{\alpha\beta}^{*(ww)} = n_\alpha n_\beta \sum_{a=0}^2 C^{(ww)} \tau_{\alpha\beta}^{*(w)}$. On the other hand, such thermal source is ineffective on the clamped edge.

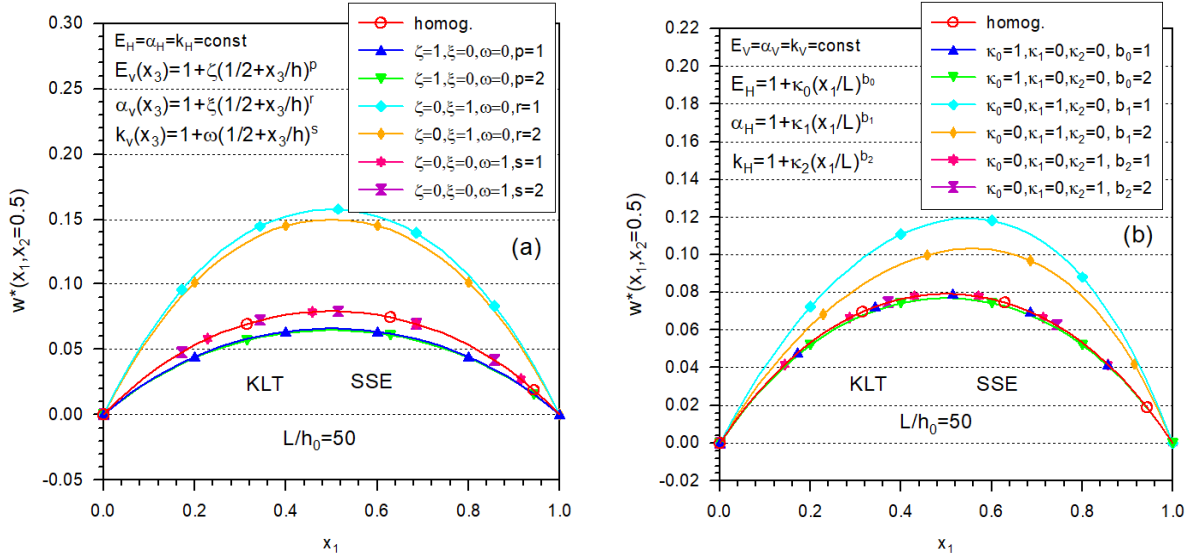


Figure 11. Deflection response to thermal loading in FGM plates with simply supported edges and various gradations of material coefficients: (a) transversal gradation; (b) in-plane gradation

The influence of gradation of material coefficients on the deflection responses in FGM plates with simply supported edges are shown in Fig. 11. It is seen that deflection response is much more expressive than in the case of plates with clamped edges.

Conclusions

The unified formulation for FGM plates is developed within stationary thermo-elasticity with including the assumptions of the Kirchhoff-Love theory as well as the 1st and 3rd order shear deformation plate theories. The functional gradation is considered in the transversal and/or in-plane direction for such material coefficients as: the Young modulus, coefficient of linear thermal expansion, and heat conduction coefficient. Moreover, the plate thickness can be continuously variable too. For the derived 2D formulation, the numerical implementation is developed with making use the strong formulation and meshless approximation of spatial variations of field variables. The original system of the governing PDE is decomposed into the system of the 2nd order PDE, in order to decrease the order of derivatives in the original system. The numerical simulations are employed for study the coupling effects in FGM plates subjected to three kinds of stationary loading: (i) uniform transversal mechanical loading; (ii) simple tension applied in the plane of plate; (iii) thermal loading. Individual as well as combined gradations of material coefficients and the plate thickness are considered. The coupling between the bending and in-plane deformation modes is explained and particular coupling effects are documented and discussed.

Acknowledgements

The financial support of the Slovak Research and Development Agency under the contract No. APVV-14-0440 is greatly acknowledged.

Appendix

The integrations required in definition of semi-integral fields can be performed in closed form and expressed in terms of two kinds of integrals

$$d_{(0)a} := \int_{-1/2}^{1/2} z^a dz = \frac{1 + (-1)^a}{(a+1)2^{a+1}} \quad (\text{A.1})$$

$$d_{(p)a} := \int_{-1/2}^{1/2} z^a \left(\frac{1}{2} \pm z \right)^p dz = (\pm 1)^a \int_0^1 \left(y - \frac{1}{2} \right)^a y^p dy = (\pm 1)^a \sum_{k=0}^a \binom{a}{k} \left(-\frac{1}{2} \right)^k \frac{1}{p+1+a-k}$$

In view of the definitions (10), (3), (6) and (A.1), the semi-integral fields are given as

$$T_{\alpha\beta}(\mathbf{x}) = C^{(uu)}(\mathbf{x})\tau_{\alpha\beta}^{(u)}(\mathbf{x}) + C^{(u\varphi)}(\mathbf{x})\tau_{\alpha\beta}^{(\varphi)}(\mathbf{x}) + C^{(uw)}(\mathbf{x})\tau_{\alpha\beta}^{(w)}(\mathbf{x}) + \sum_{a=0}^2 C^{(u\mathcal{G}_a)}(\mathbf{x})\tau_{\alpha\beta}^{(\mathcal{G}_a)}(\mathbf{x}) \quad (\text{A.2})$$

$$T_{3\alpha}^{(w\varphi)}(\mathbf{x}) = C^s(\mathbf{x}) \left[w_{,\alpha}(\mathbf{x}) + \varphi_{\alpha}(\mathbf{x}) \right]$$

$$M_{\alpha\beta}^{(\varphi)}(\mathbf{x}) = C^{(\varphi u)}(\mathbf{x})\tau_{\alpha\beta}^{(u)}(\mathbf{x}) + C^{(\varphi\varphi)}(\mathbf{x})\tau_{\alpha\beta}^{(\varphi)}(\mathbf{x}) + C^{(\varphi w)}(\mathbf{x})\tau_{\alpha\beta}^{(w)}(\mathbf{x}) + \sum_{a=0}^2 C^{(\varphi\mathcal{G}_a)}(\mathbf{x})\tau_{\alpha\beta}^{(\mathcal{G}_a)}(\mathbf{x})$$

$$M_{\alpha\beta}^{(w)}(\mathbf{x}) = C^{(wu)}(\mathbf{x})\tau_{\alpha\beta}^{(u)}(\mathbf{x}) + C^{(w\varphi)}(\mathbf{x})\tau_{\alpha\beta}^{(\varphi)}(\mathbf{x}) + C^{(ww)}(\mathbf{x})\tau_{\alpha\beta}^{(w)}(\mathbf{x}) + \sum_{a=0}^2 C^{(w\mathcal{G}_a)}(\mathbf{x})\tau_{\alpha\beta}^{(\mathcal{G}_a)}(\mathbf{x})$$

with the coefficients $C^{(\bullet\bullet)}(\mathbf{x})$ being given as

$$C^{(uu)}(\mathbf{x}) := 12 \frac{1-\nu}{H} \frac{D_0}{(h_0)^2} \left(d_{(0)0} + \zeta d_{(p)0} \right) D_{1H}(\mathbf{x}) \quad (\text{A.3})$$

$$C^{(u\varphi)}(\mathbf{x}) := 12 \frac{1-\nu}{H} \frac{D_0}{h_0} c_1 \zeta \left(d_{(p)1} - \frac{4}{3} c_2 d_{(p)3} \right) D_{2H}(\mathbf{x}) = C^{(\varphi u)}(\mathbf{x})$$

$$C^{(uw)}(\mathbf{x}) := 12 \frac{1-\nu}{H} \frac{D_0}{h_0} \zeta \left[(c_1 - 1) d_{(p)1} - \frac{4}{3} c_1 c_2 d_{(p)3} \right] D_{2H}(\mathbf{x})$$

$$C^{(u\mathcal{G}_a)}(\mathbf{x}) := -12 \frac{1-\nu}{H} \frac{D_0}{(h_0)^2} \theta_0 \alpha_0 \theta_0 \left(d_{(0)a} + \xi d_{(r)a} + \zeta d_{(p)a} + \xi \zeta d_{(p+r)a} \right) B_{1H}(\mathbf{x})$$

$$C^s(\mathbf{x}) := c_1 6(1-\nu) \frac{D_0}{(h_0)^2} \left\{ \left[(1-c_2)\kappa + c_2 \right] \left(d_{(0)0} + \zeta d_{(p)0} \right) - 8c_2 \left(d_{(0)2} + \zeta d_{(p)2} \right) + \right. \\ \left. + 16c_2 \left(d_{(0)4} + \zeta d_{(p)4} \right) \right\} D_{1H}(\mathbf{x})$$

$$C^{(\varphi\varphi)}(\mathbf{x}) := 12 \frac{1-\nu}{H} D_0 c_1 \left[d_{(0)2} - \frac{8}{3} c_2 d_{(0)4} + \frac{16}{9} c_2 d_{(0)6} + \right. \\ \left. + \zeta \left(d_{(p)2} - \frac{8}{3} c_2 d_{(p)4} + \frac{16}{9} c_2 d_{(p)6} \right) \right] D_{3H}(\mathbf{x})$$

$$\begin{aligned}
 C^{(\varphi w)}(\mathbf{x}) &:= -16 \frac{1-\nu}{H} D_0 c_1 c_2 \left[d_{(0)4} - \frac{4}{3} d_{(0)6} + \zeta \left(d_{(p)4} - \frac{4}{3} d_{(p)6} \right) \right] D_{3H}(\mathbf{x}) \\
 C^{(\varphi g_a)}(\mathbf{x}) &:= -12 \frac{1-\nu}{H} \frac{D_0}{h_0 \theta_0} \alpha_0 \theta_0 c_1 \left[d_{(0)a+1} - \frac{4}{3} c_2 d_{(0)a+3} + \xi \left(d_{(r)a+1} - \frac{4}{3} c_2 d_{(r)a+3} \right) + \right. \\
 &\quad \left. + \zeta \left(d_{(p)a+1} - \frac{4}{3} c_2 d_{(p)a+3} \right) + \xi \zeta \left(d_{(p+r)a+1} - \frac{4}{3} c_2 d_{(p+r)a+3} \right) \right] B_{2H}(\mathbf{x}) \\
 C^{(wu)}(\mathbf{x}) &:= 12 \frac{1-\nu}{H} \frac{D_0}{h_0} \zeta \left[(1-c_1) d_{(p)1} + \frac{4}{3} c_1 c_2 d_{(p)3} \right] D_{2H}(\mathbf{x}) = -C^{(uw)}(\mathbf{x}) \\
 C^{(wp)}(\mathbf{x}) &:= 16 \frac{1-\nu}{H} D_0 c_1 c_2 \left[d_{(0)4} - \frac{4}{3} d_{(0)6} + \zeta \left(d_{(p)4} - \frac{4}{3} d_{(p)6} \right) \right] D_{3H}(\mathbf{x}) = -C^{(\varphi w)}(\mathbf{x}) \\
 C^{(ww)}(\mathbf{x}) &:= 12 \frac{1-\nu}{H} D_0 \left\{ (c_1-1) d_{(0)2} - \frac{16}{9} c_1 c_2 d_{(0)6} + \zeta \left[(c_1-1) d_{(p)2} - \frac{16}{9} c_1 c_2 d_{(p)6} \right] \right\} D_{3H}(\mathbf{x}) \\
 C^{(wg_a)}(\mathbf{x}) &:= 12 \frac{1-\nu}{H} \frac{D_0}{h_0 \theta_0} \alpha_0 \theta_0 \left\{ (c_1-1) d_{(0)a+1} - \frac{4}{3} c_1 c_2 d_{(0)a+3} + \right. \\
 &\quad \left. + \xi \left[(c_1-1) d_{(r)a+1} - \frac{4}{3} c_1 c_2 d_{(r)a+3} \right] + \zeta \left[(c_1-1) d_{(p)a+1} - \frac{4}{3} c_1 c_2 d_{(p)a+3} \right] + \right. \\
 &\quad \left. + \xi \zeta \left[(c_1-1) d_{(p+r)a+1} - \frac{4}{3} c_1 c_2 d_{(p+r)a+3} \right] \right\} B_{2H}(\mathbf{x})
 \end{aligned}$$

where $D_{jH}(\mathbf{x}) := \left(h^*(\mathbf{x}) \right)^j E_H(\mathbf{x})$, $B_{jH}(\mathbf{x}) := \left(h^*(\mathbf{x}) \right)^j E_H(\mathbf{x}) \alpha_H(\mathbf{x})$, $h^*(\mathbf{x}) = h(\mathbf{x}) / h_0$.

References

- [1] Nowacki, W. (1975) *Dynamic Problems of Thermoelasticity*, Noordhoff International Publishing, Leyden, The Netherlands.
- [2] Sator, L., Sladek, V., Sladek, J. (2014) Coupling effects in elastic analysis of FGM composite plates by mesh-free methods, *Composite Structures* **115**, 309–322.
- [3] Sator, L., Sladek, V., Sladek, J. (2016) Elastodynamics of FGM plates by meshfree method, *Composite Structures* **140**, 100–110.
- [4] Sladek, V., Sladek, J., Sator, L. (2013) Physical decomposition of thin plate bending problems and their solution by mesh-free methods, *Engineering Analysis with Boundary Elements* **37**, 348–365.
- [5] Lancaster, P., Salkauskas, K. (1981) Surfaces generated by moving least square methods, *Math Comput* **37**, 141–158.
- [6] Sladek, V., Sladek, J., Zhang, Ch. (2008) Computation of stresses in non-homogeneous elastic solids by local integral equation method: a comparative study, *Computational Mechanics* **41**, 827–845.
- [7] Sladek, V., Sladek, J. (2010) Local integral equations implemented by MLS approximation and analytical integrations, *Engineering Analysis with Boundary Elements* **34**, 904–913.
- [8] Atluri, S.N. (2004) *The meshless method, (MLPG) for domain & BIE discretization*, Tech. Science Press, Forsyth.
- [9] Sator, L., Sladek, V., Sladek, J. (2017) Multi-gradation coupling effects in FGM plates, *Composite Structures* **171**, 515–527.
- [10] Sladek, V., Sator, L., Sladek, J. (2018) Thermoelastic analysis of bending problems in FGM plates, *Int J Comp Meth and Exp Meas* **6**, 1161–1172.
- [11] Bever, M.B., Duwez, P.E. (1972) Gradients in composite materials, *Mater Sci Eng* **10**, 1–8.
- [12] Jha, D.K., Kant, T., Singh, R.K. (2013) A critical review of recent research on functionally graded plates, *Composite Structures* **96**, 833–849.
- [13] Suresh, S., Mortensen, A. (1998) *Fundamentals of Functionally Graded Materials*, Institute of Materials, London.

- [14] Koizumi, M. (1993) The concept of FGM, *Ceramic Transactions* **34**, 3-10.
- [15] Yamanouchi, M., Koizumi, M., Hirai, T., Shioda, I. (1990) *Proceedings of the First International Symposium on Functionally Graded Materials*, FGM Forum, Tokyo.
- [16] Reddy, J.N. *Mechanics of Laminated Composite Plates: Theory and Analysis*, CRC Press, Boca Raton.
- [17] Tornabene, F. (2009) Free vibration analysis of functionally graded conical, cylindrical shell and annular plate structures with four-parameter power-law distribution, *Comput Meth Appl Mech Engng* **198**, 2911-2935.
- [18] Tornabene, F., Viola, E. (2009) Free vibration of four-parameter functionally graded parabolic panels and shells of revolution, *Eur Jour Mech A/Solids* **28**, 991-1013.
- [19] Tornabene, F., Fantuzzi, N., Baccocchi, M., Viola, E., Reddy, J.N. (2017) A numerical investigation on the natural frequencies of FGM sandwich shells with variable thickness by the local generalized differential quadrature method, *Applied Sciences* **7**, 131.
- [20] Ardestani, M.M., Soltani, B., Shams, S. (2014) Analysis of functionally graded stiffened plates based on FSDT utilizing reproducing kernel particle method, *Composite Structures* **112**, 231-240.
- [21] Barber, J.R. (2010) *Elasticity*, Springer Science+Business Media B.V., New York.
- [22] Lurie, A.I. (2005) *Theory of Elasticity*, Springer-Verlag, Berlin.
- [23] Mantari, J.L., Oktem, A.S., Guedes Soares, C. (2012) Bending response of functionally graded plates by using a new higher order shear deformation theory, *Composite Structures* **94**, 714-723.

2016-2017 Central Italy Earthquake: Seismic Assessment of "Pietro Capuzi" School in Visso (Marche)

†*Chiara Ferrero¹, Paulo B. Lourenço², and Chiara Calderini¹

¹Department of Civil, Chemical and Environmental Engineering (DICCA), University of Genoa, Genoa, Italy.

²ISISE, Department of Civil Engineering, University of Minho, Guimarães, Portugal.

*Presenting author: chiara.ferrero@edu.unige.it

†Corresponding author: chiara.ferrero@edu.unige.it

Abstract

A prolonged seismic sequence struck the regions of Central Italy between August 2016 and January 2017, causing several fatalities and widespread damage to the built environment. The main objective of this work was to study the structural and seismic behavior of “Pietro Capuzi” school, located in Visso, in the Marche region, which was severely damaged by the 2016-2017 Central Italy Earthquake. A 3D finite element (FE) model of the entire school was prepared, adopting a macro-modelling approach to represent masonry materials. An eigenvalue analysis was initially performed in order to identify the dominant modes of vibration of the structure and calibrate the numerical model according to the results of the dynamic identification tests. Afterwards, non-linear static analyses were performed on the calibrated FE model to evaluate the seismic response of the structure. Finally, the numerical results obtained in terms of failure mechanisms and seismic capacity were compared with the real damage experienced by the building. The numerical model proved to accurately predict the seismic response exhibited by the structure during the past seismic events.

Keywords: Unreinforced masonry, finite element modeling, seismic assessment, non-linear analysis

Introduction

Recent and past earthquakes have shown that unreinforced masonry buildings are prone to damage by seismic actions. However, the seismic assessment of such structures is a highly demanding task, not only for the potential complexity of this kind of buildings, but also for the lack of data usually available regarding geometry, construction details and material mechanical characterization. This study was aimed at assessing the seismic performance of “Pietro Capuzi” school, located in the municipality of Visso (Marche, Italy), which was severely damaged by the 2016-2017 Central Italy earthquake, also named the Amatrice-Visso-Norcia seismic sequence. The school was an excellent case study to investigate the seismic capacity of an existing masonry structure for several reasons. Firstly, “Pietro Capuzi” school is part of the public buildings permanently monitored by the Seismic Observatory of Structures (hereafter named OSS). As a result, the entire seismic sequence affecting the structure was recorded by the existing system of accelerometers [1], thus providing valuable information in terms of seismic input at the base as well as vibrations experienced at different levels. Secondly, the inspection and extensive experimental campaign that were performed in 2011 on behalf of the OSS supplied detailed documentation regarding geometry, structural configuration, construction details as well as some information about the mechanical properties of masonry [2][3][4]. Furthermore, dynamic identification tests were also carried out in 2011 to characterize the dynamic response of the structure [5]. Such documentation, acquired from the OSS, was integrated with the information about geometry, structural configuration, past interventions and seismic damage supplied by the Italian Network of

University Laboratories in Seismic Engineering (RELUIS) and the University of Genova, in charge of post-earthquake surveys [6][7]. In addition, the photographic record of the progressive damage experienced by the building during the entire seismic sequence was provided by the Italian Department of Civil Protection [8]. In conclusion, the availability of such detailed data, which is not very common, allowed for the preparation of a detailed FE model of the entire structure. This was then updated on the basis of the modal parameters determined experimentally by the OSS. Furthermore, the detailed description of the damage experienced by the structure during the entire seismic sequence provided the opportunity to validate the numerical model of “Pietro Capuzi” school through a comparison between the damage simulated and the damage pattern observed.

Description of the building

Overall configuration and use

Built in the 1930s, “Pietro Capuzi” nursery and primary school (Figure 1) is a stand-alone structure located in the urban area of Visso, in the Marche region. The building has four levels, three of them above ground (raised ground floor, first floor and attic) and a basement partially sub-grade (Figure 1c). The three levels above ground have a plan area of approximately 605 m² each, whereas the basement has a plan area of approximately 120 m². The basement, raised ground floor and first floor have an inter-storey height of 3.09 m, 4.44 m and 4.26 m respectively. The attic has a maximum height of about 3.19 m, resulting in a total height above ground of about 13.50 m in correspondence of the main façade. The total built volume is estimated at about 4800 m³.

The building does not present any narrowing or enlargement of the horizontal section along the height in the part above ground. However, it is strongly irregular in plan since it presents a T-shape configuration given by the connection of two orthogonal bodies (hereafter named Body A and Body B as shown in Figure 1b). Body A, characterized by an elongated rectangular shape and oriented in NW-SE direction (hereafter called X direction as indicated in Figure 1b), hosts most of the classrooms, while the areas used as offices, the canteen and toilets are allocated in Body B. A staircase connecting the basement, raised ground floor and the first floor is located on the east side of Body B.

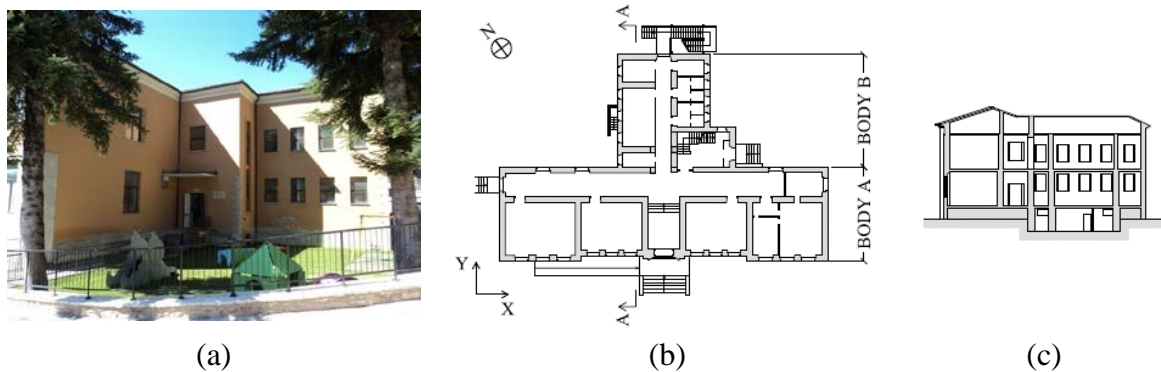


Figure 1 – “Pietro Capuzi” school: (a) external view, (b) plan of the raised ground floor, and (c) section AA [2].

Structural configuration and technical details

The structure of the building consists of load-bearing masonry walls that extend upwards from the basement until the attic. The prevalent type of masonry is a stone masonry, whereas

the sporadic presence of solid brick masonry can be observed in some pillars in the staircase as well as in portions of walls filled with solid bricks during past interventions. The thickness of the walls varies with height, ranging from about 70-85 cm at the basement to approximately 70 cm at raised ground floor and 50-70 cm at first floor. Foundations are made of the same stone masonry of the upper levels that extend downwards until a depth of approximately 60 cm from the planking level of the basement.

Concerning horizontal diaphragms, the slabs of the raised ground floor and first floor are lightweight slabs, with a thickness of 30 cm each, while the slab between first floor and the attic is a steel-clay slab. In the areas of the raised ground floor where there is no basement, the slab on grade consists of a concrete slab resting directly on the ground. The staircase is made of reinforced concrete, and it is supported on load-bearing masonry walls on three sides, and three masonry pillars on the internal short side. The hipped roof has a timber structure consisting of purlins and rafters that are supported by a system of trusses or by the inner walls extending until the top of the roof. In correspondence of ridges and valleys, hip rafters are present. Above the timber structure, clay tiles are located. A reinforced concrete bond beam is located at the height of the spring line of the roof.

Past damage and previous strengthening interventions

“Pietro Capuzi” school was damaged by the seismic events that hit the regions of Umbria and Marche between 1997 and 1998. Cracks mainly appeared in the staircase where the two bodies composing the building connect to each other. This led to a strengthening intervention performed in the 1990s, which was aimed at: (1) repairing the seismic damage, (2) improving the seismic capacity of the building, and (3) solving static problems due to the decay of some parts of the timber roof. As for the latter, the original roof timber structure was replaced with new glue laminated timber elements in the central part of Body A above the classrooms. As for the seismic damage, the major cracks were repaired using the technique of “cuci and scuci”. Regarding the seismic improvement, the interventions carried out included: (1) addition of metallic profiles studded to the internal side of masonry walls at the height of the slab between first floor and attic, (2) connection of the roof timber elements to the perimeter walls by means of metallic plates anchored with bars, (3) creation of a steel frame around the openings of the main façade at the first level, (4) addition of metallic tie-rods, (5) injections of lime mortar in some piers where pipes were filled in the past with solid brick masonry, and (6) filling with solid brick masonry of the spans between the pillars in the staircase.

Seismic damage and monitoring data

The 2016-2017 Amatrice-Visso-Norcia seismic sequence started on August 24th, 2016 with the Amatrice earthquake (M_w 6.0), which hit a vast area of the Central Apennines producing almost 300 casualties and widespread damage to the built environment. The sequence was characterized by nine mainshocks with moment magnitudes higher than 5, which occurred on August 24th, 2016 (M_w 6.0 and M_w 5.4), October 26th, 2016 (M_w 5.4 and M_w 5.9), October 30th, 2016 (M_w 6.5) and January 18th, 2017 (four shakes with $M_w \geq 5.0$).

Damage suffered in the 2016-2017 Amatrice-Visso-Norcia seismic sequence

“Pietro Capuzi” school suffered severe damage due to the Amatrice-Visso-Norcia seismic sequence. A detailed description of the damage experienced by the building, updated on the date of December 8th, 2016, was reported in [6][7]. As shown in Figure 2a, the walls were affected by severe cracks at both the raised ground floor and first floor, especially in the Y direction where the damage is more severe and widespread than in the X direction. The building mainly presents diagonal shear cracks developing through the entire thickness of the

walls in both piers and spandrels. Furthermore, a shear-sliding mechanism producing an important residual deformation was activated in the northwest (NW) side of the building (Figure 2b). Flexural cracks can also be observed on the southeast (SE) side of Body B (Figure 2c). In addition to the in-plane damage, an out-of-plane mechanism occurred in the NW corner of the building, resulting in the partial collapse of the masonry walls and adjacent attic slabs (Figure 2d).

With regards to the other slabs, two local collapses occurred in the slab between the raised ground floor and first floor (Figure 2e). Furthermore, the steel slab between first floor and attic experienced an extensive collapse because of the out-of-plane mechanisms as well as cracks along the metallic profiles in the portions of slab still standing. For the connection between vertical walls and slabs, horizontal cracks were present in the external walls at the slab locations. In the interior, severe cracks were observed in correspondence of the connection between walls and slabs (Figure 2f).

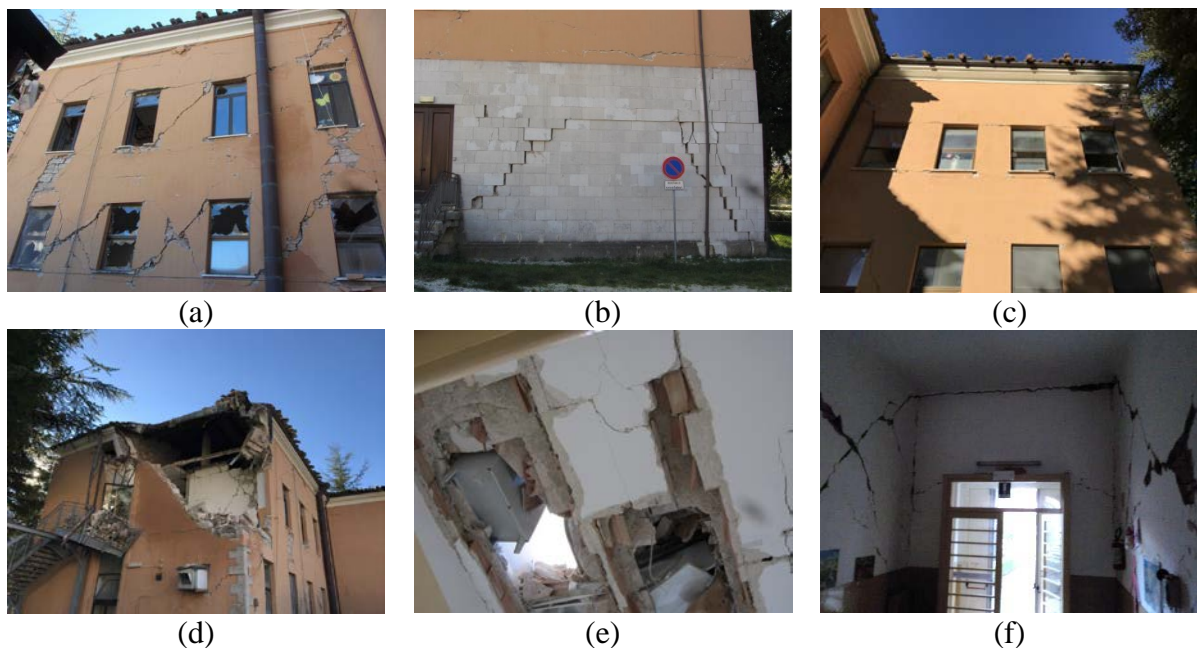


Figure 2 – Damage state updated on December 8th, 2016: (a) diagonal shear cracks in piers and spandrels, (b) shear-sliding mechanism, (c) out-of-plane mechanism, (d) flexural cracks, (e) local collapse of the slab between raised ground floor and first floor, (f) cracks in correspondence of the connection between walls and slabs [6].

A partial reconstruction of the damage development experienced by “Pietro Capuzi” school during the Amatrice-Visso-Norcia seismic sequence was carried out thanks to the photographic documentation provided in [8]. After the seismic event of August 24th, the building exhibited an in-plane response with cracks occurring in both masonry piers and spandrels (Figure 3a). In addition to moderate cracks widespread throughout the entire building, severe cracking appeared in some piers of internal and external walls. As shown in Figure 3b, the earthquakes that occurred on October 26th, 2016 produced significant deterioration of the damage level. Not only did the severity of the in-plane damage increase significantly, but an out-of-plane mechanism also occurred. As for the in-plane response of masonry walls, it was observed that, more than in the creation of new cracks, the deterioration of the damage state mainly consisted in the enlargement of the cracks produced by the seismic event of August 24th. As shown in Figure 3c, the damage state did not change significantly after the shake of October 30th with respect to October 26th.

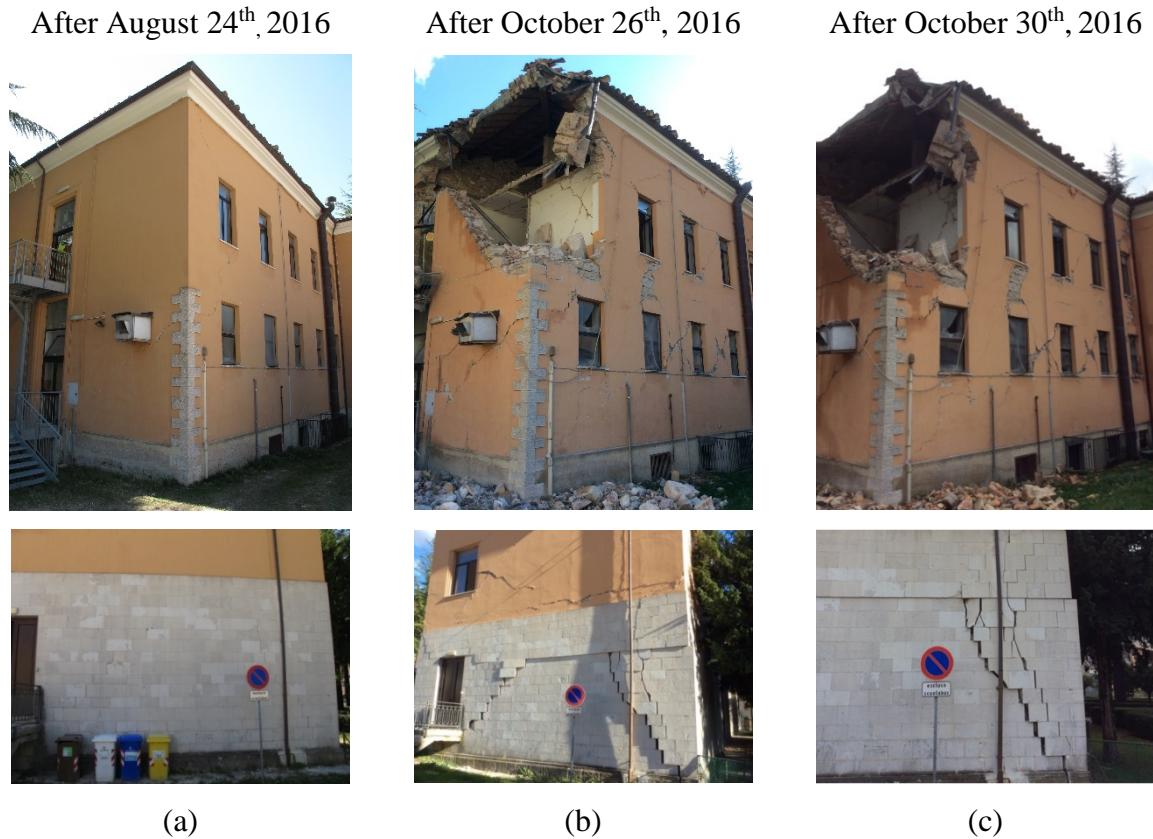


Figure 3 – Observed damage in “Pietro Capuzi” school after the earthquakes occurred on: (a) August 24th, 2016, (b) October 26th, 2016, and (c) October 30th, 2016 [8].

Permanent Monitoring by the Seismic Observatory of Structures

As part of the OSS, “Pietro Capuzi” school is permanently monitored by a system of accelerometers located in different parts of the structure: one tri-axial accelerometer is located in the basement to measure the seismic input, and ten bi-axial accelerometers are placed at the first and second levels at the intrados of the slabs to record the accelerations experienced by the structure. For all of these sensors, the OSS provided the authors with the time histories in acceleration and displacement obtained for the seismic events occurred on August 24th (M_w 6.0), October 26th (M_w 5.4 and 5.9) and October 30th (M_w 6.5) [1][6]. These recordings allowed characterization of the seismic input at the base of the structure and aided in drawing significant conclusions about the accelerations experienced by the building.

For the four seismic events considered, Table 1 reports the distance of the epicenter from “Pietro Capuzi” school as well as the values of the peak ground acceleration (PGA) obtained for the two horizontal components XX and YY, and vertical component ZZ. The structure was subjected to significant values of horizontal ground motion, in particular the highest values of horizontal PGA were recorded during the M_w 5.9 earthquake of October 26th, which are equal to 0.36 g and 0.47 g in the XX and YY directions, respectively.

Table 1 – Distance from the epicenter and PGA for the seismic events of August 24th, October 26th and October 30th.

Seismic event	Distance from epicenter [km]	PGA x [g]	PGA y [g]	PGA z [g]
2016/08/24_M _w = 6.0	28	0.33	0.32	0.14
2016/10/26_M _w = 5.4	7	0.30	0.21	0.41
2016/10/26_M _w = 5.9	4	0.36	0.47	0.31
2016/10/30_M _w = 6.5	10	0.29	0.30	0.33

Figure 4 shows a comparison between the 5% damped acceleration response spectra of the horizontal components of motion in XX and YY directions for the four earthquakes considered, and the elastic response spectrum in accordance with the Italian Building Code (hereafter called NTC2008 [9]) for a return period of 712 years and soil of type B [3]. It is observed that the seismic events of August 24th (M_w 6.0) and October 26th (M_w 5.9), which caused the most severe damage to the structure, resulted in significant amplifications of the acceleration with respect to the elastic spectrum of NTC2008. Furthermore, it is important to note that the YY component of the acceleration spectrum of October 26th earthquake exceeded the code spectrum for all the significant ranges of period, thus agreeing with the severe damage suffered by the structure in this direction.

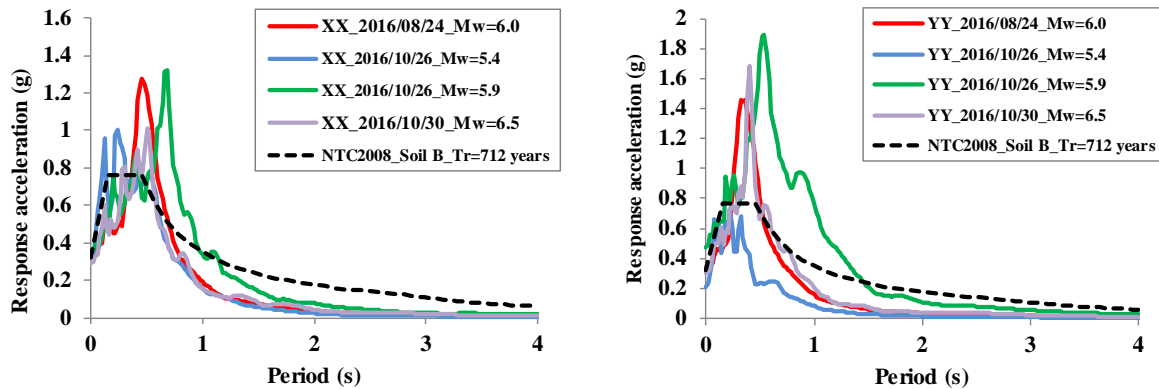


Figure 4 - Comparison between the response spectra (XX and YY component) of the seismic events of August 24th, October 26th and October 30th and the elastic spectrum of NTC2008 for soil B and return period Tr=712 years (damping 5%).

The monitoring system installed in the school provided useful information to characterize the seismic performance of the building during the entire seismic sequence. Table 2 reports the values of the maximum acceleration recorded on the structure (PSA), the amplification factor ($\alpha_{PSA/PGA}$, calculated as the ratio between PSA and PGA), and the maximum inter-storey drift (D_{max}, i.e. the maximum ratio of the relative floor displacement to the height of the corresponding floor), for the seismic events on August 24th (M_w 6.0), October 26th (M_w 5.4 and 5.9) and October 30th (M_w 6.5). First, it was observed that the structure underwent values of acceleration that were significantly amplified with respect to the maximum acceleration measured at the base. Indeed, very high values of amplification factors, even higher than 4, were obtained. Furthermore, some interesting observations could be drawn analyzing the values of inter-storey drift produced during the seismic sequence. With this aim, the OSS

already provided the following reference values regarding drift-damage relation for masonry buildings: (1) no damage for drift ranging from 0 to 2‰, (2) slight damage for drift ranging from 2‰ to 4.5‰, (3) moderate damage for drift ranging from 4.5‰ to 8‰, and (4) severe damage for drift higher than 8‰ [10]. The highest values of drift were produced by the M_w 5.9 earthquake of October 26th, for which values of 13.53 and 16.01 were obtained in X and Y direction respectively. These values were significantly higher than the ones resulting from the shakes of August 24th and October 26th (M_w 5.4). According to the reference values provided by OSS, the values of drift obtained for the M_w 5.9 earthquake of October 26th corresponded to serious damage, while moderate and slight damage were respectively associated to the shakes of August 24th and October 26th (M_w 5.4). Note that the values of drift obtained for the M_w 6.5 earthquake of October 30th were less meaningful since the building had been severely damaged by the previous seismic events. These results were consistent with the damage level exhibited by the building during the entire seismic sequence. Furthermore, the higher values of drift obtained in the Y direction for the seismic events of October agree with the more severe damage observed in the structure in the Y direction with respect to the X direction. Consequently, the inter-storey drift was found to represent a reliable measure of the damage suffered by the structure after the earthquake.

Table 2 – Parameters characterizing the seismic response of the structure for the seismic events of August 24th, October 26th and October 30th.

Seismic event	PSA x [g]	PSA y [g]	$\alpha_{PSA/PGA,x}$	$\alpha_{PSA/PGA,y}$	Dmax x [‰]	Dmax y [‰]
2016/08/24_ M_w = 6.0	1.05	0.80	3.21	2.52	6.10	4.23
2016/10/26_ M_w = 5.4	1.27	0.78	4.27	3.70	3.95	2.46
2016/10/26_ M_w = 5.9	1.41	1.33	3.86	2.81	13.53	16.01
2016/10/30_ M_w = 6.5	1.36	1.47	4.65	4.89	5.52	10.98

Numerical model

Preparation of the FE model

A 3D finite element (FE) model of “Pietro Capuzi” school was created in Midas FX+ Version 3.3.0 Customized Pre/Post-processor for DIANA software [11]. A macro-modelling approach was used to represent masonry, which was considered as a composite material without any distinction between units and mortar [12]. Geometry definition is essential when dealing with the modeling of a complex structure [13] since any adopted strategy may entail different final results. “Pietro Capuzi” school has three levels above the ground and a basement partially sub-grade. Modeling building portions below ground is often controversial, especially when detailed information about the foundations is not available. Hence, the first decision to be made was the basement modeling strategy, and three different models were considered, as illustrated in Figure 5. As shown in Figure 5a, the basement occupies only a part of the plan area of the first level of Body B. Model A neglects the portion of the basement underground (Figure 5b), whereas model B and C take it into consideration by extending the walls of the basement downwards. The main distinction between models B and C is that the first one considers only the walls where the basement is located according to the geometrical survey (Figure 5c), whereas the second one assumes an equal height underground (1.47m) along all the walls of Body B, in agreement with past modelling strategies [4]. Note that the passive earth pressure exerted on the walls of the basement was not considered due to the limited

height of the portion of the basement underground. The choice of the model to use for structural analyses was based on the comparison between the experimental and numerical responses of the different models in terms of frequencies and mode shapes (see the following paragraph). Note that the three models only differ for the presence of the basement, therefore the description of geometry and materials presented below is applicable to all.

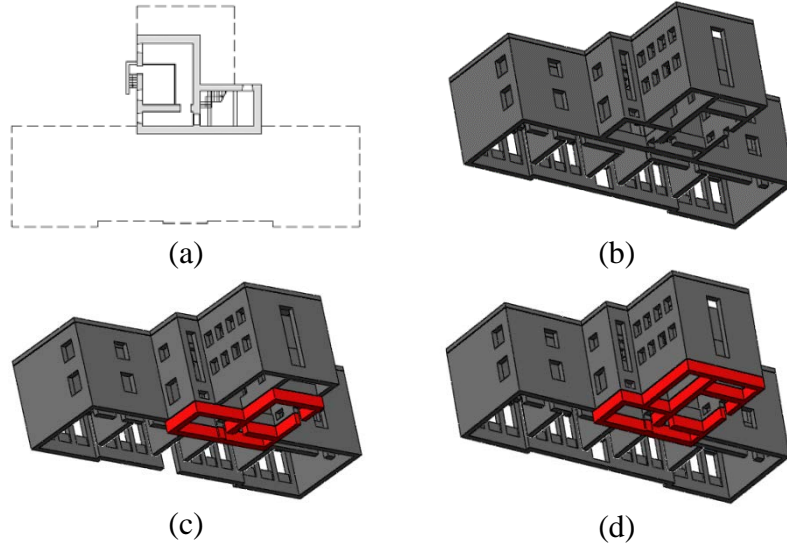


Figure 5 – (a) Plan of the basement, and geometry of the different FE models prepared for the school: (a) model A, (b) model B, (c) model C.

Masonry walls were modeled using solid FEs; in particular, four-node three-sided isoparametric solid tetrahedron elements (TE12L) were adopted [11]. Although the use of shell elements for the walls would have resulted in a significant reduction of the number of degrees of freedom and, consequently, a more limited computational effort, the strategy of using solid FEs was adopted because of the presence of masonry panels characterized by a similar length in both axial and transversal directions.

Regarding diaphragms, three-node triangular (T15SH) and four-node quadrilateral (Q20SH) isoparametric curved shell elements were adopted to model the intermediate slabs [11]. The roof was built as an ensemble of inclined surfaces supported by masonry walls (Figure 6a). Since a detailed survey of the roof timber structure was not available, it was believed that this solution might allow an adequate load distribution on the perimeter walls. In the case of the roof, only three-node triangular isoparametric curved shell elements were used to assure a better-quality mesh and prevent the creation of elements with undesirable shapes in the corners.

Finally, 1D elements were used to model beams and tie-rods. Two-node, three-dimensional class-I beam elements (L12BE) were adopted for the reinforced concrete beams located at the entrance and in the staircase as well as the reinforced concrete bond-beam present at the top of the building. Two-node regular directly integrated (1-point) truss elements (L2TRU) were adopted to mesh metallic tie-rods [11].

The geometry and final mesh of the school are shown in Figure 6 for model B. In total, the numerical model is composed by 180.567 nodes and 696.997 elements in the case of model A, 187.988 nodes and 727.853 elements as for model B, and 189.256 nodes and 730.953

elements in the case of model C. Regarding the boundary conditions, a clamped boundary condition was adopted at the base of the walls, and all the degrees of freedom were restrained.

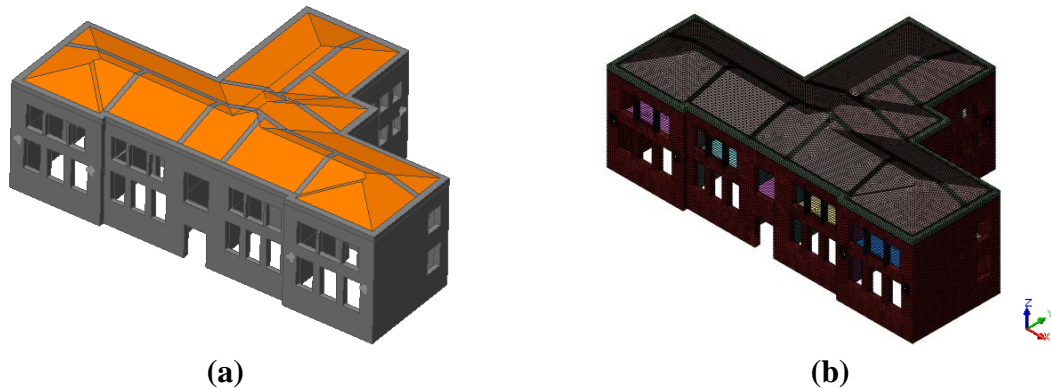


Figure 6 – Three-dimensional model: (a) geometry and (b) mesh discretization.

Material and diaphragm properties

A description of the properties adopted for materials and diaphragms is presented below. To perform numerical analyses, a non-linear behavior was adopted exclusively for masonry materials, whereas a linear elastic behavior was employed for slabs, reinforced concrete beams and tie-rods. The physical non-linear behavior of masonry was simulated by means of the Total Strain Rotating Crack Model that is available in DIANA [11]. To represent masonry behavior in tension and compression, an exponential stress-strain relationship and a parabolic stress-strain relationship were adopted, respectively. Three different types of masonry were identified in the building and represented in the FE model: (1) cut stone masonry with good texture, (2) stone masonry injected during past interventions, and (3) solid brick masonry.

The physical and mechanical properties of the different types of masonry adopted in the numerical model are reported in Table 3. The elastic properties of masonry materials were defined according to the prescriptions given by the Italian Circolare for knowledge level LC2 [14], considering also the qualitative and quantitative information obtained from past inspections and tests [3][6]. Some corrective coefficients, as indicated in [14], were adopted to improve the properties of stone and brick masonry due to the presence of injections and good mortar, respectively. As for the inelastic properties, they were determined on the basis of information and recommendations available in literature [15][16].

Table 3 - Material properties of the different types of masonry adopted in the numerical model.

Material property	Stone masonry	Stone masonry (+ injections)	Brick masonry (+ good mortar)
Specific weight [kN/m ³]	21	21	18
Elasticity modulus [MPa]	1740	2610	2250
Poisson's ratio [-]	0.2	0.2	0.2
Compressive strength [MPa]	2.67	4.00	4.00
Compressive fracture energy [N/mm]	4.27	6.40	6.40
Tensile strength [MPa]	0.108	0.163	0.190
Tensile fracture energy [N/mm]	0.024	0.024	0.024

The physical and mechanical properties adopted for reinforced concrete and steel are presented in Table 4. Since no physical or mechanical characterization was available, the properties of these materials were derived from NTC2008 [9].

Table 4 – Material properties of reinforced concrete and steel.

Material property	Concrete	Steel
Specific weight [kN/m ³]	25	78.5
Modulus of elasticity [MPa]	31500	210000
Poisson's ratio [-]	0.2	0.29

Regarding the slabs, the only properties to determine were bending and axial stiffness since a linear elastic behavior was adopted for them. These properties are automatically calculated in DIANA on the basis of the elasticity modulus of the material assigned to the slabs and the thickness of the FEs (curved shell elements in this case) used to model them. In this study, the values of an equivalent modulus of elasticity E_{eq} and an equivalent thickness h_{eq} to input in DIANA were calculated based on the real bending and axial stiffness of a strip of slab as wide as the spacing between the principal elements composing the slab structure.

An isotropic or orthotropic material was adopted for two-way and one-way slabs, respectively. For the former, the same value of stiffness was adopted in the two in-plane directions, whereas for the latter, stiffness equal to the 10% of the principal direction was assumed in the secondary direction. As for the stiffness in the orthogonal direction (vertical direction), it was estimated so that it fulfilled the requirements of orthotropic elasticity reported in [11].

Table 5 presents the in-plane axial stiffness in both principal and secondary directions (E_1h and E_2h) and the bending stiffness in the vertical direction ($E_3h^3/12$) for the different types of slabs.

Table 5 - Properties of diaphragms adopted in the numerical model.

Type of diaphragm	Type of material	E_1h (kN/m)	E_2h (kN/m)	$E_3h^3/12$ (kNm)
Lightweight concrete slab (one-way)	orthotropic	3.001E+06	3.001E+05	7.051E+03
Lightweight concrete slab (two-way)	isotropic	2.644E+06	2.644E+06	1.061E+04
Steel slab (one-way)	orthotropic	3.822E+05	3.822E+04	6.618E+02
Roof	isotropic	1.889E+06	1.889E+06	5.666E+02

Eigenvalue analysis and model updating

A first eigenvalue analysis was carried out to obtain the natural frequencies and modes shapes of the three different models prepared for “Pietro Capuzi” school. In order to identify which model better simulated the real behavior of the structure, the numerical results were compared with the ones derived from the dynamic identification tests performed by [5] in terms of natural frequencies and mode shape. The OSS provided the natural frequencies as well as the mode shape vectors for three vibration modes. For these modes, the deformed shape, shown

in Figure 7a for the first level of the building, was obtained in this study by drawing in scale on the plan of the building the modal displacements associated to each mode shape vector. The modes identified experimentally were global modes corresponding to: (1) a translational mode in the transversal direction of the building (mode 1 - 3.175 Hz), (2) a torsional mode (mode 2 - 3.755 Hz), and (3) a translational mode in the longitudinal direction (mode 3 - 4.047 Hz).

In order to gather more information about the modal response of the structure and visualize animations of the mode shapes, a further dynamic identification was performed in ARTeMIS Modal 5.0 software [17] by processing the signals recorded by [5] during the dynamic tests carried out in 2011. Though richer in terms of more possible frequencies, this dynamic identification whose results are reported in [18], was consistent with the one carried out by [5]. The latter was used by the authors to perform the updating of the numerical model. Hence, the model updating was focused on the three modes identified in [5], which were compared with the ones obtained numerically that presented a similar mode shape (Figure 7b). Note that only the mode shapes obtained for model B are presented in Figure 7b due to the similarity of results obtained for the three models.

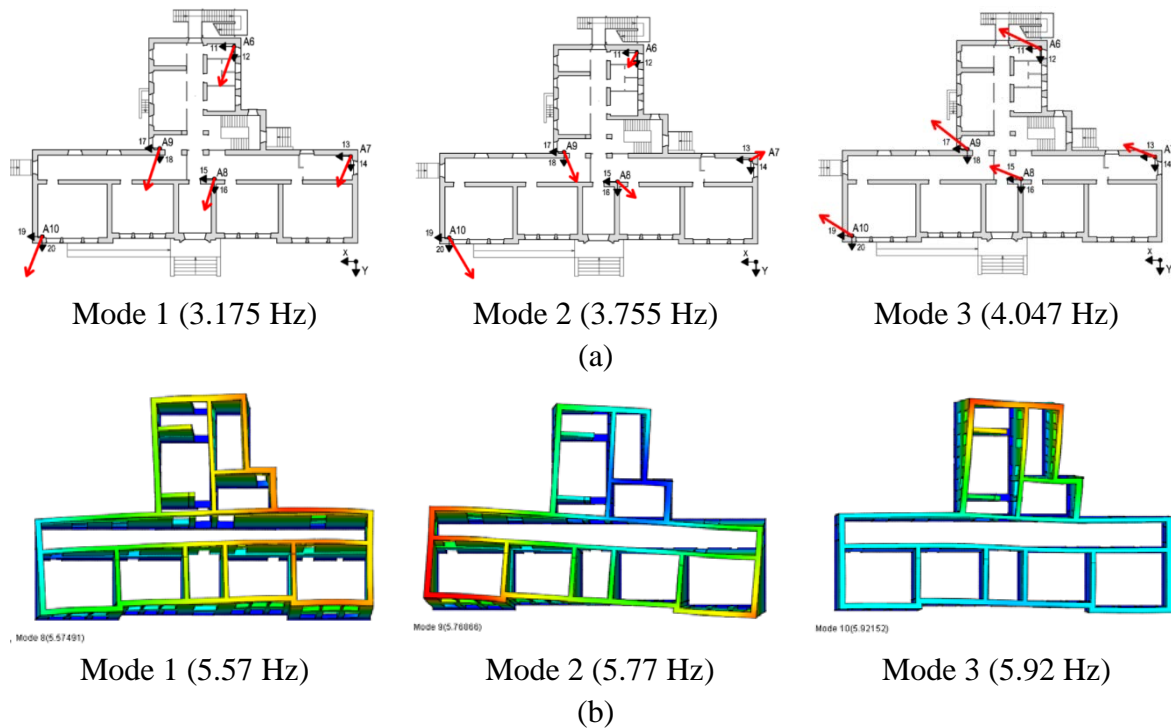


Figure 7 - Mode shapes obtained (a) experimentally and (b) numerically (for model B).

Subsequently, for the three global modes identified, a comparison between numerical and experimental results was performed in terms of frequency error and modal assurance criterion (MAC). The latter is a statistical indicator that is normally used to compare the mode shapes obtained from analytical or numerical models with the ones obtained experimentally [19].

Table 6 presents the average frequency errors and MACs for the three models. A similar frequency error of about 60%, calculated on average between the frequencies of the three modes, was obtained for the three different models. As for the average of MACs, the value obtained for model C (0.53) was much lower than the ones obtained for model A (0.68) and model B (0.69). Since the average frequency error was slightly lower for model B (58.5%)

when compared with model A (59.8%), model B was considered the model that better simulates the real behavior of the structure, and it was used for model updating and further structural analyses.

Table 6 - Comparison of numerical and experimental frequencies and mode shapes

	Model A	Model B	Model C
Average frequency error [%]	59.8	58.5	56.2
Average MAC	0.68	0.69	0.53

As shown in Table 6, the average of the frequency errors for model B is significantly high, consequently model updating was necessary to calibrate the numerical model of the school. Two different strategies of model updating were adopted considering different parameters as variables to tune and using an iterative procedure in order to minimize frequency error and MAC.

Table 7 presents the results obtained for the two calibrations carried out. The first calibration was performed adjusting the properties of masonry materials and diaphragms. Concerning masonry materials, the elasticity modulus of stone and brick masonry was considered for the updating process. As for the slabs and roof, the axial stiffness was varied while keeping the bending stiffness constant. However, this last strategy was disregarded since considerable reductions of the axial stiffness of the diaphragms produced only limited decreases in the values of the numerical frequencies, but, at the same time, resulted in a sharp decrease of the average MAC. On the other hand, since the natural frequencies obtained numerically were significantly higher than the experimental ones, a sharp reduction of the initial values adopted for the elasticity modulus of masonry was needed to reach a reasonable average frequency error of about 6%. The latter was derived considering an elasticity modulus of 700 MPa, 1050 MPa and 905 MPa respectively for stone masonry, stone masonry with injections and brick masonry. Although these values were still within the range of values provided by the Italian Circolare [14], such a significant reduction may indicate that masonry was poorly built, was damaged when the dynamic tests were performed, or soil conditions significantly influenced the measures of frequency values. For these reasons, a further calibration was carried out considering soil-structure interaction and adopting a finite stiffness for the soil.

To model the soil, plane interface elements (T18IF [11]) were placed at the base of the walls in the numerical model. The values of normal stiffness modulus k_n and shear stiffness modulus k_t that must be input for the interfaces were adopted as the variable to update. On the basis of the values of dynamic Young's modulus and dynamic shear modulus of the soil (as reported in [3]), the reference values of the normal stiffness modulus and shear stiffness modulus of the interfaces were assumed as $1.91\text{E}+05 \text{ kN/m}^3$ and $6.83\text{E}+04 \text{ kN/m}^3$, respectively. A new eigenvalues analysis was performed on the numerical model with interfaces at the base. Model updating was then carried out adjusting the values of normal and shear stiffness moduli of the interfaces until achieving an error between experimental and numerical frequencies lower than 5% for both each mode and on average. Such an error was obtained adopting a value of $1.32\text{E}+05 \text{ kN/m}^3$ for the normal stiffness modulus and a value of $4.71\text{E}+04 \text{ kN/m}^3$ for the shear stiffness modulus.

Table 7 – Results obtained from different updating strategies.

Updating variables	Frequency error [%]			MAC			Average f. error [%]	Average MAC
	Mode 1	Mode 2	Mode 3	Mode 1	Mode 2	Mode 3	All modes	All modes
E_{masonry} reference values	75.2	53.0	47.2	0.73	0.77	0.59	58.5	0.70
E_{masonry} updated values	13.5	-1.3	-3.5	0.73	0.79	0.63	6.1	0.72
k_n, k_t interfaces reference values	14.4	5.2	6.1	0.88	0.57	0.70	8.6	0.72
k_n, k_t interfaces updated values	4.3	-2.8	-2.0	0.89	0.55	0.71	3.1	0.72

According to the results reported in Table 7, it can be concluded that the average of MACs is almost insensitive to the updating strategy employed, whereas the model updating based on the stiffness of interfaces resulted in a better matching between the natural frequencies obtained experimentally and numerically. Not only is the error, on average, almost the half of the one obtained by updating the elasticity modulus of masonry, but the error in the frequency of the first mode also comes down to a value of 4.3%, which is significant lower than the 13.5% reached with the first calibration.

In conclusion, the calibrated FE model with interfaces was chosen to perform further structural analyses. For this model, Figure 8 presents the three global mode shapes obtained numerically and corresponding to the experimental ones.

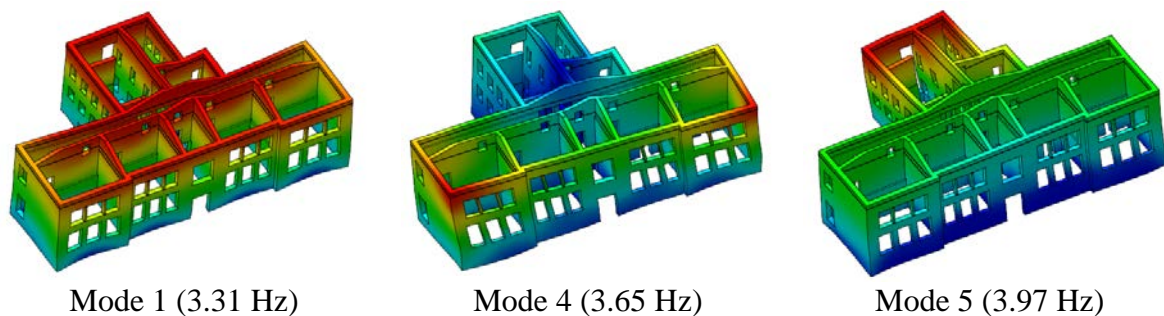


Figure 8 – Mode shapes of the three global modes obtained for the calibrated model with interfaces.

Non-linear static analyses

Pushover analyses were performed to evaluate the seismic performance of “Pietro Capuzi” school. A lateral distribution pattern proportional to the mass of the structure was adopted to apply horizontal loads. The analyses were performed along the X and Y global axes of the numerical model, corresponding respectively to the longitudinal and transversal directions of the structure, in both positive and negative directions. Eight control nodes, located at the top of the building and characterized by large displacements but small deformations, were used to plot the capacity curves (Figure 9). The damage was associated to high values of principal

tensile strains, which indicated cracking [20]. The principal crack width strain (mm) was also taken into consideration to evaluate the damage state.

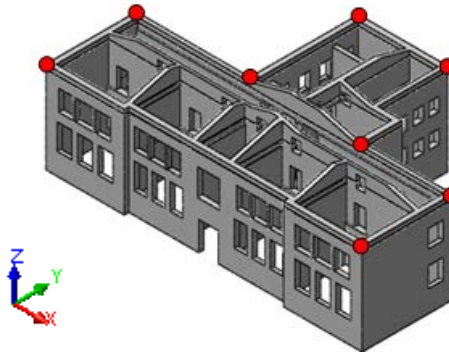


Figure 9 – Control nodes for the pushover analyses.

Figure 10 shows the capacity curves obtained for the pushover analyses performed in +X, -X, +Y and -Y. As for -Y direction, it is important to note that the curve exhibited a long, almost flat, plateau, consistent with a shear or rocking failure mode of masonry elements, without material crushing or geometrical non-linear effects. This made the last point of the curve physically less meaningful, but only related to convergence issues. For this reason, the ultimate capacity of the structure was estimated by taking into consideration the value of ultimate displacement indicated by the Italian Circolare [14] for the verification with respect to ultimate (life safety) limit state for existing masonry buildings. Consequently, the capacity curve presented in Figure 10 was interrupted when the values of ultimate displacement was overcome. Note that the same verification was performed for the analyses carried out in +Y, -X and -X direction, but in these cases the ultimate displacement was not exceeded.

As shown in Figure 10, the building exhibits a higher stiffness and lateral load-carrying capacity in X direction than in Y direction. The maximum values of lateral load factor obtained from the analyses in -X and -Y are, indeed, equal to 0.55 g and 0.44g, respectively. It is observed that the structure is characterized by a lower capacity in +Y and +X directions with respect to -Y and -X directions, respectively. However, this can be attributed to the fact that the curves in +Y and +X directions were plotted for smaller ultimate displacements, as the curves almost coincide in the positive and negative directions.

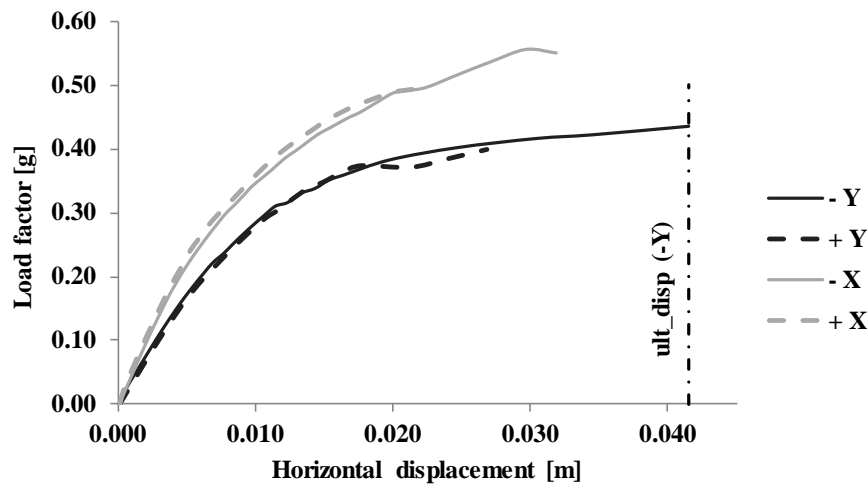


Figure 10 - Capacity curves obtained from the analyses performed in +X, -X, +Y and -Y directions.

Figure 11 presents the damage pattern of masonry walls in terms of principal crack width strain for the pushover analyses in -Y and -X direction, for which the maximum values of horizontal load factor were obtained. A global failure mechanism is observed in both the directions of analysis, with masonry walls mostly affected by diagonal shear cracks. Flexural cracks are also observed, especially in the spandrels. Furthermore, it is observed that, despite the higher values of horizontal load applied in -X direction, the damage is more severe in the analysis in -Y direction, as shown by damage distribution and higher values of principal crack width strain.

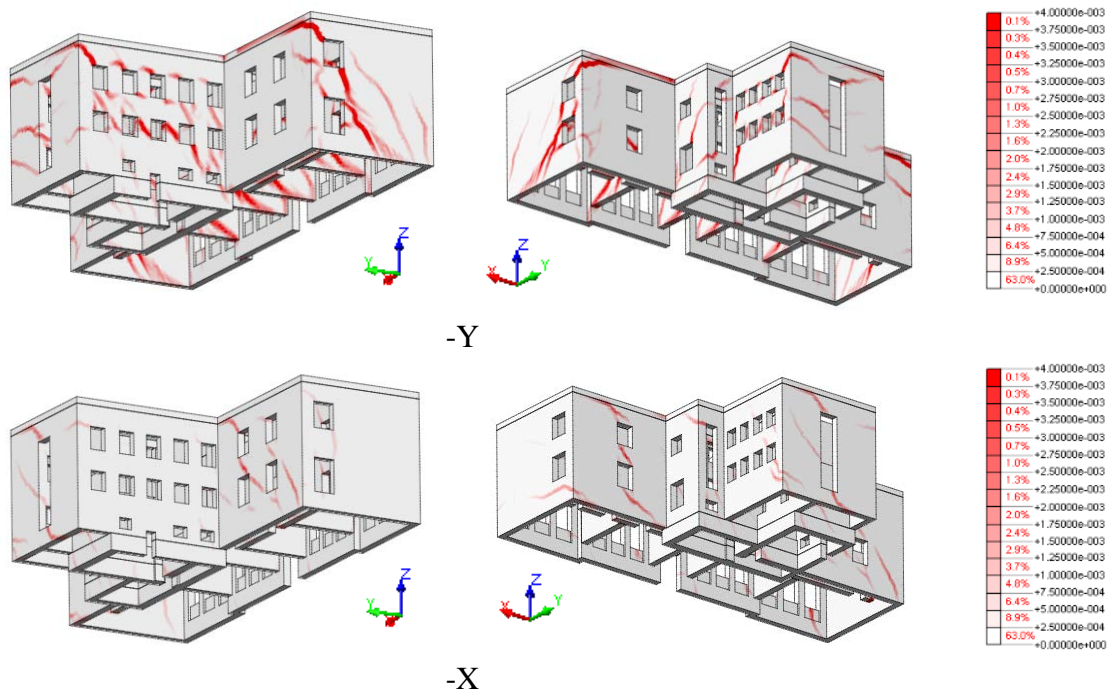


Figure 11 - Principal crack width strain for pushover analyses in -Y and -X directions.

Finally, even if pushover analyses are not able to fully represent the simultaneous action of an earthquake in different directions, a comparison can be attempted between the results obtained from pushover analyses and the seismic performance exhibited by the structure during the Amatrice-Visso-Norcia seismic sequence. According to the damage survey, “Pietro Capuzi” school was characterized by an in-plane response with severe cracks, mainly due to shear, in both internal and external walls. Consequently, the damage mechanisms obtained from the numerical model are in good agreement with the crack pattern observed in the structure. As for the out-of-plane mechanism activated in one corner of the building after the seismic event of October 26th, 2016, it is to note that such a mechanism was difficult to obtain from the numerical model due to the presence of the concrete bond-beam at the top of the structure. Furthermore, it possibly resulted from a local disintegration of the material, which it is not considered in the FE model.

Regarding the seismic capacity, the maximum values of horizontal load that the structure was found to bear from pushover analysis are consistent with the values of PGA recorded during the seismic event of October 26th. In particular, this is clear in -Y direction where the maximum lateral load applied in the FE model was equal to 0.44 g and it was associated to a severe damage state compatible with the one observed in the building in the walls oriented in Y direction (Figure 12a). As for X direction, according to the results of pushover analyses, the building exhibited a higher seismic capacity, with a maximum lateral load equal to 0.55g, but

less severe damage. From Figure 12b, the PGA recorded in X direction was much lower and, indeed, the damage was less severe in X direction than in Y direction. In conclusion, the numerical model was found to accurately predict the seismic performance of the building in terms of horizontal maximum load and failure mechanisms.

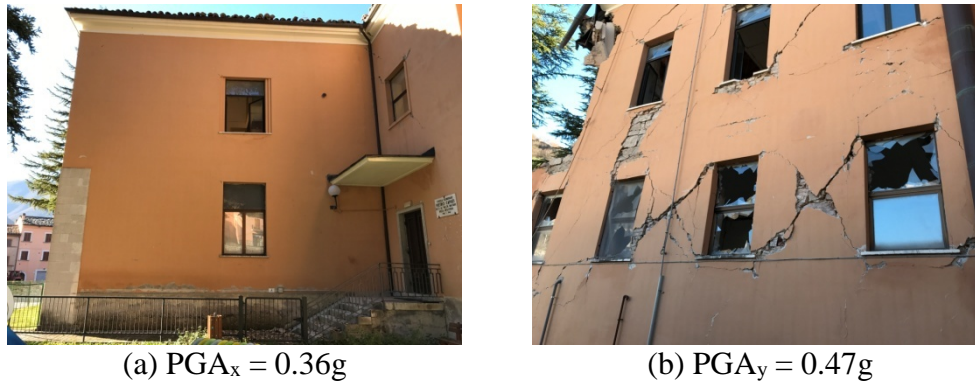


Figure 12 - Comparison between the damage distribution obtained in the walls oriented in X (a) and Y (b) after the Mw 5.9 earthquake of October 26th.

Conclusions

This paper presented the seismic assessment of “Pietro Capuzi” school in Visso, Marche, which was severely damaged by the 2016-2017 Amatrice-Visso-Norcia seismic sequence. First, the progressive damage suffered by the building after the subsequent earthquakes was analyzed in terms of development of the crack pattern and inter-storey drift. Secondly, on the basis of the documentation available in terms of geometry, construction details and material properties, a 3D macro FE model of the entire structure was prepared and calibrated according to the results of past identification dynamic tests. Finally, pushover analyses were performed in order to assess the seismic performance of the building. In both X and Y directions, the school exhibited a global in-plane response with cracks occurring in masonry piers and spandrels. A higher seismic capacity was observed in X direction with respect to Y direction. From the comparison between the numerical damage and the observed crack pattern, it was possible to conclude that the numerical model was capable to realistically simulate the seismic performance exhibited by the building during the past seismic events in terms of capacity to bear horizontal loads as well as failure mechanisms.

Acknowledgements

The authors are grateful to the Seismic Observatory of Structures, RELUIS and Ing. Serena Cattari of the University of Genova for the material made available for the development of the present work. Support was also provided by the SAHC Consortium and the University of Minho in Guimarães (Portugal).

References

- [1] DPC Osservatorio Sismico delle Strutture OSS Download Service <http://www.mot1.it/ossdownload>, accessed 12/04/2017.
- [2] S.G.M. S.r.l., *Rilievo geometrico e strutturale - Scuola elementare "Pietro Capuzi"- Visso (MC)*, 2010 (in Italian).
- [3] S.G.M. S.r.l., *Rilievo proprietà meccaniche e rilievo geometrico costruttivo - Scuola elementare e materna "Pietro Capuzi" - Visso (MC)*, Relazione P4134-20/11, 2011 (in Italian).
- [4] S.G.M. Engineering, *Modellazione e analisi numerica - Scuola elementare e materna "P. Capuzi" - Visso (MC)*, 2011 (in Italian).

- [5] CESI S.p.A., *Caratteristiche del sistema di monitoraggio sismico - Scuola materna "P. Capuzi", Visso (MC)*, Rapporto di installazione, 2010 (in Italian).
- [6] Cattari, S., Degli Abbati, S., Ottonelli, D., Sivori D., Spacone, E., Camata, G., Marano, C., Da Porto, F., Lorenzoni, F., Penna, A., Graziotti, F., Ceravolo, R., Matta, E., Miraglia, G., Spina, D. and Fiorini, N., *Task 4.1 Workgroup_Report di sintesi sulle attività svolte sugli edifici in muratura monitorati dall'Osservatorio Sismico delle Strutture, Linea Strutture in Muratura*, ReLUIS report, Rete dei Laboratori Universitari di Ingegneria Sismica, 2017 (in Italian).
- [7] Cattari, S. and Sivori, D. (2018) Lessons from the 2016 Central Italy Earthquakes: the seismic behavior of two masonry schools in Visso and Caldarola (Marche region), *Bulletin of Earthquake Engineering* (to be submitted).
- [8] DPC, *Photographic survey performed during Amatrice-Norcia-Visso seismic sequence*, 2016.
- [9] D.M. 14/1/2008, *Norme tecniche per le costruzioni (NTC 2008)*, Ministero delle Infrastrutture e dei Trasporti, 2008 (in Italian).
- [10] DPC, Oss - *Osservatorio Sismico delle Strutture*, <http://www.protezionecivile.gov.it/jcms/it/osservatorio.wp>, accessed 12/04/2017.
- [11] TNO DIANA BV, *DIANA Finite Element Analysis User's Manual Release 9.6*, Delft, The Netherlands, 2014.
- [12] Lourenço, P. B., *Computational strategies for masonry structures*, PhD Thesis, Delft University of Technology, Netherlands, 1996.
- [13] Silva, L. C., Mendes, N., Lourenço, P. B. and Ingham, J. (2018) Seismic Structural Assessment of the Christchurch Catholic Basilica, New Zealand, *Structures* **15**, 115-130.
- [14] Circolare 2 febbraio 2009, n. 617, *Istruzioni per l'applicazione delle "Nuove norme tecniche per le costruzioni" di cui al decreto ministeriale 14 gennaio 2008*, 2009 (in Italian).
- [15] Angelillo, M., Lourenço, P. B. and Milani, G. (2014) *Masonry behaviour and modelling*, in Angelillo M (eds) *Mechanics of Masonry Structures*. CISM International Centre for Mechanical Sciences, vol 551. Springer, Vienna
- [16] Lourenço, P. B., *A user/programmer's guide for the micro-modelling of masonry structures*. Relatório nº 03.21.1.31.35, Universidade Técnica de Delft, Delft, Países Baixos and Universidade do Minho, Guimarães, 1996.
- [17] SVS, *ARTEMIS Modal, 5.0.0.1*, User's manual, Denmark, 2016.
- [18] Ferrero, C., *2016 Central Italy Earthquake: Seismic Assessment of "Pietro Capuzi" School in Visso (Marche)*, SAHC Master Thesis, University of Minho, Guimarães, Portugal, 2017.
- [19] Pastor, M., Binda, M. and Harcarik, T., *Modal Assurance Criterion* (2012) *Procedia Engineering* **48**, 543-548.
- [20] Lourenço, P. B., Trujillo, A., Mendes, N. and Ramos, J. L., *Seismic performance of the St. George of Latins church: Lessons learned from studying masonry ruins* (2012) *Engineering Structure* **40**, 501-518.

Rock mounted iconic lighthouses under extreme wave impacts: Limit Analysis and Discrete Element Method

***†Athanasios Pappas¹, Dina D'Ayala¹, Alessandro Antonini², and Alison Raby²**

¹Civil, Environmental and Geomatic Engineering – University College London, UK.

²School of Marine Science and Engineering, Plymouth University, UK.

*Presenting author: a.pappas@ucl.ac.uk

†Corresponding author: a.pappas@ucl.ac.uk

Abstract

This paper deals with the resilience of rock mounted lighthouses impacted by extreme waves. The investigated lighthouse of Wolf Rock is built with large and meticulously dovetailed granite blocks. For this structural typology, uplift of blocks and separation of joints is possible under intense wave impacts. Sliding between subsequent courses of stones is limited by the existence of vertical keys between the blocks. The lateral forces that can trigger uplift and rocking, a highly nonlinear behaviour, are calculated with the use of Limit Analysis method through an iterative procedure written in Python programming language. Three different overturning mechanisms, which correspond to different hypotheses about the working section of the lighthouse, are considered for the uplift. The most conservative limits, i.e. with only a portion of the circular section taking the load, are representative for small and medium waves which can cause a small uplift and very local opening of a joint. The mechanism that considers the whole section, i.e. less conservative, is more representative for very strong wave impacts. The sliding failure mechanism is also considered. The Limit Analysis thresholds are validated with Discrete Element (DE) time-history analyses using the commercial software 3DEC. Intense uplift and rocking take place, and are especially present when the less conservative Limit Analysis thresholds are exceeded. The study also revealed the beneficial role of the vertical keys, without which the lighthouse would fail due to intense sliding before the overturning mechanisms are activated.

Keywords: Limit analysis, Discrete Element Method, historic lighthouse, rocking, wave impact

Introduction

Lighthouses on hostile and exposed rocks around the British Isles and Ireland have been resisting the impacts of extreme waves for over a century. However, the history of these landmarks of engineering has not been smooth. Lighthouse engineering has evolved after repeating collapses of under-designed structures and the subsequent upgraded design. The majority of the surviving rock-mounted lighthouses in the area are built based on an ingenious design: a tapered masonry structure with large-scale interconnected blocks, proposed by John Smeaton who designed a lighthouse of Eddystone in the mid-18th century. Prior to this design, three other lighthouses on the same rock had failed. The first presented unrepairable damages in its first winter, the second collapsed after a winter storm in its fourth year, and the third caught fire nearly 50 years after its construction. Though more resilient than their predecessors, plenty of the existing rock-mounted lighthouses have manifested uplift and motion after intense wave impacts. The

importance of the lighthouse network to the safety of navigation, in combination with the heritage value of these iconic lighthouses, provided the motivation for this structural analysis. The uplift and rocking behaviour of slender structures was first introduced by Housner [1]. His work evidenced that the structural behaviour of bodies capable of uplifting differs significantly to the one of continuous structures. Later studies verified the complexity of the rocking behaviour [2]–[4]. However, all of these studies are focused on base excitation and not on lateral wave impacts. Although plenty of research has been devoted on the estimation of wave impacts [5]–[7], little has been done regarding wave impacted rock lighthouses [8].

This paper presents the application of the limit analysis method on a masonry lighthouse. At first, the limits of wave impact intensity for each impact height are calculated. The activated mechanisms, i.e. overturning or sliding, are calculated for Wolf Rock lighthouse. Then, the limit analysis results are confronted with the time-history numerical results of a Discrete Element Method (DEM) model for wave impacts calculated for the specific location and structure.

Wolf Rock lighthouse

The rock mounted lighthouse of Wolf Rock is located around 13 km south-west from the point of Land's End, in the menacing coastline of Cornwall, UK. The construction of the current lighthouse which survives till today was initiated in 1862 and finished in 1869. Prior to this structure, the Wolf Rock had seen a series of four beacons failing after violent storms and plenty of designs for new lighthouses being abandoned as insufficient to resist the colossal wave impacts [9].

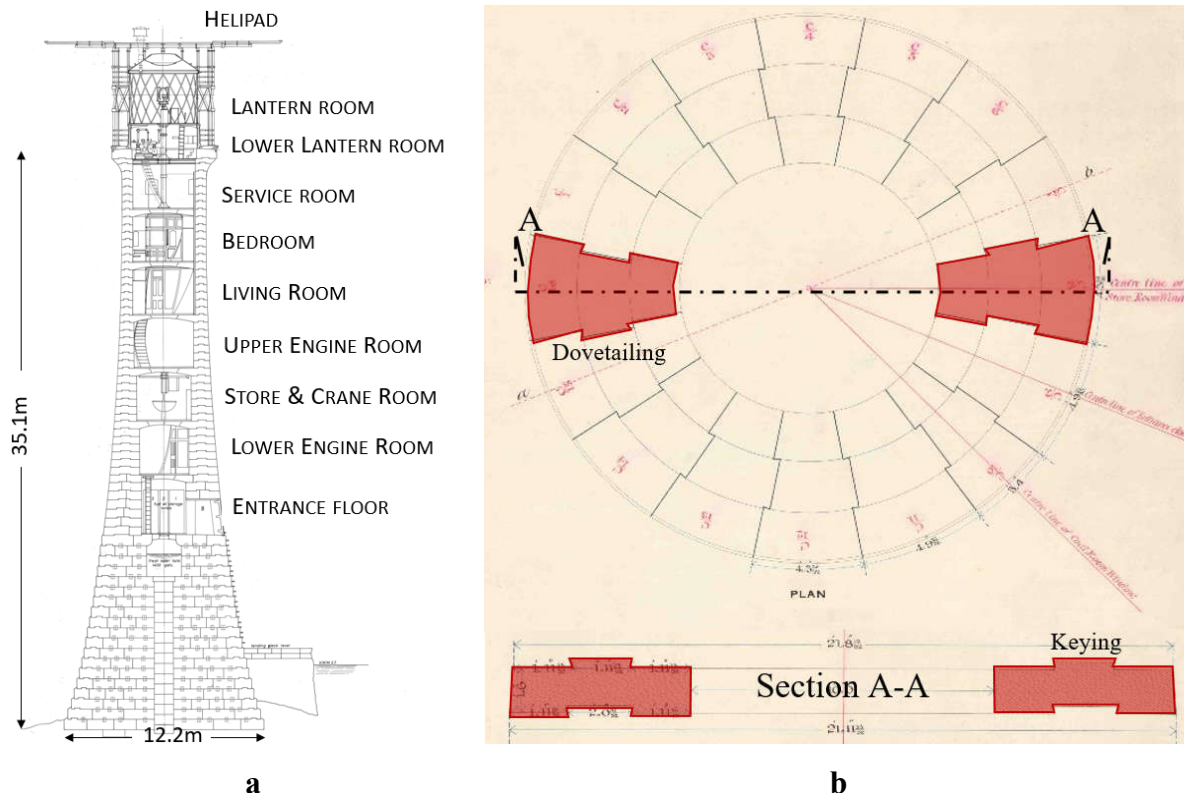


Figure 1. Wolf Rock lighthouse: (a) original section drawing and (b) details of dovetailing and keying for a course of stones.

The lighthouse consists of a granite masonry body that is 35.1 m high and has a 6.1 m high lantern at the top. The diameter of the granite body is 12.2 m at the base and gradually decreases

to 5.2 m near the top. The masonry structure consists of 6 vaulted levels, plus the lantern structure on the top. The wall thickness varies between 2.37 m at the entrance level and 0.69 m at the upper level. Finally, a steel frame helideck was constructed on the top of the masonry body in the early 1970s for facilitating the movement of personnel and supplies.

The horizontal and vertical interlocking of the granite blocks through dovetails in the vertical courses and keys in the horizontal courses is shown in Figure 1b. In this structural typology, apart from uplift, no other high relative movement between blocks is possible without fracture of the dovetailed connections. Sliding, along the horizontal joint between two successive courses of stones, is also blocked by the vertical key connections. The original drawing suggest that the height of the key is around 7.6 mm high. Although most lighthouses of this typology have the same keying technique, Wolf Rock bears additional vertical connections on the lower third of its height. This is an evidence that the designer engineer, James Nicholas Douglas, was indeed concerned about the severity of the wave impacts which could cause uplift or sliding.

Extreme wave loading

Until now, a theoretical description of the loading condition induced by the breaking waves on emerged cylinders is not available. Thus, due to this lack and to the impellent necessity to perform a survivability assessment of these ancient structures, the method of Wienke and Oumeraci [5] is applied as final tool to describe the impulsive wave load. The method consists of five main steps: *i)* The extreme offshore wave climate is identified by means of statistical extreme analysis. The Generalised Pareto Distribution, in combination of Peak Over Threshold method, is used through Bayesian inference [10] in order to identify the extreme significant wave heights (H_s) Figure 3, while the relationship between the significant wave height and peak period (T_p) is described by means of exponential equation, Figure 3 [11], [12]. The final result is a set of return periods, significant wave heights and peak periods describing the offshore wave climate.

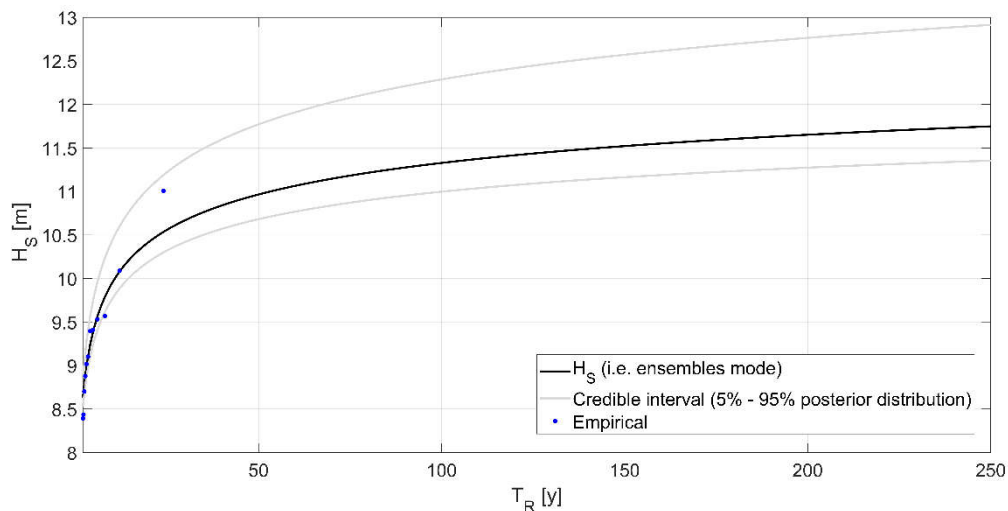


Figure 2. Wolf Rock Lighthouse offshore significant wave height (H_s) vs return period

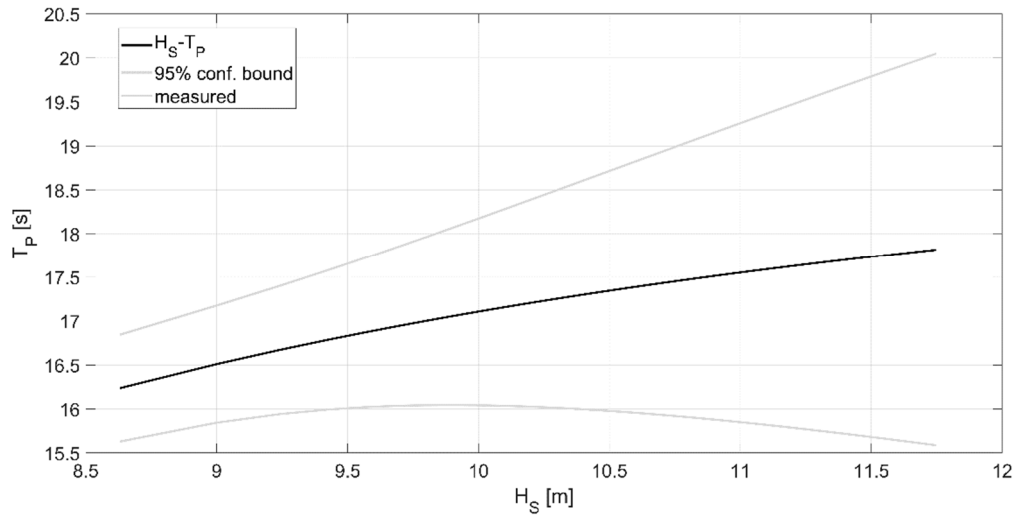


Figure 3. Wolf Rock Lighthouse offshore peak periods (T_P) vs significant wave height (H_s)

ii) In the surrounding area of the Wolf Rock, the bottom topography can be represented by very steep bottom that never reach shallower water condition. The approach of Goda [13] was applied to calculate the wave transformation from offshore to the lighthouse site identifying the local significant wave height, ($H_{s,L}$). iii) The effects of the restricted depth-to-height ratio and of breaking wave on the maximum wave height are considered by means of Battjes and Groenendijk's method [14], therefore, the design breaking wave height is assumed to be $H_{0.1\%}$. iv) Finally, the crest elevation with respect to the still water level is calculated according to Hansen's method [15]. v) Wienke and Oumeraci's method [5] is, then, applied considering the variation of the average radius of the lighthouse along to the calculated impact area. Load distribution is kept constant both in horizontal and vertical direction, while frontal area affected by the load distribution is considered included between $\pm 30^\circ$ from the wave dominant direction. The process is summarised in Figure 4.

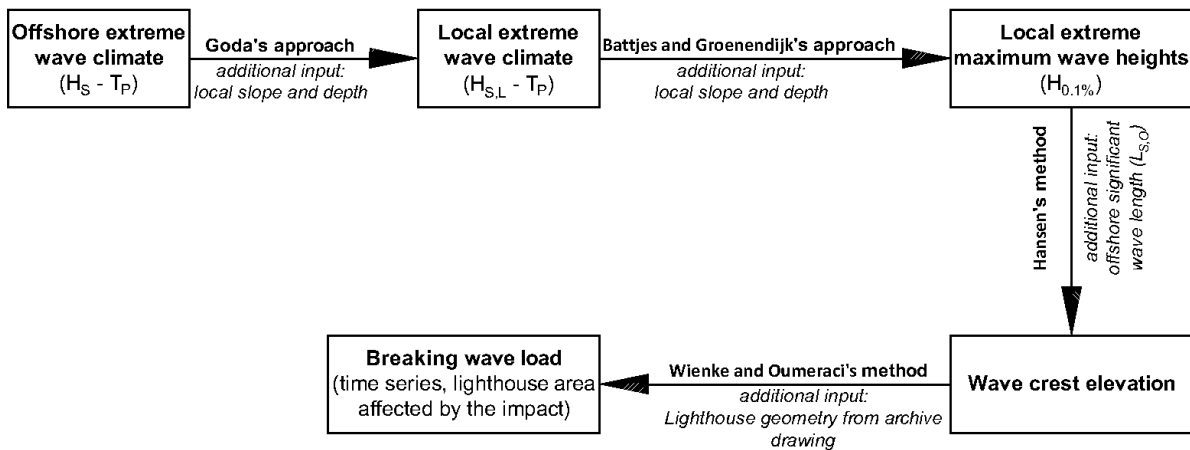


Figure 4. Summary of the process leading to the breaking wave load description.

For a wave impact with return period equal to 50 years, the total impact duration is equal to 0.075 s, and the maximum impact force, at $t = 0$, is 43031 kN. The force is applied between the 21st and the 35th course on a frontal section of 60° with uniform distribution. Thus, the load is applied on 15 courses and the force resultant is at a height of 16.37 m from the base of the structure and 12.1 m from the sea level. The time-history of the total force that this impulsive wave applies on the structure is presented in Figure 5.

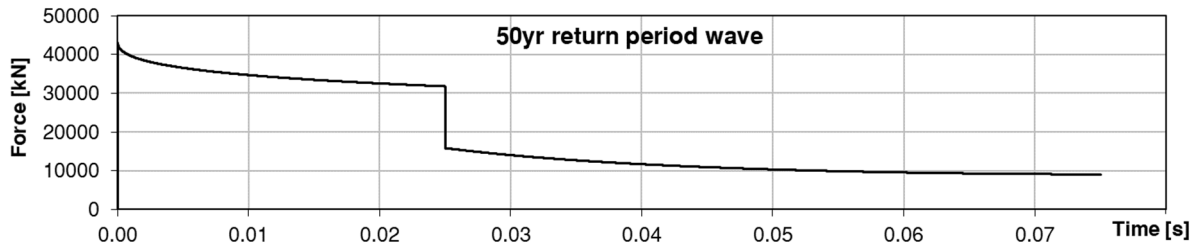


Figure 5. Time-history of the impulsive wave total force applied on the structure

Limit analysis

The limit analysis method calculates the magnitude of lateral force that is necessary for triggering a failure mechanism such as overturning or sliding (Figure 6). For overturning, the equilibrium of moments around a rotation hinge is calculated between the stabilisation forces, i.e. self-weight, and the external forces. For sliding, the equilibrium of horizontal forces is calculated by comparing the stabilisation forces, i.e. friction in horizontal joints, and the external forces. Regarding the overturning, three different mechanisms were considered. The first takes into account the whole section of the lighthouse (Figure 6a), the second considers only the front half section (Figure 6b), and the last mechanism considers only a frontal section of 60° (Figure 6c) which coincides with the impact section of the impacting wave [5]. Although the last two mechanisms (180° and 60°) are not realistic since the lighthouse is not fractured and therefore behaves as a continuous body, their calculation is useful for estimating the magnitude of external force that can cause a partial uplift. It has to be stated that the activation of an overturning mechanism is reversible. This means that for small duration impacts there can be some uplift and rocking but this does not necessarily mean overturning and collapse [1]–[3]. Note also that the existence of vertical keys for this lighthouse prevents any large sliding and therefore collapse due to sliding. Nevertheless, finite sliding due to small gaps between the vertical keys of the joints is still possible.

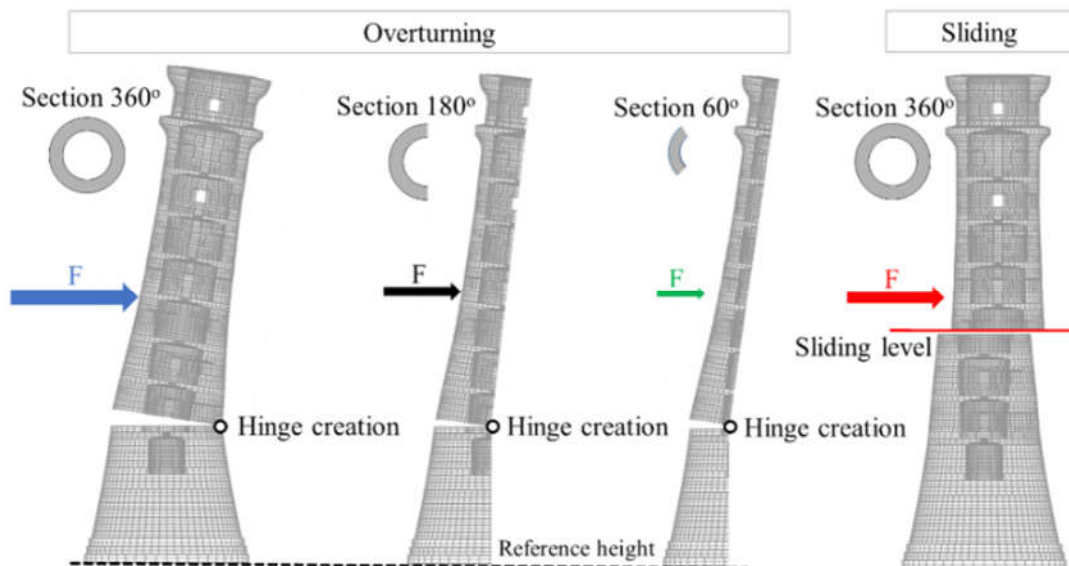


Figure 6. Failure mechanism for limit analysis: (a) overturning for 360° section, (b) overturning for 180° section, (c) overturning for 60° section, (d) sliding for 360° section

A drawback of the limit analysis is that the magnitude of activation force that is necessary for triggering the overturning mechanism depends on the selection of the hinge. Therefore, many hinge positions have to be tried in order to find the mechanism with the lowest activation force. Similarly, for the sliding mechanism, different joint levels have to be considered. To perform these calculations, an iterative procedure was written in Python 3.6 programming language. The self-weight of each course of stones is calculated based on the detailed geometrical data obtained during the archival research. All possible positions of horizontal activation forces are considered. Subsequently, all possible hinge or sliding levels are regarded for each external force scenario. The results for each height position of external force and the necessary magnitude for activation of the respective failure mechanism are presented in Figure 7.

The vertical axis in Figure 7 corresponds to the impact height and the horizontal axis shows the impact force that is necessary for the activation of each failure mechanism. The limit analysis curves for overturning (continuous blue, dashed black and green dash-dot green) and the sliding (dot red) are presented. The activation force decreases for increasing impact heights, which illustrates the importance of the impact height to the structural stability. The marker at force equal to 43031 kN and impact force equal to 16.37 m represents the resultant force of the 50 years return period wave which stands much higher than all limit analysis curves for overturning. This suggests that intense uplift and rocking is expected for this wave impact. In the same graph, the sliding forces, due to the wave impact, that act on each course are plotted as continuous red line. If this continuous line crosses the red dot line that represents the sliding limit, activation of sliding mechanism is possible. However, if vertical keying is present, such a mechanism cannot be activated unless the joint opens more than the height of the vertical keys or there is a rupture of the keys. The results for the 50 years wave shown in Figure 7 suggest that sliding would have been detrimental without the presence of keying.

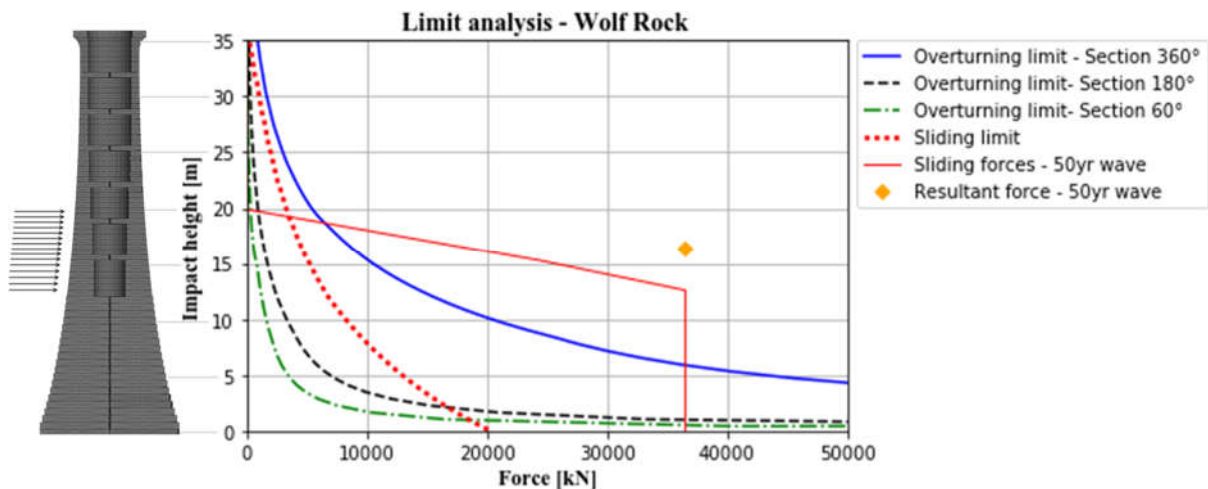


Figure 7. Limit analysis curves for overturning (continuous, dashed and dash-dot lines) and sliding (dotted line)

Numerical analysis

The numerical modelling of the structure was based on the Discrete Element Method (DEM) introduced by Cundall [16]. The method regards the structure as an assembly of discrete blocks and the solution is based on a time algorithm of sufficiently small time-steps. This solution scheme is described by Cundall (1976) as the DEM cycle. The cycle consists firstly on the calculation of the block motion in terms of velocity and acceleration, which are assumed to be constant within a given time-step. As the blocks move relatively, new contacts between blocks are detected and the relative contact velocities and forces are updated with the use of a force–

displacement law. Finally, the new forces for each block centroid are calculated and the new block motion is updated with the application of Newton's second law. The relatively simple theory behind the DEM circle makes the method particularly efficient for reproducing structural response of rigid bodies that is characterised by large displacements and separation between blocks [17].

The three-dimensional numerical model of the lighthouse is developed with the use of the DEM software 3DEC [18]. Each course of stones is modelled as an independent rigid block. The vertical keys were also modelled in detail, hence impeding large sliding unless significant uplift takes place. The Coulomb friction law is implemented for the joints between blocks with zero cohesion and an angle of friction equal to 30° . Moreover, the joint is given normal stiffness equal to $5.93 \cdot 10^{10}$ Pa/m and shear stiffness equal to $4.45 \cdot 10^{10}$ Pa/m. The specific weight of the masonry blocks is taken equal to 2463 kg/m^3 , which corresponds to granite similar with the one used for Fastnet lighthouse [19]. Additional mass is added to the top course in order to account for the mass of the lantern and helideck that were not introduced to the model. Based on previous experimental work [19], mass proportional Rayleigh damping is adopted ($\alpha = 1.57$, $\beta = 0$). The time-step for each DEM circle is equal to $3.41 \cdot 10^{-5}$.

Limit analysis vs DEM results

The DEM model is tested for various impact intensities by scaling the original wave time-history. The maximum intensity of the scaled waves is presented herein normalised for the force calculated by the limit analysis for the 360° section. For instance, the original wave calculated for 50 years return period corresponds to a normalised force of 4.94. A total of 26 analyses are carried out for normalised wave forces varying between 0.01 and 4.94.

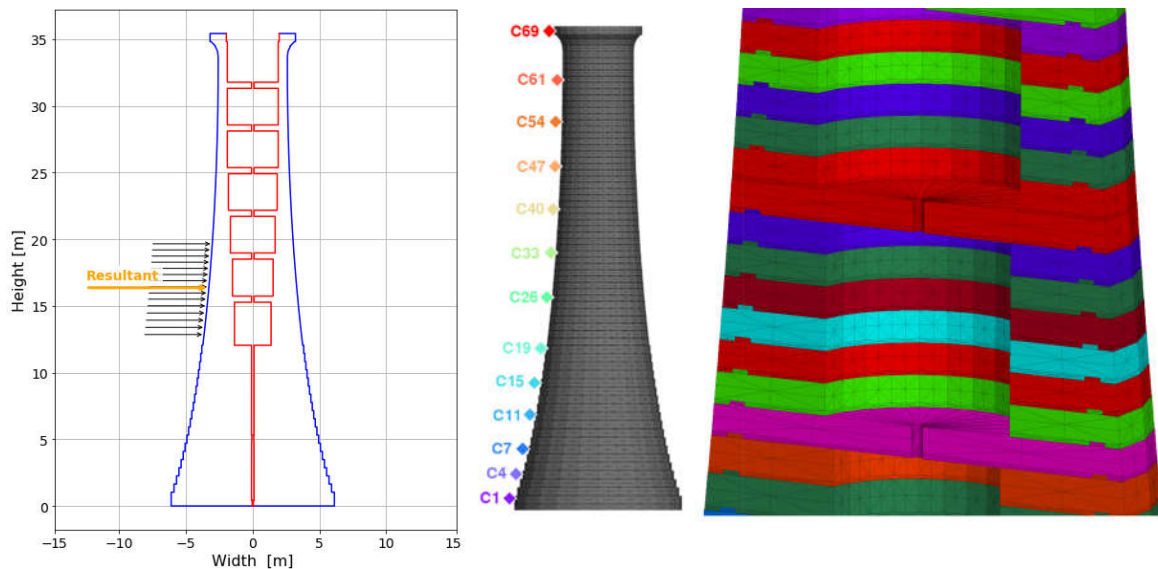


Figure 8. (a) Wave impact on Wolf Rock; (b) DEM model and control points; (c) DEM model section

The qualitative results of the tests suggest that the DEM model is able to reproduce a realistic structural response for the wave impact. Although the courses are modelled as rigid blocks, the structure is deformable due to the joint interfaces which behave as unidirectional springs. The stiffness of the joints was calculated in order that the model behaves with an equivalent modulus of elasticity equal to 30 GPa in the vertical direction [19].

The curves presented in Figure 9 and Figure 10 testify that the correlation between the impact forces and the structural response in terms of displacements is not linear. Both the vertical and

the horizontal displacement curves begin as linear for small impact intensities but they become parabolic for stronger impacts. Note that the transition from linear to parabolic shape is gradual and thus cannot be attributed to the exceedance of a strict threshold. The transition phase is clearly presented Figure 10 for normalised forces ranging between 0.01 and 0.5. The finding that there is no specific threshold over which the nonlinearity is triggered, supports the approach of using multiple assumptions (60°, 180°, 360° section), for the limit analysis as guides for the structural assessment. The trend of the vertical and horizontal displacement curves is linear for forces near the most conservative limit analysis assumption (60° section), and becomes nonlinear much earlier than the less conservative threshold (360° section). This unclear boundary between linearity and nonlinearity can be explained by the dynamic nature of the impact force that has a duration of only 0.075 s. For impacts of much higher duration, the structural response would appear more discretised and would resemble the assumptions of the limit analysis.

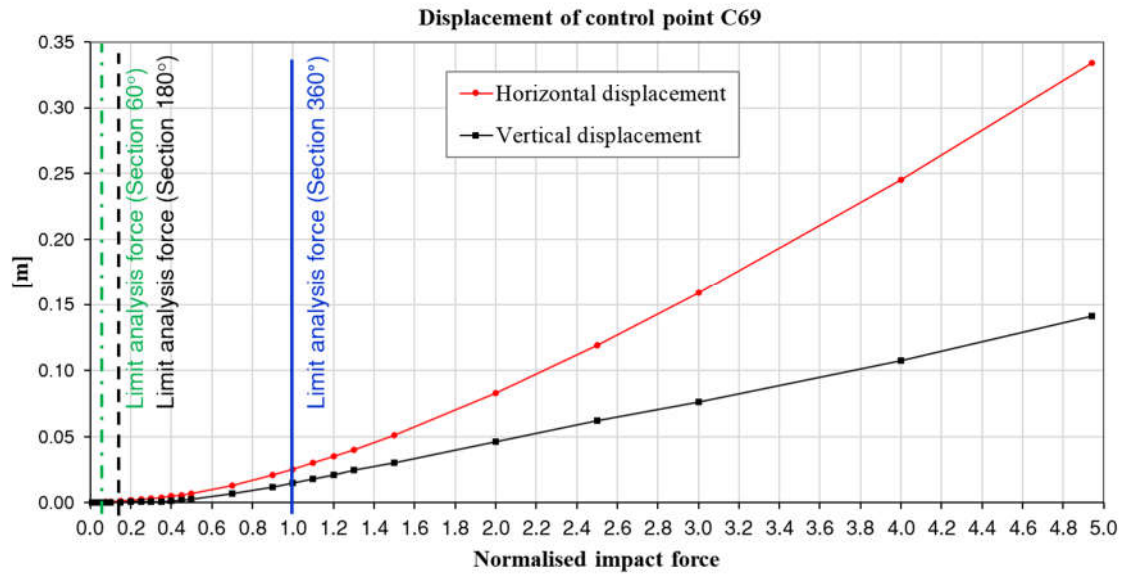


Figure 9. Maximum horizontal and vertical displacement from DEM analyses together with limit analysis threshold values; range of normalised impact forces 0.01 - 4.94

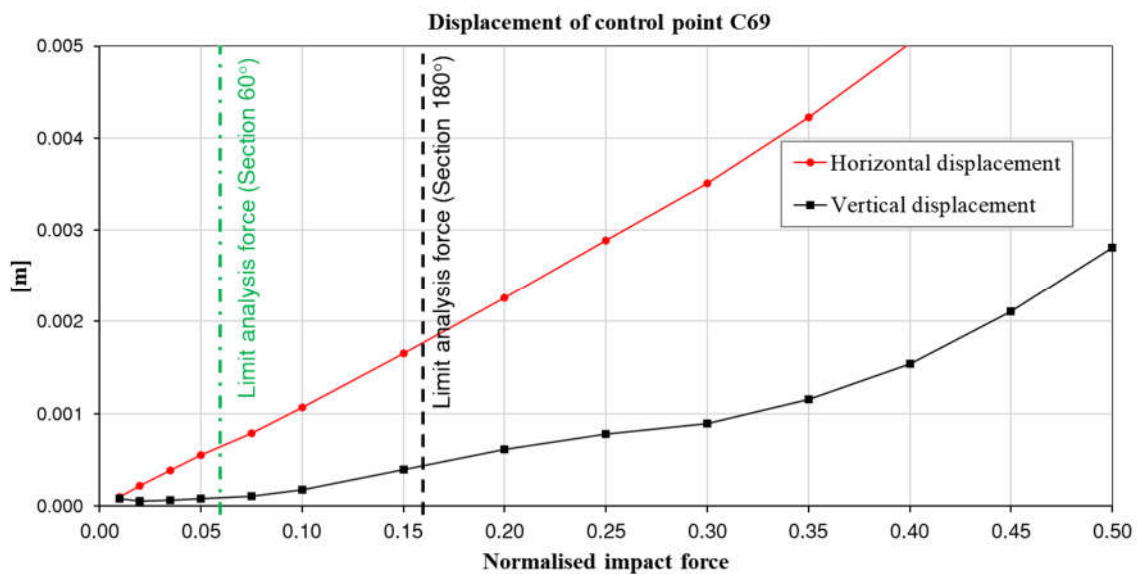


Figure 10. Maximum horizontal and vertical displacement from DEM analyses together with limit analysis threshold values; range of normalised impact forces 0.01 – 0.50

In order to explain the change from linear to parabolic trend for the maximum displacements, it is worth comparing the structural response for a strong impact, i.e. the 50 years return period wave, with the response for a weak generic wave of normalised peak force equal to 0.01. The structural response in terms of vertical and horizontal displacements of the DEM model for the estimated impact of the 50 years wave is shown in Figure 11 and Figure 12 respectively. This is a particularly intense impact with normalised force equal to 4.94. The results yield intense rocking and opening of the horizontal joints. The response time-histories reveal a time lag of the peak displacements between the upper and lower part of the lighthouse. The vertical displacements reveal that the joint opening is initiated between the control points C7 and C19 (Figure 11). The maximum uplift is found on the highest control point, i.e. C69, and is recorded around 0.28 s after a first peak found on the lower courses. Similarly, the maximum horizontal response is equal to 0.14 m and is found at C69 with a time lag of 0.2 s compared to a peak recorded on the lower courses (Figure 12).

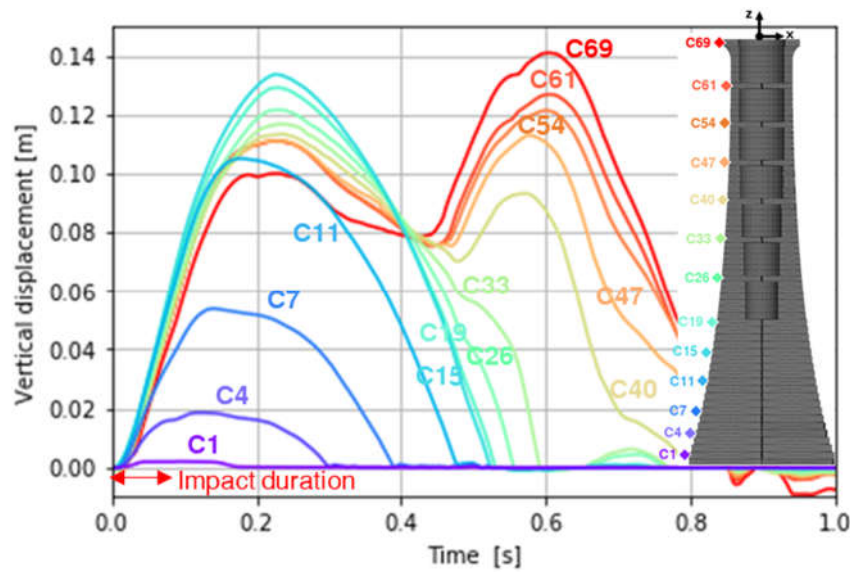


Figure 11. Structural response of the DEM model for the wave of 50 years return period recorded at the control points: vertical displacement

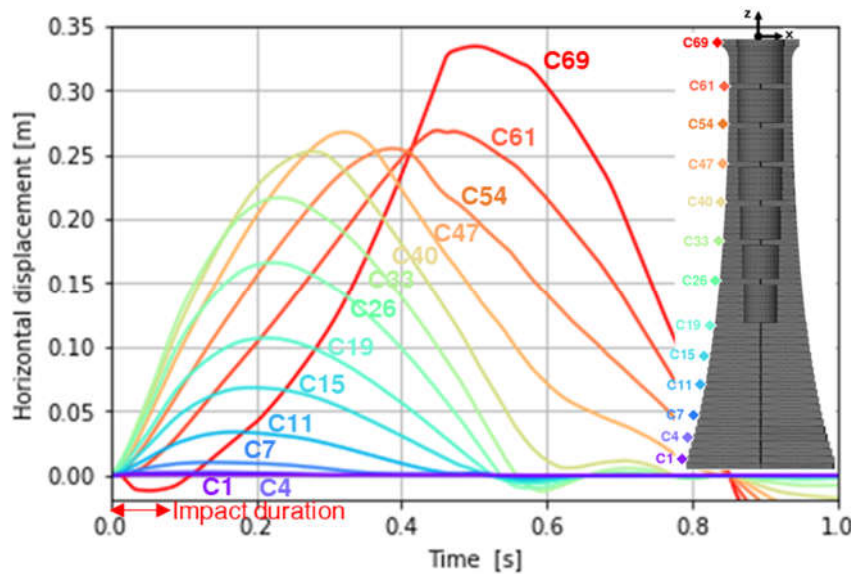


Figure 12. Structural response of the DEM model for the wave of 50 years return period recorded at the control points: horizontal displacement

The structural response of the weak impact is presented in Figure 13. This graph shows the horizontal displacement of the control points for a duration of 1.5 s, which includes the impact time (initial 0.075 s) and a damped post-impact free-vibration. It is worth noticing the intense phase difference of the upper versus the lower control points for the beginning of the motion. The higher frequencies (dominating in the lower courses) though are gradually damped out and all areas of the lighthouse pass to an in-phase vibration. Compared to the intense rocking caused by the 50 years wave, this behaviour resembles the dynamic oscillation of a quasi-linear elastic structure. Therefore, this qualitative comparison suggests that the structural response of the lighthouse passes from quasi-linear elastic oscillation to highly nonlinear rocking for increasing wave impact intensity.

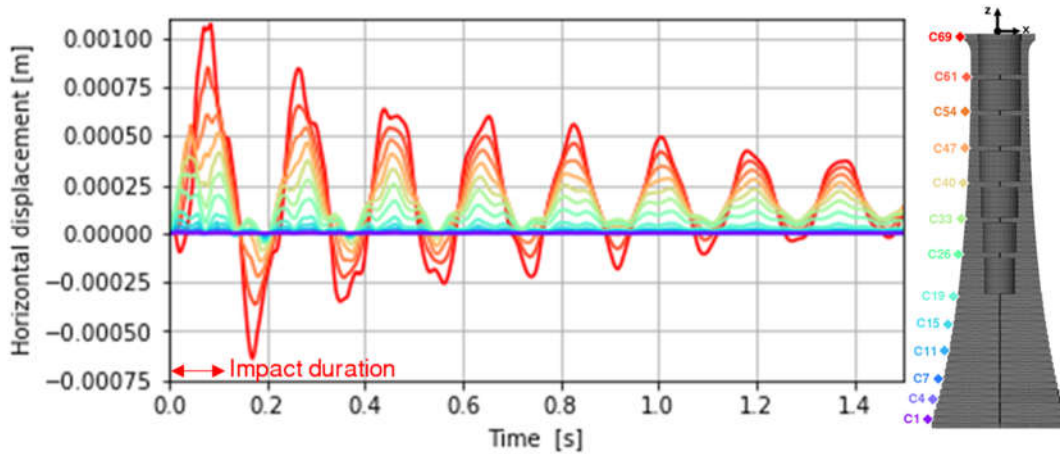


Figure 13. Horizontal displacement at control points for normalised wave impact force equal to 0.01

Structural assessment for the 50 years wave impact

According to the DEM results for the 50 years return period wave impact, the maximum horizontal displacement calculated on the top of the masonry body (control point C69) is equal to 0.33 m. For the same point, the maximum uplift is equal to 0.14 m, which is around twice the height of the vertical key. However, due to the elasticity of the structure, accounted by the stiffness of the joints, this uplift is not concentrated at a single joint, which would be catastrophic, but is a sum of multiple joint separations. The opening of joints is concentrated on two areas of the lighthouse (Figure 14a). The first involves the 4th till the 19th course, below the impact, where the joints open on the side of the impact. The second area is more localised and consists of the 31st till the 33rd course, just above the resultant of the impact forces, which open on the opposite side of the impact (Figure 14b). On the impact side, the maximum vertical separation between successive courses of stones is equal to 0.015 m and is recorded between the 9th and 10th course. This uplift is smaller than the height of the vertical keys, hence sliding is unlikely to happen in this area. On the opposite side, the maximum vertical separation between successive courses of stones is equal to 0.032 m and is recorded between the 32nd and 33rd course. Although sliding forces are not applied to this area, this uplift corresponds almost half of the key height. This suggests that the structural response is particularly intense.

According to limit analysis for the 360° section, joint opening on the impact side takes place at the joints below the impact forces. Once the overturning mechanism has been activated, accelerations due to inertial forces are applied to the rotating body, i.e. the courses above the open joints, which can cause new joint openings and overturning on the opposite direction. The same

Python code that was created for the limit analysis was enhanced in order to predict joint opening on both sides of the structure. The code predicts opening between the 13th and 20th course on the impact side and the 25th and 26th course on the opposite side. Although these openings are at the same regions as the ones predicted by the DEM analysis, the two methods do not perfectly align. This is expectable due to parameters such as damping, friction at the keys, and mainly the elasticity that are taken into account only in the DEM model. The elasticity of the structure allows limited deformation and therefore redistribution of the internal forces that influence the structural response by slightly shifting or expanding the joint opening areas.

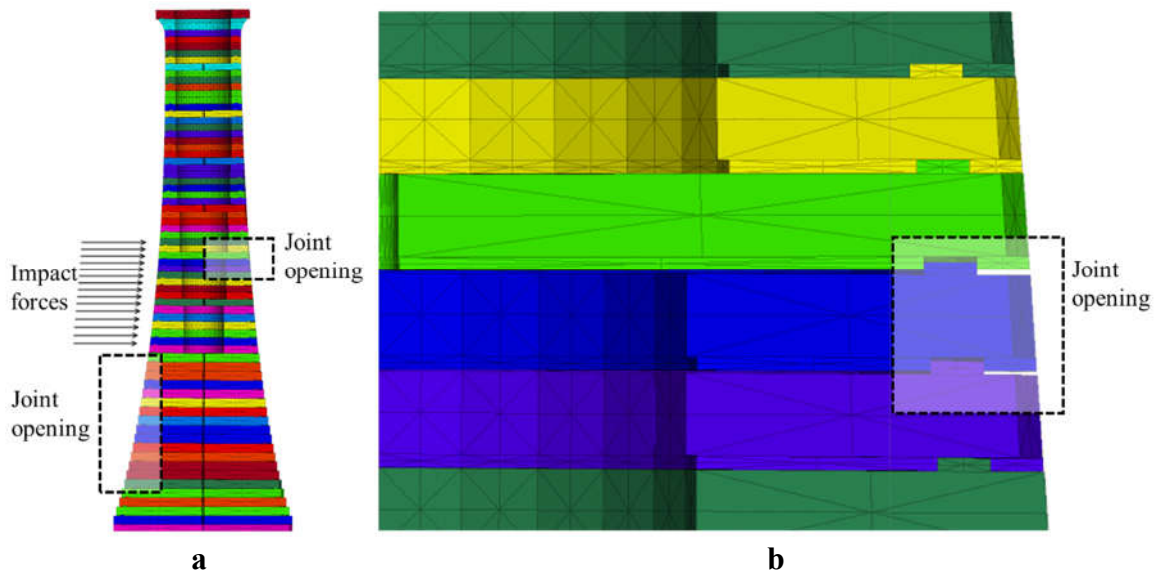


Figure 14. Joint opening at 0.075 s for the 50 years wave impact: (a) overall aspect; (b) detailed view between the 31st and 33rd joint

Conclusions

The coupling of limit analysis with DEM time-history analyses for wave impacts offers valuable knowledge about the structural behaviour of the Wolf Rock lighthouse. The main findings of this paper are summarised below:

- The iterative code developed for the limit analysis, offers important understanding about the structural behaviour of the lighthouse. The limit analysis results show that the combination of bigger diameter and greater weight near the bottom makes the lighthouse able to resist significantly bigger forces if the impact area is near the bottom. Moreover, the huge importance of the vertical keying to the stability of the lighthouse is revealed. Although an overturning mechanism can be reversible, meaning that exceeding of the threshold does not necessarily mean damage, the sliding mechanism is not reversible.
- The DEM analysis produces very satisfactory results and is able to reveal the crucial areas of the structure where opening of horizontal joints takes place. The same qualitative structural behaviour is shown with the iterative code for the limit analysis. Though in similar areas, the joint opening is not identical for the two methods. This is mainly because of the elasticity of the DEM model that changes the distribution of the internal forces.
- The structural response of the lighthouse resembles a quasi-linear elastic vibration for small impacts and becomes highly nonlinear with intense rocking and joint separation for stronger impacts. Moreover, the parametric analysis shows that the amplitude of the

maximum vertical and horizontal displacements follows a linear trend with for small impacts and gradually becomes parabolic for stronger impacts. This transition is not strictly correlated with any of the limit analysis thresholds. For the most conservative limit analysis assumption, i.e. overturning section of 60°, the trend is still linear and becomes parabolic before the exceedance of for the least conservative threshold, i.e. 360° section.

- The 50 years wave impact that was calculated based on the climatic wave conditions for Wolf Rock which has a steep bottom topography, causes intense rocking and opening of the horizontal joints. This wave causes joint opening both on the side of the impact, below the impact area, and on the opposite side, above the impact area. The maximum separation between successive joints is 0.032 m which is roughly half the height of the vertical keys. A maximum horizontal displacement equal to 0.33 m and uplift equal to 0.14 m are found for the control point on the top of the masonry body. Both the limit analysis and the DEM analysis confirm that this is a particularly intense wave. Wolf Rock lighthouse would not be able to survive such an impact without the presence of vertical keying.

Acknowledgments

The authors would like to thank the UK Engineering and Physical Sciences Research Council (EPSRC) for the financial support of the STORMLAMP project EP/N022947/1 and the Trinity House for the provision of detailed drawings and for physical access to the Wolf Rock lighthouse.

References

- [1] G. Housner, "The behavior of inverted pendulum structures during earthquakes," *Bull. Seismol. Soc. Am.*, vol. 53, no. 2, pp. 403–417, 1963.
- [2] C. Yim, A. Chopra, and J. Penzien, "Rocking response of rigid blocks to earthquakes," *Earthq. Eng. Struct. Dyn.*, vol. 8, pp. 565–587, 1980.
- [3] Y. Ishiyama, "Motions of rigid bodies and criteria for overturning by earthquake excitations," *Bull. New Zeal. Soc. Earthq. Engineering*, vol. 17, no. 1, pp. 24–37, 1984.
- [4] P. Spanos and A. Koh, "Rocking of rigid blocks due to harmonic shaking," *J. Eng. Mech.*, vol. 110, no. 11, pp. 1627–1642, 1984.
- [5] J. Wienke and H. Oumeraci, "Breaking wave impact force on a vertical and inclined slender pile - Theoretical and large-scale model investigations," *Coast. Eng.*, vol. 52, no. 5, pp. 435–462, 2005.
- [6] W. Allsop, M. Calabrese, and D. Vicinanza, "Wave Impact Loads - Pressures And Forces," no. February 2013, 2001.
- [7] K. Tanimoto, S. Takahashi, T. Kaneko, and K. Shiota, "Impulsive Breaking Wave Forces on an Inclined Pile Exerted by Random Waves," *20th Conf. Coast. Eng. Taipei, Taiwan*, vol. 20, pp. 2288–2302, 1986.
- [8] Q. Trinh, A. Raby, D. Banfi, M. Corrado, B. Chiaia, Y. Rafiq, and F. Cali, "Modelling the Eddystone Lighthouse response to wave loading," *Eng. Struct.*, vol. 125, pp. 566–578, 2016.
- [9] C. Nicholson, *Rock lighthouses of Britain*, 2nd ed. Wakefield, UK, 1983.
- [10] L. Cheng, A. AghaKouchak, E. Gilleland, and R. . Katz, "Non-stationary extreme value analysis in a changing climate," *Clim. Change*, vol. 127, no. 2, 2014.
- [11] U.S. Army Coastal Engineering Research Center, "Shore protection manual." Government Publishing Office, Washington, DC, 1977.
- [12] A. Antonini, R. Archetti, and A. Lamberti, "Wave simulation for the design of an innovative quay-wall: the case of Vlora's harbour," *Nat. Hazards Earth Syst. Sci. Discuss.*, vol. 17, no. 1, pp. 127–142, 2017.
- [13] Y. Goda, *Random seas and design of maritime structures*, 2nd ed. Singapore: World Scientific, 2000.
- [14] J. Battjes and H. Groenendijk, "Wave height distributions on shallow foreshores," *Coast. Eng.*, vol. 40, no. 3, 2000.
- [15] J. B. Hansen, "Periodic waves in the surf zone: Analysis of experimental data," *Coast. Eng.*, vol. 14, no.

- 1, 2005.
- [16] P. Cundall, "Explicit finite difference methods in geomechanics," in *Proceedings of the 2nd International Conference on Numerical Methods in Geomechanics*, 1976, pp. 132–150.
 - [17] J. V. Lemos, "Discrete Element Modeling of Masonry Structures," *Int. J. Archit. Herit.*, vol. 1, no. 2, pp. 190–213, May 2007.
 - [18] Itasca Inc., "3DEC 5.0: 3-Dimensional Distinct Element Code, Theory and Background." 2013.
 - [19] A. Pappas, D. D'Ayala, A. Antonini, J. Brownjohn, and A. Raby, "Numerical modelling of Fastnet lighthouse based on experimental dynamic identification," in *International Conference on Advances in Construction Materials and Systems ICACMS-2017*, 2017.

Simulations of Dynamic Concrete Fracture with Brazilian Splitting Test

***Josef Květoň¹ and Jan Eliáš¹**

¹Faculty of Civil Engineering, Brno University of Technology, Czech Republic

*Corresponding author: kveton.j@fce.vutbr.cz

Abstract

Concrete fracture is a phenomenon that material science deals with since its very beginning. With increase of computational capacity, linear models previously used for the description of material behavior in the vicinity of stress concentrators are being replaced with complex models using nonlinear material laws. Such models can be continuous, but for the better description of material behavior at the scale, where heterogeneities are recognized, discrete formulations are usually adopted.

The contribution presents simulations of concrete fracture in dynamic regime using rigid body spring network. The material is represented by interconnected rigid cells of convex polyhedral shape. In the dynamic regime, for relatively slow loading rates, use of implicit time integration scheme for solving system of equations is quite common. For faster loading rates, the explicit scheme is usually used, however, in the presented contribution, the implicit scheme is conveniently used as well. The time step length is then not restricted by the stability of the system, but only by the desired accuracy. Model is applied to dynamic simulations of Brazilian splitting discs experimentally investigated in literature.

Keywords: Concrete Fracture; Implicit Dynamics; Discrete Modeling; Meso-scale; Strain Rate

Introduction

It is well known that response of a tested specimen is highly dependent on loading rate. To obtain material properties such as tensile, flexural or compressive strength, the loading rate is usually of order 10^{-4} m/s. Using higher loading rate may affect the resulting values of obtained material properties. It is due to change in mechanism of failure. For a quasi-static loading, damaged volume of material is relatively small, in a case of tensile failure, the initial micro-cracking finally localizes into one relatively narrow highly damaged region – macro-crack. With increasing loading rate, the energy accumulated in a material volume is too high to be consumed by one crack only and multiple cracking occurs. The total damaged volume is then larger and damage is diffused in a wider region.

Attempts to describe this phenomenon of strain-rate dependency can be found in literature [1,7,8]. The relative increase (with respect to quasi-static rate) of material characteristic (strength, fracture energy, elastic modulus) is described by Dynamic Increase Factor (DIF). However, this phenomenon is associated not only with the increase in material resistance, but also with material inertia. General use of DIF is complicated, since it is not applicable to different geometry than the one it was obtained at.

It is therefore convenient to use the model that already incorporates the influence of inertia. Strain rate dependency of the material properties can then be investigated.

The contribution presents simulations of dynamic tests on Brazilian Splitting Discs. The simulations are calculated using discrete particle model. The time dependent response is obtained solving equations of motion using implicit time integration scheme according to

Newmark [6]. The performance of the mathematical model is compared with experimental results from literature [5].

Particle model

When modeling heterogeneous materials behavior, discrete approach is particularly convenient. It is because of its ability to represent internal material structure of concrete. The material volume is discretized into polyhedral particles that are assumed to be ideally rigid and their interaction is prescribed at their contact facets. The model particles then represent larger concrete/mortar aggregates with surrounding cement paste. The particle shape is obtained by Voronoi tessellation applied on a set of points randomly placed in the volume domain within prescribed minimum distance l_{\min} , which is related to maximum aggregate size. Each particle has 3 translational and 3 rotational degrees of freedom. Omitting direct representation of smaller mineral aggregates because of the computational reasons affects the model response, however, phenomena occurring below the resolution of the model are captured by nonlinear constitutive law.

The model is adopted according to [2] using further simplifications from [3]. In its basic version, only 4 parameters are used for material description; Elastic modulus E_0 and tangential to normal stiffness ratio α for elastic behavior and parameters of tensile strength f_t and fracture energy for tensile failure G_f in nonlinear regime. Other parameters such as compressive strength etc. are derived from these according to recommendations in [2]. It is important to note that all of these parameters are applied at meso-scale and their values differ from overall macro-scale material properties. The approximate relation between meso-scale elastic parameters and macro-scale Youngs modulus and Poissons ratio can be obtained from principle of virtual work [4]

$$E_0 = \frac{E}{1-2\nu} \quad \alpha = \frac{1-4\nu}{1+\nu} \quad (1)$$

As has been stated, constitutive law is applied at the contacts of discrete particles. The contact behavior is dependent on straining direction in both elastic and nonlinear regimes. After reaching the elastic limit, the damage model is applied to describe loss of integrity of material. The stress-strain relation depends on damage parameter D

$$\mathbf{s} = (1-D)E_0\boldsymbol{\alpha}\mathbf{e} \quad D \in \langle 0,1 \rangle \quad (2)$$

Here \mathbf{s} and \mathbf{e} are meso-scopic stress and strain vectors respectively with elements corresponding to normal and two tangential directions and $\boldsymbol{\alpha}$ is diagonal matrix with the first diagonal element 1 (for normal direction) and remaining diagonal elements α . Initially, parameter D is equal to zero. After reaching the elastic limit, the parameter increases up to 1, which indicates stress free crack. The evolution of damage is crucial part of the nonlinear behavior of the model. It takes into account combination of normal and tangential straining. Since its description is quite complex, the interested reader is referred to [3].

Dynamics

The solution of equations of motion $\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{F}$ is provided by an unconditionally stable time integration scheme according to Newmark [6]. \mathbf{M} , \mathbf{C} and \mathbf{K} stay for mass, damping and stiffness matrix respectively, \mathbf{F} is a loading vector. The damping matrix is not taken into account, since the system is damped in inelastic regime by the energy dissipation at contact facets. The solution is based on dynamic equilibrium at the end of each time step

(time $t+\Delta t$). The values of accelerations $\ddot{\mathbf{u}}$ and velocities $\dot{\mathbf{u}}$ at this time are estimated as numerical derivatives of displacements \mathbf{u}

$$\ddot{\mathbf{u}}_{t+\Delta t} = \frac{1}{\beta \Delta t^2} (\mathbf{u}_{t+\Delta t} - \mathbf{u}_t) - \frac{1}{\beta \Delta t} \dot{\mathbf{u}}_t - \left(\frac{1}{2\beta} - 1 \right) \ddot{\mathbf{u}}_t \quad (3)$$

$$\dot{\mathbf{u}}_{t+\Delta t} = \dot{\mathbf{u}}_t + \Delta t (1-\gamma) \ddot{\mathbf{u}}_t + \gamma \Delta t \ddot{\mathbf{u}}_{t+\Delta t} \quad (4)$$

where β and γ are parameters of the Newmark method, which, to keep the solution unconditionally stable, need to be kept within limits $\gamma, \beta \in \langle 0, 1 \rangle$ and $2\beta \geq \gamma \geq 0.5$. Omitting damping and substituting these into equation of motion, following system is obtained

$$\left(\mathbf{K} + \frac{1}{\beta \Delta t^2} \mathbf{M} \right) \mathbf{u}_{t+\Delta t} = \mathbf{F}_{t+\Delta t} + \mathbf{M} \left(\frac{1}{\beta \Delta t^2} \mathbf{u}_t + \frac{1}{\beta \Delta t} \dot{\mathbf{u}}_t + \left(\frac{1}{2\beta} - 1 \right) \ddot{\mathbf{u}}_t \right) \quad (5)$$

Each grain contributes to global mass matrix by 6x6 matrix. This particle mass matrix can be further divided into 4 smaller 3x3 submatrices. Upper left, relating forces and translations is diagonal containing value of particle mass at each diagonal element. Lower right, connecting rotations with moments consists of moments and products of inertia. Since the Voronoi cell center does not necessarily need to be particle center of gravity, static momentum needs to be taken into account to obtain balance, which is included in upper right and lower left submatrices.

No strain rate dependency at the level of constitutive equation was used for simulations presented in this contribution, it is assumed that the meso-level model correctly accounts for inertia even within the fracture process zone. Viscous effects due to presence of water are neglected.

Simulations

Simulations presented in this section were calculated using mathematical models corresponding to experimental series on Brazilian splitting tests reported in [5]. Simplified models of Brazilian discs were supported (respectively loaded) by line of boundary aggregates as shown in Fig. 1. The loading is applied by increasing deformation under prescribed loading rate. It certainly does not fully correspond to experimental loading, which was applied by Hopkinson bar setup. Loading by pressure (force) wave would be more realistic, but in our case, we can exactly specify the desired strain rate by prescribing it.

The experiments were performed on concrete and mortar discs with diameter $D = 70$ mm and with thickness $T = 30$ (further referred to as thin) and 55 mm (further referred to as thick). Model geometry was set accordingly. Two materials (concrete and mortar) were used in experimental study, therefore also two sets of material parameters were used for the modeling. Model parameters used in calculations are listed in Tab. 1. Parameter l_{\min} was set according to maximum aggregate size used in experiments. The material fracture parameters were estimated according to maximum load obtained by quasi-static tests on thin discs. For quasi-static simulations, the prescribed displacement rate of $1.67 \cdot 10^{-6}$ m/s was applied. Value of material meso-scale tensile strength was set to the value of macro-scopic tensile strength of the material and meso-level fracture energy was then found by fitting the model response to match the peak load only, since the post-peak behavior is hard to capture in the case of splitting test. The fit was performed separately for both simulated materials – concrete and mortar.

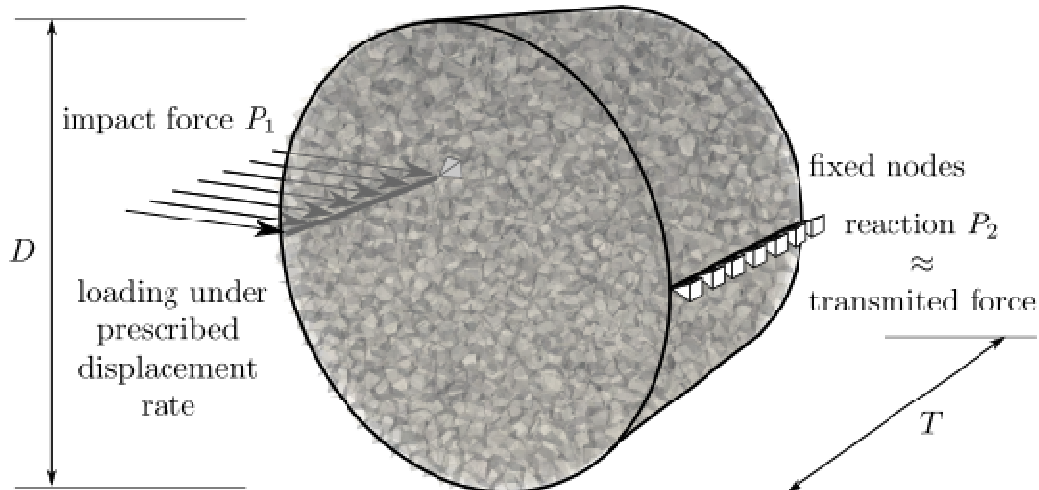


Figure 1 Model setup.

Table 1 Material parameters used from simulations of concrete and mortar discs.

	E_0 [GPa]	α	f_t [MPa]	G_f [N/m ²]	ρ [kg/m ³]	l_{min} [mm]
concrete	44	0.237	2.64	9.93	2400	7.50
mortar	40	0.237	3.47	40.66	2020	2.36

In the case of Brazilian splitting disc tests, the material tensile strength f_{tu} is estimated according to eq. (6), P_1 and P_2 are impact and transmitted force respectively, that are different in the case of dynamic loading, but equal in the case of quasi-static loading

$$f_{tu} = \frac{P_1 + P_2}{\pi TD} \quad (6)$$

After finding the best possible set of parameters to match the quasi-static response on thin specimens, the response of thick specimens was calculated. Lower strengths were obtained for thick discs, which corresponds to experimental data. Values of peak load and tensile strengths obtained by the model are listed in Tab. 2 together with experimental values. Note that rows corresponding to thin specimens are from the fitted response.

Table 2 Peak load and tensile strength obtained by the model compared to experiments.

		experiments [5]		simulations	
		peak load [kN]	strength [MPa]	peak load [kN]	strength [MPa]
concrete	thin	8.71	2.64	<i>fit</i> 8.71	2.64
	thick	15.71	2.60	15.24	2.52
mortar	thin	11.44	3.47	<i>fit</i> 11.44	3.47
	thick	20.26	3.35	23.01	3.80

The dynamical simulations were performed under wide range of displacement rates in correspondence to the experimental test. On the graph in the left part of Fig. 2 load-displacement curves are plotted for strain rates up to 150 s⁻¹. Significant delay in transmitted force can be observed as well as different increase of *impact* force P_1 compared to the increase in *transmitted* force P_2 . Greater increase in the first one is mainly caused by inertia.

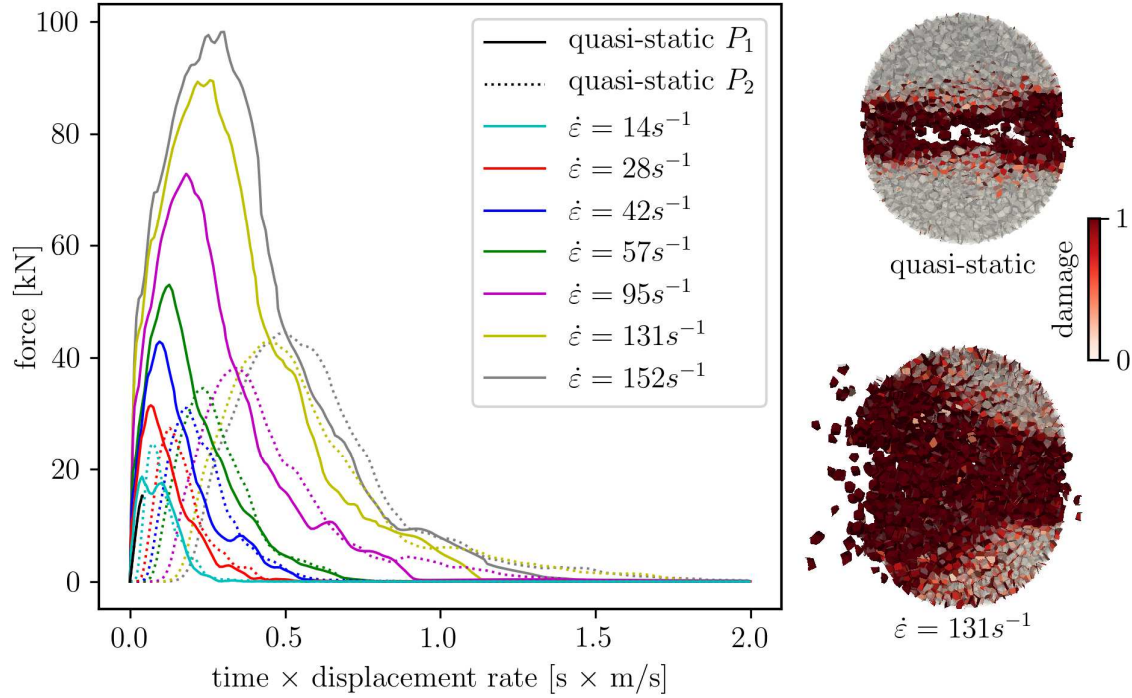


Figure 2 Force vs. displacement curve for thin mortar specimens (left) damaged volume obtained by simulation with 10 \times magnified displacements (right).

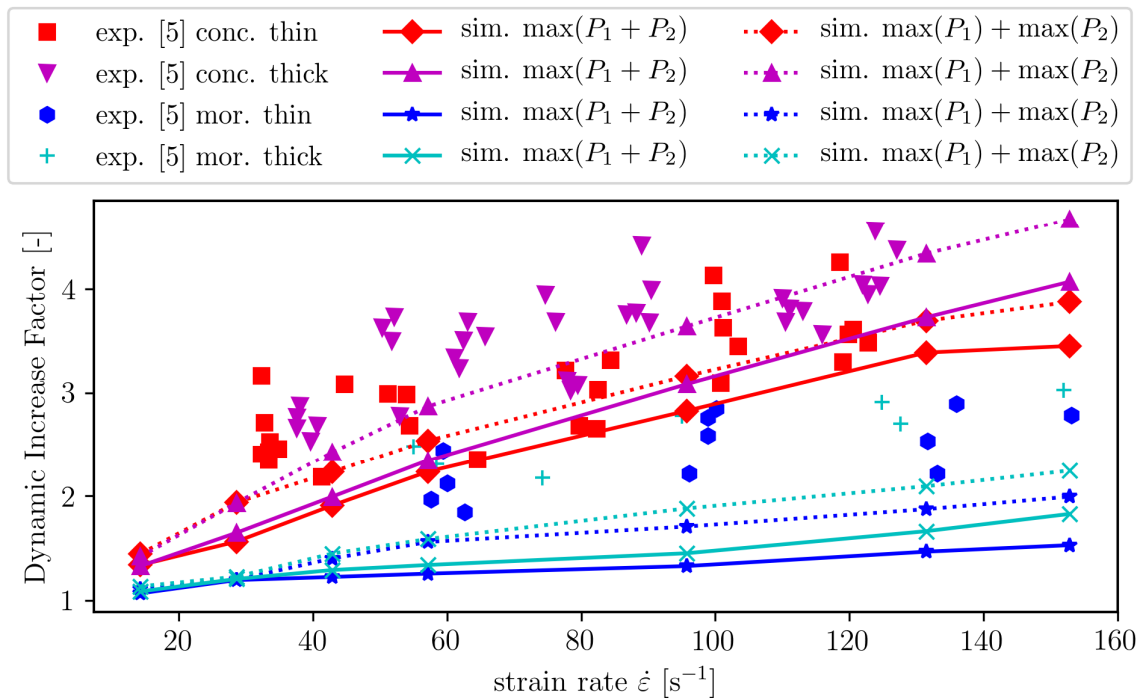


Figure 3 Dynamic increase factor of material tensile strength obtained by the model compared with experiments [5].

The smaller increase in *transmitted* force P_2 can be attributed to the change in material resistance due to greater strain rate.

Right part of Fig. 2 shows crack patterns for quasi-static and fast loading rate. In these pictures, displacements are 10× magnified. Particle facets are colored according to value of damage D . It can be clearly seen that in case of quasi-static loading, the model predicts the specimen to be split into two parts with relatively localized crack path, whereas in case of fast rate, the damaged volume is much greater, compare with [5].

Fig. 3 displays relation between material dynamic tensile strength according to Eq. (6) and strain rate obtained by the discrete particle model compared to the experimental data from [5]. Value of DIF is calculated by dividing dynamic and static strength. Since forces evolve in time, it is important to note the difference between $\max(P_1 + P_2)$ and $\max(P_1) + \max(P_2)$, even though the latter gives us results closer to the experimental evidence, the former should be taken into account for DIF estimation. The difference is not only in case of different material, but also for different thickness of the specimen, which corresponds to experimental evidence. However, even though the trends in increasing value of DIF are present, desired match is not obtained.

Conclusion

Discrete particle model was used for simulations of Brazilian splitting discs loaded under various strain rates. No strain rate dependency of constitutive law was imposed. It has been shown that the model is able to predict the increase in material strength, however, the desired match is not obtained. Authors believe that it is caused by the effects under the resolution of the model, thus finer resolution should be used for investigation of this phenomena. If course resolution is kept, the omitted part of the strain rate effect should be incorporated in the constitutive law phenomenologically.

Acknowledgement

Financial support of the project FAST-J-18-5412 is gratefully acknowledged.

References

- [1] BRARA, A., and KLEPACYKO, J. Experimental characterization of concrete in dynamic tension. *Mechanics of Materials* 38, 3 (2006), 253 – 267.
- [2] CUSATIS, Gianluca a Luigi CEDOLIN. Two-scale study of concrete fracturing behavior. *Engineering Fracture Mechanics*. 2007, 74(1-2), 3-17. DOI: 10.1016/j.engfracmech.2006.01.021
- [3] ELIÁŠ, J. Boundary Layer Effect on Behavior of Discrete Models. *Materials*. 2017, 10(2), 157-. DOI: 10.3390/ma10020157
- [4] ELIÁŠ, J. Adaptive technique for discrete models of fracture. *International Journal of Solids and Structures*. 2016, 100-101, 376-387. DOI: 10.1016/j.ijsolstr.2016.09.008
- [5] JIN, X, HOU, C., FAN, X., LU, C., YANG, H., SHU, X. and WANG, Z. Quasi-static and dynamic experimental studies on the tensile strength and failure pattern of concrete and mortar discs. *Scientific Reports*. 2017, 7(1), -. DOI: 10.1038/s41598-017-15700-2
- [6] NEWMARK, N. A method of computation for structural dynamics. University of Illinois, Urbana, 1959.
- [7] WU, H., ZHANG, Q., HUANG, F., and JIN, Q. Experimental and numerical investigation on the dynamic tensile strength of concrete. *International Journal of Impact Engineering* 32, 1 (2005), 605 – 617.
- [8] YAN, D., and LIN, G. Dynamic properties of concrete in direct tension. *Cement and Concrete Research* 36, 7 (2006), 1371 – 1378.

Computational Package for the Simulation of Plasma Microscopy Properties and Ion Beam-Plasma Interaction in High Energy Density Plasmas

†*Juan Miguel Gil^{1,2}, Rafael Rodríguez^{1,2}, Guadalupe Espinosa¹, and Pablo R. Beltrán¹

¹IUNAT, Departamento de Física, Universidad de Las Palmas de Gran Canaria, Spain.

²Instituto de Fusión Nuclear, Universidad Politécnica de Madrid, Spain.

*Presenting and †corresponding author: juanmiguel.gil@ulpgc.es

Abstract

Plasma microscopy properties and ion beam-plasma interaction are fundamental in many topics in plasma physics. Fundamental research and modelling in plasma atomic physics, like radiative properties and particle and laser beams-plasma interaction, continues to be essential for providing basic understanding and advancing on many different topics relevant to high-energy-density systems, particularly for nuclear fusion and astrophysics plasmas. In this work, we present a versatile computational package for the simulation and calculation of atomic and radiative properties of high energy density plasmas as well as the properties of ion-beam-plasma interaction processes. This computational package combines a set of theoretical and numerical approximations which yield substantial savings in computing running time, still comparing well with more elaborated codes and experimental data. Finally, calculations of several relevant plasma magnitudes for various plasma situations are shown and compared.

Keywords: Non-local thermodynamic equilibrium plasmas, Plasma atomic properties, Plasma radiative properties, Simulation and generation of databases of plasma properties, Ion beam-plasma interaction.

Introduction

Fundamental research and modelling in plasma atomic physics, like radiative properties and particle and laser beams-plasma interaction, continues to be essential for providing basic understanding and advancing on many different topics relevant to high-energy-density (HED) systems, particularly for nuclear fusion and astrophysics plasmas. Thus, in the field of inertial confinement fusion the radiative properties are the responsible of the absorption by the dopants in the fuel ablator of the thermal radiation in the indirect drive scheme. In magnetic confinement fusion devices, where the radiation emitted by the impurities could lead to thermal instabilities or to disruptions in the plasma edge. On the other hand, beam-matter interaction experiments are one of the key tools to investigate the fundamental physics properties of matter under extreme conditions, like high-energy-density plasmas, and a detailed theoretical description of the interactions allows us to diagnose the temperature and density, to obtain information about either the dynamic structure function or opacities and the equation of state of the plasmas. Also, the beam-matter interaction is one of the most essential problems in the nuclear fusion research area, in particular, in the field of beam-driven inertial confinement fusion, such as heavy-ion fusion or proton and ion fast ignition, where a precise knowledge of the energy deposition of the beam particles is required to design the fusion target. Therefore, the plasma properties are essential to analyze and explain both experiments and observations and also in their radiative-hydrodynamics numerical simulations. Furthermore, the simulations of the plasma properties in HED physics require the

development of complex theoretical models and their computational implementation for the generation of large plasma properties databases in a wide range of plasma conditions, as for example, atomic or opacity data. These plasma properties involve the calculation of a huge number of atomic levels (around 10^5) and atomic processes (around 10^7), by solving Dirac equations to obtain the atomic structure of each quantum atomic level of the ions considered in the plasma simulations, as well as, the cross sections of the atomic processes in the plasma. It is also necessary, to solve a very large set of coupled rate equations to obtain the average ionization of the plasma and the abundances of the atomic levels in both local thermodynamic equilibrium (LTE) and non-local thermodynamic equilibrium (NLTE) regimes. Moreover, this set must be solved for each plasma condition, i.e. density and temperature, the system must be resolved, and in a hydrodynamic simulation the profile of plasma conditions could involve around 10^3 of them.

In this work we present a versatile computational package based on three coupled codes named MIXKIP, RAPCAL [1-2] and STOPP, and their capabilities for the simulation of properties of the high energy density plasmas. MIXKIP code calculates the atomic structure and atomic kinetic of the ions in the plasma, RAPCAL code calculates the radiative properties and STOPP code calculates the energy losses of the ion beams crossing the plasma. This computational package combines a set of theoretical and numerical approximations which yield substantial savings in computing running time, still comparing well with more elaborated codes and experimental data. In order to show the capabilities of the package, calculations of several relevant plasma magnitudes for various plasma situations are shown and compared.

Theoretical and computational models

The determination of the atomic properties is the first step to be solved in the calculation of plasma properties in the HED plasmas. Our atomic model performs a detailed description (ion-by-ion) in the relativistic detailed configuration accounting approach (RDCA). So that each configuration i (either ground or excited one) of the ion with charge state ζ is characterized by one mono-electronic configuration, denoted by

$$C_{\zeta i} = \left\{ (c_{\zeta ik} \equiv nljm)_{k=1, \dots, N}^{w_{\zeta ik}} \right\}; \sum_{k=1}^{M_{Z, \zeta}} w_{\zeta ik} = N_{\zeta} = Z - \zeta$$

where n , l , j , m are the principal quantum number, orbital momentum, total angular momentum and third component of the total angular momentum respectively; N_{ζ} is the number of bound electrons of the ion, $w_{\zeta ik}$ denotes the orbital occupation integer number, and finally, $M_{Z, \zeta}$ is the number of orbitals used in the representation of the configurations which is fixed depending on parameters like temperature, density, charge state or nuclear charge among others. For a given ion, for each relativistic configuration, we solve the Dirac equation for each occupied orbital which gives us its relativistic mono-electronic wave function. The corresponding energy is obtained using the density-functional theory (DFT) for each ion in the context of Kohn-Shan equations [3] and assuming the local density approach (LDA) for the exchange and correlation energy as well as for the exchange and correlation potential [4]. This scheme for isolated detailed atom has been extended for atoms and ions in plasmas (non-isolated ion), both for weakly and strongly coupled plasmas, where the external potential of the bound electron system is due to the nucleus of the ion and the plasma surrounding. The Dirac equation for each orbital k of the ion with atomic number Z , in the charge state ζ and excited state i -th is given by

$$[c\vec{\alpha} \cdot \vec{p} + c^2\beta + U_{ef,\zeta i}(r; Z, \bar{Z}(n, T))]\varphi_{\zeta ik}(\vec{r}) = \varepsilon_{\zeta ik} \varphi_{\zeta ik}(\vec{r})$$

where c is the light speed, α y β are the Dirac matrices, ε 's are the mono-electronic level energies and φ 's are the mono-electronic bi-spinor wave functions which depend on the major and minor component, P and Q respectively, which are given by

$$\varphi_{\zeta ik}(\vec{r}) \equiv \varphi_{nljm}(\vec{r}) = \frac{1}{r} \begin{pmatrix} P_{nlj}(r)\Omega_{ljm}(\theta, \phi) \\ iQ_{nlj}(r)\Omega_{l'jm}(\theta, \phi) \end{pmatrix}$$

being $l' = l \pm 1$ and Ω_{ljm} the spherical bi-spinors. Moreover, major and minor component can be obtained from coupled radial equations it is follows

$$\begin{aligned} \frac{dP_{nlj}}{dr} &= -\frac{\kappa}{r}P_{nlj} + \frac{1}{c}(\varepsilon_{nlj} - U_{ef,\zeta i}(r; Z, \bar{Z}(n, T)))Q_{nlj} \\ \frac{dQ_{nlj}}{dr} &= \frac{\kappa}{r}Q_{nlj} - \frac{1}{c}(\varepsilon_{nlj} - U_{ef,\zeta i}(r; Z, \bar{Z}(n, T)) - 2c^2)P_{nlj} \end{aligned}$$

with $\kappa = (l - j)(2j + 1)$. Finally, U_{ef} is the self-consistent non isolated effective potential which take into account the interaction of one bound electron with the nucleus of the ion, with the rest of bound electrons and the plasma surrounding. This effective potential depends on the atomic configuration and some plasma parameters, and for non-isolated ions in weakly coupled plasmas, we use the non-isolated atomic potential developed in [5], which is given by

$$U_{ef,\zeta i}(r; a) = -\frac{1}{r} \left\{ (N_\zeta - 1) (\Phi_{\zeta i}(r) - \eta_{\zeta i}(r; a)) + [Z - N_\zeta + (N_\zeta - 1)\eta_{\zeta i}(0; a)]e^{-ar} + 1 \right\}$$

with the screening functions of the non-hydrogenic ions given by

$$\Phi_{\zeta i}(r) = \frac{1}{(N_\zeta - 1)} \left\{ Z - N_\zeta + 1 - r \left[\sum_k w_{\zeta ik} \int \frac{|\varphi_{\zeta ik}(r')|^2}{|\vec{r} - \vec{r}'|} d\vec{r}' + U_{ex,\zeta i}(r) \right] \right\}$$

with

$$\rho_{\zeta ik}(r) \equiv |\varphi_{\zeta ik}(r)|^2 = \frac{1}{r^2} (P_{\zeta ik}(r) + Q_{\zeta ik}(r))^2$$

and

$$\eta_{\zeta i}(r) = \frac{1}{2}a \int_0^\infty e^{-a|s-r|} \phi_{\zeta i}(s) ds$$

being a the inverse of the Debye radius which depend on the electron density n_e and temperature T as well as on the average ionization \bar{Z} of the plasma, and is given by

$$a = \left[\frac{4\pi n_e}{T} \frac{(\bar{Z} + \bar{Z}^2)}{\bar{Z}} \right]^{1/2}$$

For strongly coupled plasmas the non-isolated atomic potential used [6,7] is given by

$$U_{ef,\zeta i}(r; a) = -\frac{1}{r} (N_\zeta - 1) (\Phi_{\zeta i}(r) - \Phi_{\zeta i}(r; R_o)) + \frac{\bar{Z}}{2R_o} \left(1 - \frac{r^2}{R_o^2} \right)$$

at $r < R_o$ and $U_{ef,\zeta j}(r; a) = 0$ at $r > R_o$, being R_o the sphere-ion radius which is given by

$$R_o = \left[\frac{3\bar{Z}}{4\pi n_e} \right]^{1/3}$$

From the wave functions and level energies of all orbitals of each configuration, we calculate the relativistic configuration energy $E_{\zeta j}$, given by

$$E_{\zeta i}(r) = \sum_k w_{\zeta ik} \varepsilon_{\zeta ik} - \frac{1}{2} \sum_k w_{\zeta ik} \int |\varphi_{\zeta ik}(r')|^2 U_{\zeta i}(r) d\vec{r} d\vec{r}'$$

where $U_{\zeta i}(r)$ is the atomic potential due to the bound electrons of the ion, which can be obtained as $U_{\zeta i} = (Z/r) + U_{ef,\zeta i}(r; a = 0)$. Oscillator strength of the transition, from the mono-electronic orbital k to k' of the atomic configurations ζi and $\zeta i'$, respectively, is evaluated in the electric dipole approximation as it follows

$$f_{\zeta i, \zeta i'} = w_{\zeta ik} \left(1 - \frac{w_{\zeta i' k'}}{(2j' + 1)} \right) f_{\zeta ik, \zeta i' k'}$$

with

$$f_{\zeta ik, \zeta i' k'} = \frac{2}{3} (\varepsilon_{\zeta j' k'} - \varepsilon_{\zeta j k}) (2j' + 1) \times \\ \times \begin{pmatrix} j & 1 & j' \\ 1/2 & 0 & -1/2 \end{pmatrix} \left\{ \int_0^\infty [P_{\zeta ik}(r) P_{\zeta i' k'}(r) + Q_{\zeta ik}(r) Q_{\zeta i' k'}(r)] r dr \right\}$$

where $\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ is the 3-j symbol. Another important magnitude overall in the interaction of particle beams with plasma is the mean excitation energy $I_{\zeta j}$ given by

$$\ln I_{\zeta i} = \frac{1}{N_\zeta} \sum_k w_{\zeta ik} \ln I_{\zeta ik}$$

where the mean excitation energy of each orbital is obtained as

$$I_{\zeta ik} = \sqrt{\frac{E_{c, \zeta ik}}{r_{\zeta ik}^2}}$$

being

$$E_{k, \zeta ik} = \varepsilon_{\zeta j k} - \int \varphi_{\zeta ik}^*(r) U_{ef, \zeta i}(r; Z, \bar{Z}(n, T)) \varphi_{\zeta ik}(r) d\vec{r}$$

and

$$r_{\zeta ik}^2 = \int \varphi_{\zeta ik}^*(r) r^2 \varphi_{\zeta ik}(r) d\vec{r}$$

Energies and oscillator strengths play a fundamental role in the photon absorption and emission processes from the plasma and in the spectral properties of the plasma, and mean excitation energies, in the energy deposition of ion beams in the plasmas.

Dirac equation is solved by a fourth order Runge-Kutta method in a linear or exponential radial mesh, and taking into account the semi-classical non-relativistic approximation at the origin and at practical infinity. There are many atomic equations (around 10^6) as configurations considered, and they are coupled with atomic rate equations by means of the average ionization of the plasma, and therefore, they must be solved iteratively. At high

temperature and low density, the effective potential tends to the isolated one, and the atomic and rate equations became uncoupled. On the other hand, although the number of atomic levels for a given isolated ion is infinite, this number should be finite for the atoms in a plasma if we want to obtain a satisfactory simulation of their radiative properties. This is due to the coulomb interaction between the bound electrons and the surrounding plasma, so we have to make a previous selection of the levels. However, there is not a priori criterion to determine which configurations should be included in the model. In general, the kind of configurations to include depends on the plasma conditions, the presence of external radiation fields or the interaction with particle beams. The experience achieved, based on the large number of cases studied during the development of the computational package, has led us to consider a complete enough set of configurations which allow us to obtain reasonable average ionization and ion abundances or populations. This set of configuration must be extended to obtain reasonable radiative properties of the plasma. In this case, the criterion employed was based on a rule of thumb in which the configurations included for each ion in the model are those with energies up to twice the ionization energy of the ground configuration of the ion. As said before, as the plasma density increase, screening effects due to neighbouring electrons and ions modify the atomic potential, and therefore, the ionization energy of the ground configuration. This effect is commonly modelled in plasma physics through the so-called continuum lowering, that represents the depression of the potential with respect to the isolated situation due to the electric fields generated by the plasma charged particles. This effect can restrict the number of configurations and, then, it determines the final set of atomic levels used for each plasma conditions. In this work, the model used for the continuum lowering is based on that provide by our self-consistent non isolated potentials, or based on the widely used one developed by Stewart and Pyatt [8]. Another question is related with the degree of detail of the atomic description. The most detailed description is the so-called detailed level accounting (DLA) approach. However, this description entails very large computational times and, therefore, it is only useful for chemical elements of low atomic number. This computational package has been designed to provide atomic data generated in the RDCA approach from low to high Z plasmas by solving the scheme explained before. Furthermore, this package has been also designed to work with external atomic data tables or codes. The current external atomic source is FAC code [9], which is designed to provide atomic data in DCA and DLA approaches.

Once the atomic data have been obtained, they are used in the second step which is the determination of the plasma level populations. The models commonly used for their calculation are based in the so-called collisional-radiative (CR) models [10], which are valid for plasmas either at LTE or NLTE. In these models, one has to solve a set of coupled atomic rate equations (one per each relativistic configuration in the present model) given by

$$\frac{dN_{\zeta i}(\mathbf{r}, t)}{dt} = \sum_{\zeta' i'} N_{\zeta' i'}(\mathbf{r}, t) \mathbb{R}_{\zeta' i' \rightarrow \zeta i}^+ - \sum_{\zeta' i'} N_{\zeta i}(\mathbf{r}, t) \mathbb{R}_{\zeta i \rightarrow \zeta' i'}^-$$

where $N_{\zeta i}$ is the population density of the atomic configuration or level i of the ion with charge state ζ . The terms $\mathbb{R}_{\zeta' i' \rightarrow \zeta i}^+$ and $\mathbb{R}_{\zeta i \rightarrow \zeta' i'}^-$ take into account all the atomic processes, both collisional and radiative, that contribute to populate and depopulate the configuration i of the ion ζ , respectively. Moreover, two complementary equations have to be satisfied together with the set of atomic rate equations. First, the conservation of the ion density n_{ion} ,

$$\sum_{\zeta=0}^Z \sum_{i=1}^{M_{Z,\zeta}} N_{\zeta i} = n_{ion}$$

and, second, the charge neutrality condition in the plasma,

$$\sum_{\zeta=0}^Z \sum_{i=1}^{M_{Z,\zeta}} \zeta N_{\zeta i} = n_e$$

In our computational package we solve the atomic rate equations for mono and multicomponent plasmas. For the last situation, we have also provide as an input data, the molar fractions of the chemical elements in the mixture which are given by $x_m = n_{ion,m}/n_{ion}$ and they have to be satisfied $\sum_m x_m = 1$, with m runs over the number of chemical elements in the plasma. The atomic processes included in the present CR model are: collisional ionization and three body recombination, spontaneous decay, collisional excitation and de-excitation, radiative recombination, electron capture and auto-ionization. The cross sections of theses atomic processes in the plasma are calculated by mean of widely known analytical expressions [11] which depend on the atomic data calculated in the theoretical frame mentioned above. At this point we want to highlight that in the computational package are also implemented the Saha-Boltzmann and Coronal Equilibrium equations which provide the asymptotic behaviour of the atomic rate equations at high and low plasma electron density, respectively.

The rate coefficients of collisional processes between an ion, which goes from the state ζi to $\zeta' i'$, and one free electron, with incident energy ε , are obtained as

$$\langle v \sigma_{\zeta i \rightarrow \zeta' i'} \rangle = \int_{E_{th}}^{\infty} v \sigma_{\zeta i \rightarrow \zeta' i'}(\varepsilon) f(\varepsilon) d\varepsilon$$

where σ is the cross section of the processes, $v = \sqrt{2\varepsilon/m_e}$ and $f(\varepsilon)$ are, respectively, the speed and the distribution function of the free electrons. Free electron distribution function has to be solved from the free electron rate equation which is coupled with atomic rate equations. In our model, free electron plasma is assumed in LTE, and therefore, the free electron distribution function is given by the Boltzmann's distribution which is characterized by the plasma electron temperature. The rate coefficients of radiative processes between an ion, which goes from the state ζi to $\zeta' i'$, and one photon, with energy $h\nu$, are obtained as

$$\langle c \sigma_{\zeta i \rightarrow \zeta' i'} \rangle = \int_{E_{th}/h}^{\infty} c \sigma_{\zeta i \rightarrow \zeta' i'}(\nu) f_{ph}(\nu) d\nu$$

being f_{ph} the photon distribution function which is related with the spectral radiation intensity of the photons by mean $I_\nu = ch\nu f_{ph}(\nu)$. Radiative transfer equation provides the spectral radiation intensity from the following equation

$$\frac{1}{c} \frac{\partial I_\nu(\mathbf{r}, t, \nu, \mathbf{n})}{\partial t} + \mathbf{n} \cdot \nabla I_\nu(\mathbf{r}, t, \nu, \mathbf{n}) = -\kappa(\mathbf{r}, t, \nu) I_\nu(\mathbf{r}, t, \nu, \mathbf{n}) + j(\mathbf{r}, t, \nu)$$

where ν is the photon frequency, \mathbf{n} is a unit vector in the direction of propagation of the photons for any value of solid angle Ω , and finally, $j(\mathbf{r}, t, \nu)$ and $\kappa(\mathbf{r}, t, \nu)$ are the monochromatic emissivity and absorption coefficients, respectively. Both coefficients include electron transitions in the plasma between atom bound levels (line transitions or bound-bound contributions), between bound and free levels (photoionization and radiative recombination which are bound-free contributions) and between electron free levels (direct and inverse bremsstrahlung or free-free contributions). The expressions used to calculate them can be found elsewhere [12]. In our model, radiative transfer equation is assumed in stationary conditions, and therefore, the first sum in the left hand of the transfer equation equals zero. The formal solution of the radiative transfer equation, along the ray defined by the points of space (s_o, s) , is given by

$$I_\nu(s, \nu) = I_\nu(s_o, \nu)e^{\tau(s_o, s; \nu)} + \int_{s_o}^s j(s', \nu)e^{\tau(s', s; \nu)} ds'$$

being τ the optical depth, given by

$$\tau(s_o, s; \nu) = \int_{s_o}^s \kappa(s', \nu) ds'$$

Also, it is assumed a uniform distribution of emitting atoms and isotropic emission in the plasma in the three basic geometries (plane, cylindrical and spherical), and moreover, there are not external radiation fields. Then, spectral radiation intensity can be written as

$$I_\nu(s, \nu) = \frac{j(\nu)}{\kappa(\nu)} (1 - e^{-\kappa(\nu)(s-s_o)})$$

Two ingredients are needed to compute the spectral radiation intensity, or $j(\mathbf{r}, t, \nu)$ and $\kappa(\mathbf{r}, t, \nu)$. First, the cross sections of the radiative processes, which are obtained through atomic simulations. Secondly, the populations of the atomic levels in the plasma obtained from atomic rate equations. On the other hand, as the rate equations included the radiative processes in the plasma, i.e. spectral radiation intensity or the absorption and emissivity coefficients, one has to solve a set of coupled equations which are formed by atomic rate equations and radiative transfer equation. For optically thin plasmas, where the radiative processes can be despised in the atomic rate equations, the coupling is avoided and the atomic rate equations and radiative transfer equation can be solved separately: linear system of atomic rate equations are solved, and once the absorption and emissivity coefficients are obtained from atomic level populations, spectral radiation intensity is calculated. In the opposite case, for optically thick plasmas, the set of coupled equations are solved in the context of the escape factor formalism which avoids the explicit resolution of the radiative transfer equation. The formal solution of the transfer equation is introduced in the atomic rate equations, through rate coefficients of the bound-bound radiative processes, and now, a non-linear system of atomic rate equations is solved.

Another issue to consider is the method chosen for the solution of the atomic rate equations solver. When the electron density is taken as input parameter characterizing the population kinetics problem, then the set of rate equations constitute a linear system of M equations for the level populations, where M denotes the total number of levels included in the collisional-radiative model. Solvers can be broadly classified into two categories, direct and iterative. The direct solvers compute a solution which is guaranteed to be as accurate as the problem definition. The amount of time required to obtain a solution by such algorithms typically scale like M^3 . In a population kinetics problem, the number of levels can reach the order of 10^5 , so for such large systems of linear equation the direct methods lead to prohibitively long

run times. For the kind of problems that we are interested, the iterative methods yield an approximation to the solution significantly faster than direct method. Furthermore, iterative methods typically require less memory than direct ones and hence can be the only means of solution of the large systems of equations. Our code uses as initial population distribution the solution provides by either the corona model or the Saha-Boltzmann equations depending on whether the electron density of the case analysed is closer to the regime of low or high-density respectively.

Once the atomic data and plasma level populations the atomic configurations are obtained, by using the MIXKIP code, they are used as input data in RAPCAL code, to determine the radiative properties (spectrally resolved opacity and emissivity, specific intensity, mean Planck and Rosseland opacities or radiative power losses), or in the STOPP code to determine the energy deposition of the ion beams in plasmas (stopping power, energy losses or range). The basic equations of our radiative emissivity model to calculate $j(\mathbf{r}, t, \nu)$, has three contributions. The bound-bound contribution, $j_{bb}(\mathbf{r}, t, \nu)$, is given by

$$j_{bb}(\nu) = \sum_{\zeta} \sum_{i,j} j_{\zeta j \rightarrow \zeta i}(\nu), \quad j_{\zeta j \rightarrow \zeta i}(\nu) = \frac{h\nu}{4\pi} N_{\zeta j} A_{\zeta j \rightarrow \zeta i} N \phi_{ij}(\nu) \quad (14)$$

where we have omitted the dependence on the position and time in the formula for simplicity. $A_{\zeta j \rightarrow \zeta i}$ is the Einstein coefficient for spontaneous de-excitation between the bound states j, i of the ion ζ and h is the Planck's constant. $\phi_{ij}(\nu)$ is the line profile and in its evaluation of the line profile, natural, Doppler, and electron-impact [13] broadenings were included and also the Unresolved Transition Array width [14], which is a statistical method to take into account the atomic fine structure of the spectra in the DCA atomic approach used in this work. The line-shape function is applied with the Voigt profile that incorporates all these broadenings. The bound-free contribution to the emissivity, $j_{bf}(\mathbf{r}, t, \nu)$, is determined by means of

$$j_{bf}(\nu) = \sum_{\zeta+1,j} \sum_{\zeta,i} j_{\zeta+1,j \rightarrow \zeta,i}(\nu) \quad (15)$$

$$j_{\zeta+1,j \rightarrow \zeta,i}(\nu) = \frac{h^4 \nu^3 n_e}{2\pi c^2 \varepsilon_0^{\frac{1}{2}}} \left(\frac{1}{2m_e} \right)^{\frac{3}{2}} N_{\zeta+1,j} f(\varepsilon) \frac{g_{\zeta,i}}{g_{\zeta+1,j}} \sigma^{\text{pho}}_{\zeta+1,j \rightarrow \zeta,i}(\nu)$$

where ε is the free electron energy and m_e the electron mass. In this work, a Maxwell-Boltzmann distribution $f(\varepsilon)$ at the electron temperature is assumed. Photoionization cross section, $\sigma^{\text{pho}}_{\zeta+1,j \rightarrow \zeta,i}(\nu)$, were calculated quantum-mechanically using the FAC code in the relativistic distorted wave approach. $g_{\zeta,i}$ denotes the statistical weight of level i . Finally, for the free-free contribution to the emissivity a semi-classical expression, based on the Kramer's inverse bremsstrahlung cross section [15], was used

$$j_{ff}(\nu) = \frac{32\pi^2 e^4 a_0^2 \alpha^3}{\sqrt{3}(2\pi m_e)^{3/2} h} \left(\frac{m_e}{2\pi k_B T_e} \right)^{1/2} \overline{Z^2} n_{\text{ion}} n_e e^{-h\nu/k_B T_e} \quad (16)$$

where k_B is the Boltzmann's constant and α is the fine structure constant.

Finally, we present the basic equations of the ion beam-plasma interaction model implemented in STOPP code, which allow us to determine the energy losses and the range of the ion beam from the stopping power. We consider two contributions to the total stopping power, both bound and free electrons. Bound contribution proposed is given by

$$S_b(v_p, Z, Z_p; n_{at}, T, \bar{Z}) = \frac{Z_p^2 e^4 n_{at}}{4\pi\epsilon_0^2 v_p^2} L_b$$

where the bound stopping number L_b is given by

$$L_b = \begin{cases} (Z - \bar{Z}) \left(\log \left(\frac{2m_e}{I} \right) - \frac{2E_k}{m_e v_p^2} \right), & \text{if } v_p > v_{int} \\ (Z - \bar{Z}) \left(\frac{\alpha v_p^3}{1 + G v_p^2} \right), & \text{if } v_p < v_{int} \end{cases}$$

where Z_p is the charge of the projectile, I and E_k are the mean excitation energy and mean kinetic energy of the bound electrons in the plasma, given by

$$I = \prod_{\zeta i} I_{\zeta i}^{p_{\zeta i} N_{\zeta} / (Z - \bar{Z})}$$

and

$$E_k = (1/Z - \bar{Z}) \sum_{\zeta i} p_{\zeta i} E_{k, \zeta i}$$

being $p_{\zeta i}$ the normalized population density, α the viscosity coefficient of the bound electrons, given by

$$\alpha = 1.067 E_k^{1/2} I^2$$

and finally, G is a constant determined from the continuity condition of the bound stopping number at v_{int} , with $v_{int} = \sqrt{3E_k + 1.5I}$.

Free electron contribution to the total stopping power is given by [16]

$$S_e(v_p, Z, Z_p; n_e, T, \bar{Z}) = \frac{Z_p^2 e^4 n_e}{4\pi\epsilon_0^2 v_p^2} (G_e L_e + H_e \log X_e)$$

with the free stopping number given by

$$L_e = \log \left(\frac{R_D}{b} \right)$$

where R_D is the Debye radius, $b = Z_p e^2 / 4\pi\epsilon_0 m_e (v_p^2 + v_e^2)$ is the impact parameter, and v_e is the mean velocity of the free electrons. G_e and H_e functions are given by

$$G_e = \text{erf}(X_e / \sqrt{2}) - \sqrt{\frac{2}{\pi}} X_e e^{-X_e^2 / 2}$$

and

$$H_e = -\frac{X_e^3 \log X_e e^{-X_e^2 / 2}}{3\sqrt{2\pi}} + \frac{X_e^4}{(X_e^4 + 12)}$$

where X_e is the ratio between the projectile and free electron speeds. The ion-beam plasma interaction model implemented in STOPP code is valid to linear interaction regimen, non-degenerated free electrons, and finally, fast collisions.

Results

In this section we present some examples of the simulations carried out with the computational package based on the MIXKIP, RAPCAL and STOPP codes present before. First, we consider the spectral emission of shock waves in xenon plasmas doped with helium, at density and temperature conditions of the shock waves generated at PALS [17]. In Figure 1 we show the total emissivity and opacity, *vs* photon energy, of the mixture plasma with 90% of Xe and 10% of He, at matter density and electron temperature given by 1.5 gcm^{-3} and 10 eV, respectively. The spectral emissivity and opacity of the Xe and He as well as the total are showed, and we focus our attention in the experimental spectral window given by 35-80 eV. At this spectral energy, photons from Lyman series of the helium-like-hydrogen ion are present, and their characterization can be used to diagnose the electron density and temperature of the plasma. In

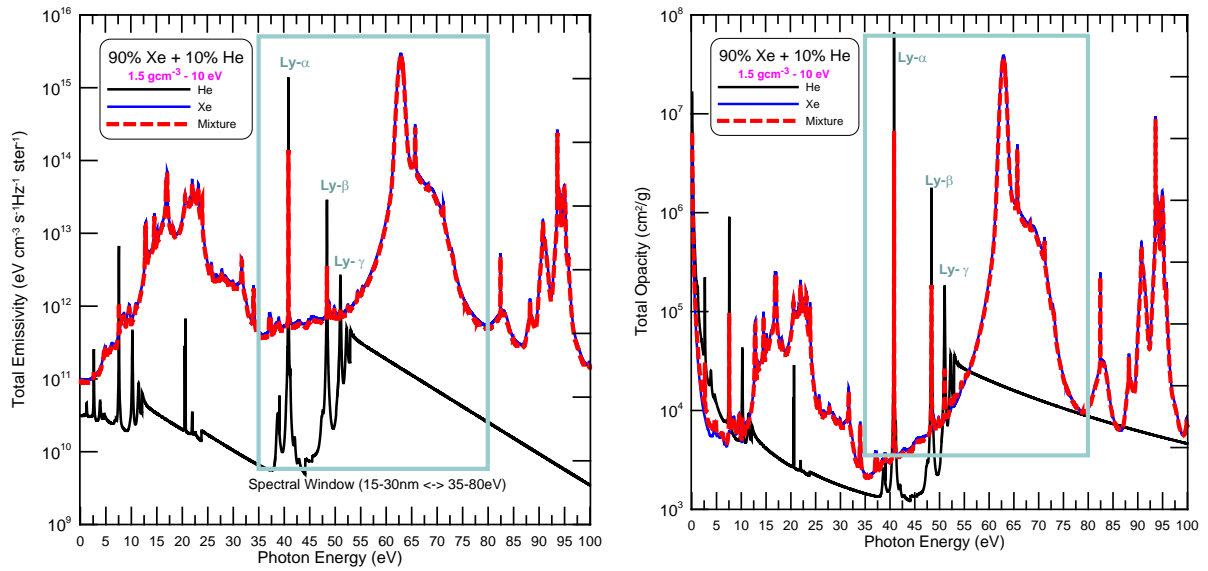


Figure 1. Spectral emissivity and opacity of Xe (90%) and He (10%) plasma mixture at 1.5 gcm^{-3} and 15 eV.

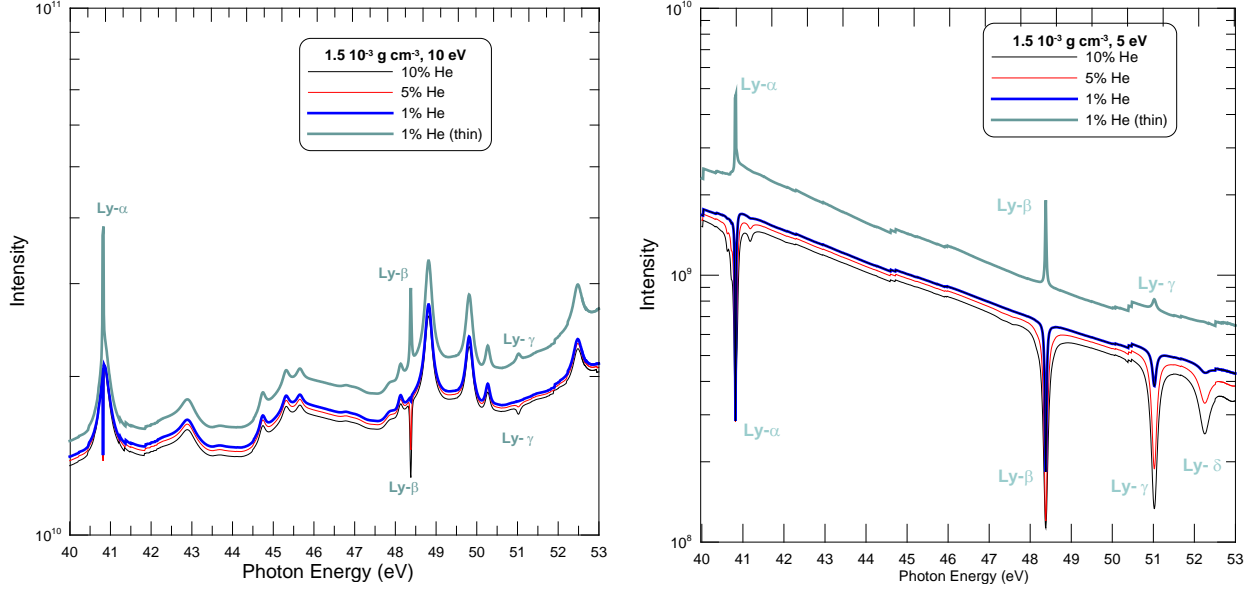


Figure 2. Specific radiative intensity emitted by Xe (90%) and He (10%) plasma mixture at 1.5 g cm^{-3} and 6 and 10 eV.

Figure 2 we show the specific radiative intensity emitted by a portion of plasma given by $0.6 \times 4 \times 0.025 \text{ mm}$, at matter density of 1.5 g cm^{-3} and temperatures of 5 and 10 eV, and finally, for different molar fractions of helium in the plasma. It can be observed that the intensity decrease when the temperature and molar fraction arise. It can also be observed the contribution to the total intensity of the Lyman series emitted from helium and the sensibility of the spectrum to the temperature.

In Figure 3 we show the stopping power and kinetic energy of proton beam at 0.5 MeV in aluminum plasma at 50 eV and different atom or ion densities, as function of depth in the plasma. It has been calculated by solving the atomic rate equation in the optically thin approximation and Saha-Boltzmann equations with the aim to simulate NLTE and LTE thermodynamic regimes in the plasma. At low densities, important differences in the stopping and kinetic energies of the proton are observed, while at high densities, NLTE and LTE simulations provide very similar results. The maximum stopping power calculated from atomic rate equations (NLTE) is lower than those obtained from Saha-Boltzmann equations (LTE) while the range is greater.

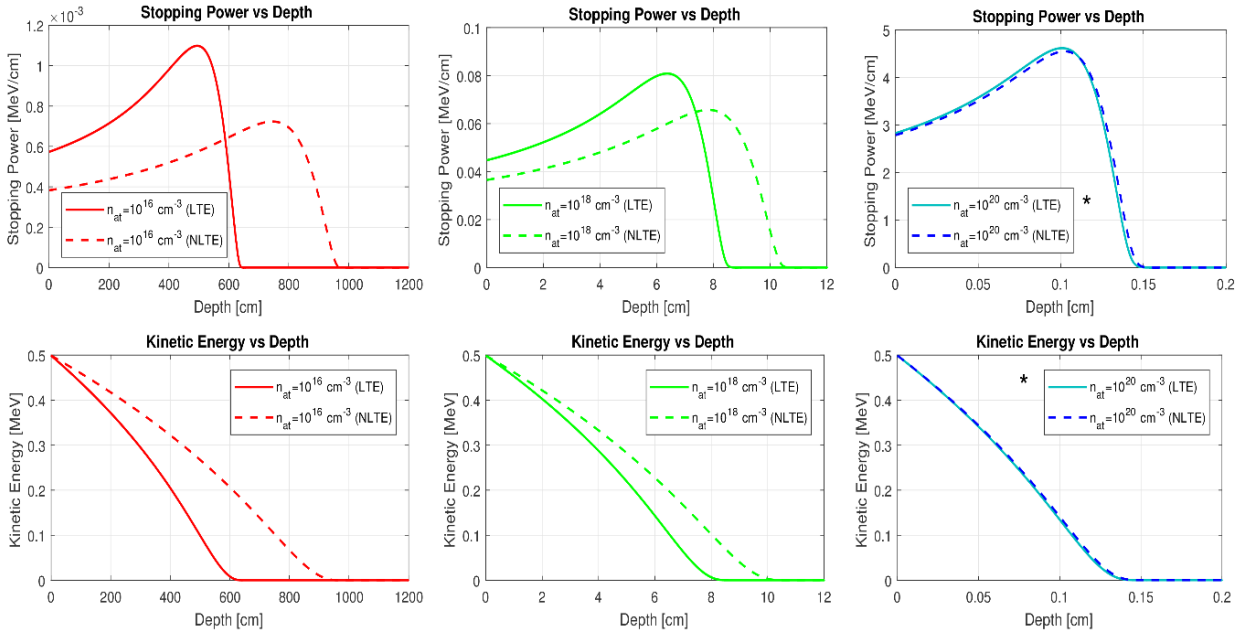


Figure 3. Stopping power and kinetic energy of proton beam at 0.5 MeV in aluminum plasma at 50 eV and different ion densities, as function of depth in the plasma.

Conclusions

In this work we have presented a versatile computational package designed to simulate plasma properties of the high energy density physics in a wide range of plasma conditions and combine a set of theoretical and numerical approximations which yield substantial savings in computing running time, still comparing well with more elaborated codes and experimental data. So, it can be simulated the atomic structure, atomic kinetic and radiative properties as well as the energy deposition of the ion beams in mono and multicomponent plasmas at LTE and NLTE thermodynamic regimens. It can also be simulated the properties of the optically thin and thick plasmas, and finally, of the non-stationary plasmas. These plasma properties involve the calculation of a huge number of atomic levels (around 10^5) and atomic processes (around 10^7), by solving Dirac equations to obtain the atomic structure of each quantum atomic configuration of the ions considered in the plasma simulations, as well as, the cross sections of the atomic processes in the plasma. It is also necessary, to solve a very large set of coupled rate equations to obtain the average ionization of the plasma and the abundances of the atomic configurations. Moreover, this set must be solved for each plasma condition, i.e. density and temperature, the system must be resolved, and in a hydrodynamic simulation the profile of plasma conditions could involve around 10^3 of them. Finally, calculations of several relevant plasma magnitudes for various plasma situations are shown and compared.

Acknowledgements

This work has been supported by the EUROfusion Consortium TASK AGREEMENT WPENR: Enabling Research IFE, Project No. AWP15-ENR-01/CEA-02, by the Project of the Spanish Government with reference FIS2016-81019-P.

References

- [1] Espinosa, G., Rodríguez, R., Gil, J.M., Suzuki-Vidal, F., Lebedev, S.V., Ciardi, A., Rubiano, J.G. and Martel, P. (2017) Influence of atomic kinetics in the simulation of plasma microscopic properties and thermal instabilities for radiative bow shock experiments, *Physical Review E* **95**, 033201.
- [2] Rodriguez, R., Florido, R., Gil, J.M., Rubiano, J.G., Suarez, D., Martel, P., Minguez, E. and Mancini, R.C. (2010) Collisional-Radiative calculations of optically thin and thick plasmas using the computational package ABAKO/RAPCAL, *Communications in Computational Physics* **8**, 185-210.
- [3] Kohn, W. and Sham, L.J. (1965) , *Physical Review A* **140**, 1133.
- [4] Gunnarsson, O. and Lundvist, B. (1976), *Physical Review B* **13**, 4274.
- [5] Gil, J.M., Martel, P., Mínguez, E., Rubiano, J., Rodríguez, R. and Ruano, F. (2002), *J. Quantum Spec. Radiat. Transfer* **75**, 539.
- [6] Gil, J.M., Martel, P., Rubiano, J., Rodríguez, R., Mínguez, E. and Doreste, L. (1997), in *Advances in laser interaction with matter and inertial fusion (World Scientific)*, pp. 335-338.
- [7] Rodríguez, R., Gil, J.M. and Florido, R. (2007), *Phys. Scr.* **76**, 418.
- [8] Stewart, J.C. and Pyatt, K.D. (1966) Lowering of ionization potentials in plasmas (1966), *The Astrophysical Journal* **144**, 1203-1211.
- [9] Gu, M.F. (2008) The flexible atomic code, *Canadian Journal of Physics* **86**, 675-689.
- [10] McWirth, R.W. P. (1978) Data needs, priorities and accuracies for plasma spectroscopy, *Physics Reports* **37**, 165-209.
- [11] Rodriguez, R., Florido, R., Gil, J.M., Rubiano, J.G., Martel, P. and Mínguez, E. (2008) RAPCAL code: A flexible package to compute radiative properties for optically thin and thick low and high-Z plasmas in a wide range of density and temperature, *Laser and Particle Beams*, **26**, 433-448.
- [12] Rodriguez, R., Florido, R., Gil, J.M., Rubiano, J.G., Suarez, D., Martel, P., Minguez, E. and Mancini, R.C. (2010) Collisional-Radiative calculations of optically thin and thick plasmas using the computational package ABAKO/RAPCAL, *Communications in Computational Physics* **8**, 185-210.
- [13] Dimitrijevic, M.S. and Konjevic, N. (1987) Simple estimates for Stark-broadening of ion lines in stellar plasmas, *Astronomy & Astrophysics* **172**, 345-349.
- [14] Bauche, J., Bauche-Arnoult, C. and Klapisch, M. (1987) Transition arrays in the spectra of ionized atoms, *Advance Atomic Molecular Physics* **23**, 131-195.
- [15] Rose, S.J. (1992) Calculation of the radiative opacity of laser-produced plasmas, *Journal of Physics B: Atomic and Molecular Physics* **25**, 1667-1681.
- [16] Thomas, P. and Meyer-ten-Vehn, J. (1991), *Phys. Rev. A* **43**,4.
- [17] Singh, R.L. et al. (2017), *High Energy Density Physics* **23**, 20-30.

Numerical simulation of metal-intermetallic laminate composites failure under dynamic loading

†*Sergey A. Zelepugin^{1,2}, Alexey S. Zelepugin^{1,2}, Alexey Popov², and Dmitri Yanov²

¹Physical and Technical Department, Tomsk State University, Russia.

²Department for Structural Macrokinetics, Tomsk Scientific Center SB RAS, Russia.

*Presenting author: szel@yandex.ru

†Corresponding author: szel@yandex.ru

Abstract

The processes of multilayer composites failure under dynamic loading were investigated. Fracture model of brittle materials subjected to high velocity impact was used. Deformation and fracture of Al₃Ti - Ti metal-intermetallic laminate composite materials under dynamic loading was numerically simulated using the finite element method.

Keywords: **MIL composites, dynamic loading, failure**

Introduction

Progress in the creation of new technological innovations mainly depends on the development and improvement of technologies for obtaining materials with required properties, so the creation of materials with desired structural and functional properties is currently an area of increased attention in materials science and technology. A new promising class of structural materials includes metal-intermetallic laminate composite materials (MILCM) which are represented by a multilayer composition with alternating metal and intermetallic layers [1]-[7]. These composite materials are attractive for use in aerospace engineering and many other areas, and methods for obtaining of MILCM allow us to use new technologies expanding the functionality of laminate composites and the area of application.

In this work the processes of high-velocity interaction of a projectile with a multilayer MILCM target were numerically investigated in axisymmetric geometry using the finite element method. The set of equations for describing unsteady adiabatic motion of an elasto-plastic medium, including nucleation and accumulation of microdamages and temperature effects, consists of the equations of continuity, motion, and energy [7]-[9]. To simulate numerically the failure of the material under high velocity impact, we applied the active-type kinetic model determining the growth of microdamages, which continuously changes the properties of the material and induce the relaxation of stresses. The strength characteristics of the medium (shear modulus and dynamic yield strength) depended on temperature and the current level of damage taking into account probabilistic approach to numerical simulation of fracture [10]-[12]. The critical specific energy of shear deformations was used as a criterion for the erosion failure of the material that occurred in the region of intense interaction and deformation of contacting bodies. To simulate the brittle-like failure of the intermetallic material under high velocity impact, we modified the kinetic model of failure and included the possibility of failure above Hugoniot elastic limit (HEL) in the shock wave and the sharp drop in the strength characteristics for the failure of material.

Formulation of the Problem

In the computations we used the target consisting from 17 composite intermetallic Al₃Ti - titanium alloy Ti-6-4 layers. Total thickness of the target was 19.89 mm. The thicknesses of intermetallic layer and the layer of titanium alloy were varied. The penetrator used was a tungsten heavy alloy rod with an initial diameter of 6.15 mm and length of 23 mm [1]. Initial impact velocity was of 900 m/s.

To simulate numerically the processes of high velocity shock loading, we use the model of an elastic-plastic damaged medium characterized by the presence of microcavities (pores, cracks). In the model the total volume of the medium W comprises the undamaged part of the medium of density ρ_c which occupies volume W_c , and microcavities of zero density which occupy volume W_f . The average density of the damaged medium is connected with the above-introduced parameters by the relationship $\rho = \rho_c(W_c/W)$. The degree of damage of the medium is characterized by the specific volume of microcavities $V_f = W_f/(W \cdot \rho)$.

A mathematical model used in the numerical code for solving high velocity impact problems is based upon a set of differential equations of continuum mechanics. The system of equations governing the nonstationary, adiabatic (for both elastic and plastic deformation) motion of a compressible medium with allowance for the evolution of microdamages comprises the continuity equation, the equation of motion, the energy equation [8] [9].

To simulate numerically the failure of the material at high velocity impact, we applied the active-type kinetic model determining the growth of microdamages, which continuously change the properties of the material and induce the relaxation of stresses:

$$\frac{dV_f}{dt} = \begin{cases} 0, & \text{if } |P_c| \leq P^* \text{ or if } (P_c > P^* \text{ and } V_f = 0), \\ -\text{sign}(P_c)K_f(|P_c| - P^*)(V_2 + V_f), & \text{if } P_c < -P^* \text{ or if } (P_c > P^* \text{ and } V_f > 0) \end{cases} \quad (1)$$

Here $P^* = P_k V_1 / (V_f + V_1)$, and V_1 , V_2 , P_k , and K_f are material constants determined experimentally. The form of condition (1) was chosen according to the experimental data. We assume that there are the fracture areas of identical initial sizes in the material with the effective specific volume V_1 . Cracks or pores are formed and grow in these fracture areas when the tensile pressure exceeds a certain critical value P^* that decreases during the growth of microdamages. The constants in (1) were chosen by comparing the results of computations and experiments concerning the recording of a free surface velocity when a specimen was loaded by planar impulses of compression. The same set of constants is used to calculate both build-up and collapse of cracks and pores (depending on the sign of P_c).

The material model includes the equation of state of the Mie-Grüneisen type that represents pressure as a function of specific volume and specific internal energy, the deviatoric elastic constitutive relationships, the von Mises yield criterion taking into account temperature effects. The strength characteristics of the medium (shear modulus and dynamic yield strength) depend on temperature and the current level of damages.

$$G = G_0 K_T \left(1 + \frac{cP}{(1+\mu)^{1/3}} \right) \frac{V_3}{(V_f + V_3)}$$

$$\sigma = \begin{cases} \sigma_0 K_T \left(1 + \frac{cP}{(1+\mu)^{1/3}} \right) \left(1 - \frac{V_f}{V_4} \right), & \text{if } V_f \leq V_4 \\ 0, & \text{if } V_f > V_4 \end{cases} \quad (2)$$

$$K_T = \begin{cases} 1, & \text{if } T_0 \leq T \leq T_1 \\ \frac{T_m - T}{T_m - T_1}, & \text{if } T_1 < T < T_m \\ 0, & \text{if } T \geq T_m \end{cases}$$

Here T_m is the melting point of the substance, and c , V_3 , V_4 , and T_1 are the constants.

To simulate the brittle-like failure of the intermetallic material under high velocity impact, we developed the model for the possibility of failure above HEL in the shock wave and the sharp drop in the strength characteristics for the failure of materials [13]:

$$\sigma = \begin{cases} \sigma_0 P_f K_T \left(1 + \frac{cP}{(1+\mu)^{1/3}} \right) \left(1 - \frac{V_f}{V_4} \right), & \text{if } V_f < V_f^k \\ \sigma_f K_T, & \text{if } V_f^k \leq V_f < V_4 \\ 0, & \text{if } V_f \geq V_4 \end{cases}, \quad (3)$$

$$P_f = \begin{cases} 1, & \text{if } \sigma_{sh} < \sigma_{HEL} \\ P_f^k, & \text{if } \sigma_{sh} \geq \sigma_{HEL} \end{cases},$$

where σ_{sh} is the stress in the shock wave ($\sigma_{sh} < 0$ for compression), P_f^k , V_f^k , σ_f , σ_{HEL} are the constants.

The critical specific energy of shear deformations is used as a criterion of the erosion failure of the material that occurs in the region of intense interaction and deformation of contacting bodies. The current value of the specific energy of shear deformations is defined from the relationship

$$\rho \frac{dE_{sh}}{dt} = S_{ij} \varepsilon_{ij}$$

The critical value of the specific energy of shear deformations depends on the conditions of interactions and is a function of the initial impact velocity

$$E_{sh}^c = a_{sh} + b_{sh} v_0$$

where a_{sh} and b_{sh} are the constants. When $E_{sh} > E_{sh}^c$ in the computational cell near the contact boundaries, the cell is assumed to be damaged and the parameters in neighboring cells are corrected with regard for the principles of conservation laws.

Results and Discussion

We consider the interaction of a projectile with a finite thickness target. The problem is formulated using the Cartesian coordinate system with initial (at $t = 0$) and boundary conditions. The initial conditions are characterized by the absence of internal stresses, and the projectile moves toward the target with a velocity v_0 . There are no external loads on free surfaces of the interacting bodies, while the conditions of sliding are implemented on the contact surfaces between the projectile and the target. The finite-element relations used to solve the formulated problem are given in [14] [15].

Fig. 1 shows the computer images with a section of the projectile and composite target at the time of 60 μ s. The thickness of intermetallic Al_3Ti layer in this case was of 0.94 mm, the thickness of the layer of Ti-6Al-4V titanium alloy was of 0.23 mm. The computations demonstrate the fact that the MIL composite target withstands the impact loading.

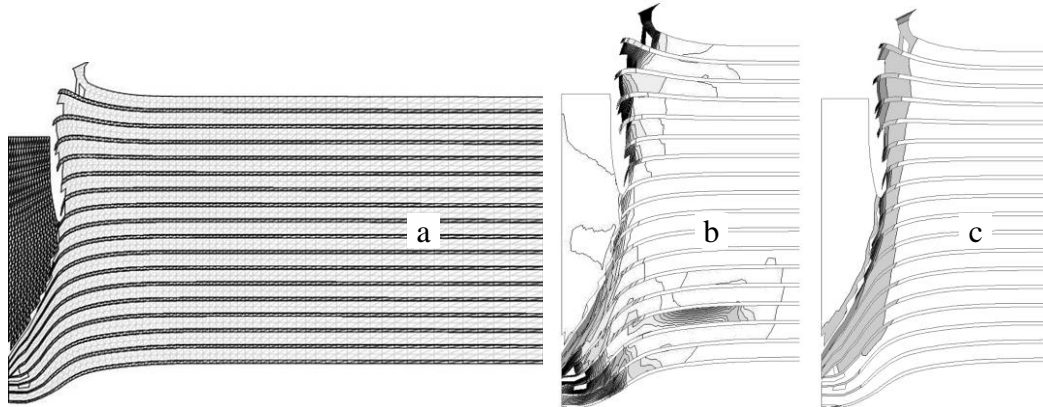


Figure 1: Computer images with a radial section of the projectile/target assembly (a), specific volume of microdamages (b) and specific shear deformation energy (c) at 60 μ s

The distribution of the damage and the deformation patterns are illustrated in Figs. 1b and 1c which show the section contours of the projectile and composite target, the contours and fields of the specific volume of microdamages (Fig. 1b) and the specific shear deformation energy (Fig. 1c). The low level of microdamages in the layers of titanium alloy shows the brittle damage stops distribution in the intermetallic layers.

Table 1: Results of simulations for target layers of different thicknesses

	Al_3Ti [mm]	Ti-6-4 [mm]	Areal density [g/cm ²]	Depth of penetration [mm]		Average velocity [m/s]	
				40 μ s	60 μ s	40 μ s	60 μ s
1	0.94	0.23	7.02	17.00	18.49	150	30
2	1.17	-	6.54	21.12	25.41	350	150
3	-	1.17	8.97	18.73	20.90	200	50
4	0.47	0.70	7.99	22.95	28.71	430	250
5	0.23	0.94	8.49	24.15	-	470	-
6	0.70	0.47	7.52	21.86	26.44	410	220
7	1.04	0.13	6.81	22.85	-	440	-

The Table 1 represents the results of simulations for target layers of different thicknesses. The results show that the depth of penetration depends on the thicknesses of intermetallic and titanium alloy layers. The MIL composite target withstands the impact loading for the 0.94 mm Al_3Ti / 0.23 mm Ti-6-4 (the ratio is about 4/1). In this case the intermetallic layer

provides the failure of the projectile and the metal layer stops the distribution of damage. In the other cases the perforation of the MIL composite target takes place. There is the same result for the uniform target made of either Al_3Ti (line 2 in the Table 1) or Ti-6-4 (3).

Conclusions

The results obtained demonstrate that destruction of the intermetallic layer is brittle compared to plastic failure of the metal layer. The computations have shown that the optimal composite target has a higher ballistic resistance in comparison with a uniform target either Al_3Ti or Ti-6-4. The optimum construction of the MIL composite should include a metal layer of sufficient thickness, which would stop the distribution of brittle damage. The results show that the depth of penetration depends on the thicknesses of intermetallic and titanium alloy layers. The composite target withstands the impact loading for the ratio about 4/1 (Al_3Ti / Ti-6-4).

Acknowledgments

This work was supported by the Russian Science Foundation (RSF), project no. 16-19-10264.

References

- [1] Vecchio, K. S. (2005) Synthetic multifunctional metallic-intermetallic laminate composites, *JOM*, **March**, 25-31.
- [2] Li, T., Jiang, F., Olevsky, E. A., Vecchio, K. S., and Meyers, M. A. (2007) Damage evolution in Ti6Al4V- Al_3Ti metal-intermetallic laminate composites, *Materials Science and Engineering A*, **443**, 1-15.
- [3] Price, R. D., Jiang, F., Kulin, R. M., and Vecchio, K. S. (2011) Effects of ductile phase volume fraction on the mechanical properties of Ti- Al_3Ti metal-intermetallic laminate (MIL), *Materials Science and Engineering A*, **528**, 3134-3146.
- [4] Cao, Y., Zhu, Sh., Guo, Ch., Vecchio, K. S., and Jiang, F. (2015) Numerical investigation of the ballistic performance of metal-intermetallic composites, *Appl. Compos. Mater.*, **22**, 437-456.
- [5] Patselov, A., Greenberg, B., Gladkovskii, S., Lavrikov, R., and Borodin, E. (2012) Layered metal-intermetallic composites in Ti-Al system: strength under static and dynamic load, *AASRI Procedia*, **3**, 107-112.
- [6] Zelepugin, S. A., Mali, V. I., Zelepugin, A. S., and Ilina, E. V. (2012) Failure of metallic-intermetallic laminate composites under dynamic loading, *Shock Compression of Condensed Matter. USA: American Institute of Physics*, **1426**, 1101-1104.
- [7] Zelepugin, S. A. and Zelepugin, A. S. (2015) Numerical simulation of multilayer composites failure under dynamic loading, *Applied Mechanics and Materials*, **756**, 408-413.
- [8] Zelepugin, S. A. and Zelepugin, A. S. (2014) Failure of multilayer composites under dynamic loading, E. Oñate, X. Oliver, A. Huerta eds. *Proceedings of the jointly organized 11th World Congress on Computational Mechanics, the 5th European Conference on Computational Mechanics and the 6th European Conference on Computational Fluid Dynamics (WCCM XI – ECCM V – ECCM VI 2014)*, Barcelona, Spain, July 20-25, 2014, **4**, 3926-3933.
- [9] Zelepugin, S. A. and Shpakov, S. S. (2009) Failure of metallic-intermetallic multilayered composite under high-velocity impact, *Mekhanika kompozitsionnykh materialov i konstruktsii*, **15**(3), 369-382.
- [10] Gerasimov, A. V. and Pashkov, S. V. (2013) Numerical simulation of the perforation of layered barriers. *Composites: Mechanics, Computations, and Applications. An International Journal*, **4** (2), 97-111.
- [11] Gerasimov, A. V., Dobritsa, D.B., Pashkov, S. V., and Khristenko, Yu. F. (2016) Theoretical and experimental study of a method for the protection of spacecraft from high-speed particles, *Cosmic Research*, **54**(2), 118-126.
- [12] Gerasimov, A. V. and Pashkov, S. V. (2015) Numerical simulation of fracture in thin-walled structures by a group of high-velocity elements, *Russian Aeronautics*, **58**(2), 137-144.
- [13] Gorelskii, V. A., Zelepugin, S. A., and Tolkachev, V. F. (2000) Experimental and numerical study of ceramics destruction by high-velocity impact, *Chemical Physics Reports*, **18**(10-11), 2211-2217.
- [14] Johnson, G. R. (2011) Numerical algorithms and material models for high-velocity impact computations, *International Journal of Impact Engineering*, **38**(6), 456-472.
- [15] Gorelski, V. A., Zelepugin, S. A., and Smolin, A. Yu. (1997) Effect of discretization in calculating three-dimensional problems of high-velocity impact by the finite-element method, *Computational Mathematics and Mathematical Physics*, **37**(6), 722-730.

Computational Methods in Architecture from theories to architectural design

Anna Irene Del Monaco¹

Dipartimento di Architettura e Progetto, Sapienza Università di Roma, ITALY.
anna.delmonaco@uniroma1.it

Abstract

Almost thirty years have passed since the mid-eighties, when the digital has burst into architectural design and production. From the initial experiments and theoretical enthusiasms, practiced in geographical, political and cultural different contexts where traditionally there was a strong relationship between design, technology, administrative institutions and industries, we have achieved some widespread and standardized results in software industry, processing, production, profession, spanning from experimental computational applications to Building Information Modeling. What will be the next conceptual steps and tools and in which specific fields it seems that the use of Computational Design be strengthened? This paper tries to reassemble the wires between the experimenters' generation, their earliest disciples strongly engaged with theoretical works, and the awake of the historians on the digital and the computation design in architecture.

Keywords: **Computation, Architecture, design.**

A retroactive awareness

Almost thirty years have passed since the mid-eighties, when the digital has burst into architectural design and production – after the personal computers spread at large scale – allowing today to read within a historical perspective the most significant past experiences and to detect issues anticipating how technology will continue to influence architecture in the near future. Since then, the use of computer technology in the design and in the building process of architecture has considerably grown.

The deanship (1988-2003) of Bernard Tschumi at Columbia University of New York, Graduate School of Architecture Planning Preservation signed a further phase of architectural postmodern culture, based on technological innovations, – digital design (research on form, style), widely disseminated by architectural journals and magazines as *Architectural Design*, and computing sciences applied architecture (BIM, prototyping, 3D printing, form-finding, parametricism) – competing with similar experiences already started at MIT (Negroponte, Media Lab 1985) and Harvard Laboratory for Computer Graphics and Spatial Analysis (Eric Teicholz, Jack Dangermond founder of ESRI), established in 1965.

From the initial experiments and theoretical enthusiasms, practiced in geographical, political and cultural different contexts where there was traditionally a strong relationship between design, technology, administrative institutions and industries, we have achieved some widespread and standardized results in software industry, processing, production, profession, spanning from experimental computational applications to Building Information Modeling.

In this new field of research – at that time – a giant as Frank Gehry certainly made digital design as the fundamental tool for linguistic exploration and for research on the construction methods to produce architecture. Gehry has played a decisive role for several decades positioning his practice – hardly imitable – between the development of digital design and the spread of software and assessment systems with the foundation of Gehry Technologies, at first getting patents and then selling the company which keeps still his name.

During the last fifteen years, after an initial phase of individual engagements (scholars, groups and institutions), the research results were spread establishing teaching courses and programs, with a special emphasis provided in cutting-edge schools: Digital Research Lab at Architectural Association London, Sci-Arch Los Angeles, Angewandte Wien, MIT Boston, Columbia New York, Cornell Ithaca, ETH Zurich. But recently courses focused on topics as Algorithmic Architecture and scripting, Computer-Aided Design and Computer-Aided Manufacturing (CAD/CAM) technologies emerged also in other architectural schools not traditionally engaged in these researches (Stuttgart, Chalmers, Cardiff, etc) with the intention to expand or reinforce their academic programs.

However, after an in depth analysis related to the institutional and industrial relationships and academic networks you would understand how decisive were the linkages and nexus between institutions, market, academia, simplifying the discourse between clients and researchers, politics and visionaries. For example, how important the relationship between Ian McHarg and Jack Dangermond (founder of ESRI and McHarg's former student), was for the work on “map overlay” concept in landscape discipline and the development of GIS. How consistent was the work of Harvard Laboratory for Computer Graphics with the American Census Bureau and the geographers' digital research at territorial scale. How strategic was for the American architectural profession to instruct young architects in the use of digital 3D software to standardise the drawing production for the construction industry passing by the organization of the drafting, designing, making physical models, drawing work in the architectural office and the acquisition of a specific software license. Therefore, also the strong commitment of Nicholas Negroponte, Leon Groisser and Jerome Wiesner (MIT president) with the defence research, the National Science Foundation and private corporations. In a seven year fund raising period, infact, the Negroponte's Architecture Machine Group collected 40 million dollars – in which 40 corporations pledged their pre-commercial research budget – with the support Dean Lawrence Anderson who sought to challenge traditional architectural pedagogy: the Beaux-Arts teaching method exercised what he called a “residual influence [that] remains as an incubus that dampens our enthusiasm for any panacea”¹. Not to talk about Buckminster Fuller and his geodesic domes and the interest on them of the American Army.

Imitation-Modeling

What will be the next conceptual steps and tools and in which specific field the use of Computational Modeling in architecture will be strengthened?

From the point of view of Italian Architecture schools their an increased challenge to enhance the integration of computation design in their curricula since enormous revolution occurred establishing the mandatory use² of Building Information Modeling, within five years, in every public tender.

It could be useful to recall that in 2005 the School of Architecture “Ludovico Quaroni” of Sapienza University of Rome organized a Symposium³ on the education of architects with a special panel titled “Digital innovation in architecture”, chaired by Antonino Saggio, which hosted Patrik Schumaker, Vasili Stroumpakos, Manuela Gatto, Mark Goulthorpe, Ed Keller, Karl Chu, Claudio D'Amato, Lucio Barbera and Giuseppe Longo [2] a mathematician from the CNRS Ecole Normale Supérieure, Paris. Longo had been invited to offer an overall comment to the researches of the architects making extensive use of computing design. His perception was almost tough after participating to the panel presentation of the experiments presented, concerning biology and computing processes 'applied' in/to architecture during the Roman symposium – especially the “Responsive Environments” by Schumacher and the “Human Genome Project” by Karl Chu. Longo observed that: “In contrast to the problem of intelligibility in the sciences of nature, for you

¹. See Radical Pedagogies: <http://radical-pedagogies.com/search-cases/a13-architecture-machine-group-media-lab-massachusetts-institute-technology-mit/>

². New Public Works Code, active since 2019.

³. Dean Lucio Barbera organized the Symposium “Becoming Architect in the XXI Century” from January 20th to 22th 2005. Proceedings, Anna Irene Del Monaco (edited by) *Becoming Architect in the XXI Century*, Casa Editrice La Sapienza 2006.

[architects], it seems to me, that the role of the mathematical instrument – now information technology – is that of *aid to invention*, in this very complex game between science and artistic fantasy that is precisely the architectural design of a 'thing that is not there'. And 'this' determines a radical difference from the point of view of the relationship with information technology and its tools, in particular, and at the same time the problem is even richer in its internal logic deserving a reflection. The extraordinary novelty we are confronted with is the fact that today we have a machine that is the result of a very complex historical-evolutionary path. This machine was not there, in the same way that 200 million years ago there were no mammals on the face of the earth.” Longo was making reference in his discourse to the Turing machine, a mathematical abstract concept, an indispensable clarification to architects referred to the non “scientific” (intelligibility) approach in their simulation (or imitation) “games”. In particular Longo in one of his paper [3] had already discussed this arguments taking of the double pendulum digital simulations. If you click “restart” relaunching a virtual pendulum – as Schumacher's did in his “Responsive environments” presentation – with the same initial data it tends to cover all the possible trajectories (a chaos aspect) “but with a real physical pendulum, not a virtual one, in a dynamic (non-linear) system – because of the thermal fluctuation, the density of trajectories –, “once reinitialized, the systems never take the same trajectory. [...] Specifically, we will see that the digital machine proposes causal structures and the breaking of symmetry which generate them as being the central structures of the intelligibility of nature. This will enable to point out a distinction between 'imitation' and 'modeling' in terms of simulation or formalization, and therefore enable to highlight the limits and the potentialities of digital simulation. Such a machine is the culminating point of a very specific process which begins with language, but which is mainly influenced by the birth and development of the alphabet: the digital machine is at first an alphabetic machine, and then a logical and formal one”. This argument would deserve a consistent and wide discussion in architecture, especially the distinction between imitation and modeling – imitation resemble causality, Longo claims – and which is, more or less the distinction that some scholars as Achim Menges [4] raise quoting Rob Howard. As Menges (and Howard) belong to a ideal group of researchers – “a research tradition” – which understand the necessity to distinguish the process of designing to build (o “fabricate”) towards the necessity to design for in search for authoriality [5, 6], to which the Zaha Hadid and Coop Himmelb(l)au approach to architecture belong: “In parallel to the development of Computer Aided Design, which is characterized by the transfer of long-established, representational design techniques into the digital realm, research into generative Computational Design has been conducted for many decades. In the 1960s design programs such as GRASP (generation of random access site plans) developed by Eric Teicholz at the Harvard Laboratory for Computer Graphics, explored ways of generating rather than drawing design solutions and evaluating their performance (Howard 1998).”

Going back to Longo's comment on the Turing machine and recalling its conceptual newness, it is interesting to read that “in short, it is an invention which is both extraordinary and contingent to our culture, which is marked by the birth of the alphabet, of Cartesian rationality, of Fregean logic, of Hilbertian formalism” [...] The 'principles' which are inherent to Physics (modern Physics), an interval and the inevitable variation, below the threshold of measurement, suffices to vey quickly produce a different evolution. The analysis of the equations within the continuum leads to an understanding of the random aspect of chaos, whereas computational imitation makes it disappear completely, but the discrete nature of its data types.”

Longo's assumptions remind us the idea of Antoine Picon when asserts that it is the information based society which invented the computer and not *viceversa* [7] – showing in his lectures a picture of the Prudencial Company (financial, insurance, retirement, investment company) at the turn of the nineteen century – and that the concept of diagram in architecture has failed – mentioning the lost promises of working on fluxes diagrams by UN Studio –, being in the end more the development of a Beaux Arts approach than something different and that, looking at a different aspect, the Building Information Modeling system imposes the idea of “who is allowed to do what and not the implementation of the idea of a fluid world”.

Not just a happy accident !

The way of approaching 'imitation' and 'modeling' put the work of an entire generation of academic-architects under a different perspective and recalls the objective meaning of the term "Computer Aided Architectural Design..." which literary could mean "aid to invention" and provide some advice to follow especially when we teach and make analogies, comparison with the natural, physical and abstract realities.

The generation of Frank Gehry, Peter Eisenman, Zaha Hadid, Coop Himmelblau, Future System, although starting the use of computer in their fifties or sixties – and in many cases with them having no direct personal skills in the use of computer –, made a significative use of computer aided design to enhance the possibility to draw, design, conceive, assess and build their architectures. Albeit it was also a generation intersecting the computational early tools with the Deconstructivist Architecture discourse, as Philip Johnson implicitly demonstrated with his exhibition at MoMA of 1988. This is the case of a generation of architects "that told to the computer what they wanted to have!", affirms Greg Lynn in a conference at the Architectural Association of 2013⁴; the results was not "just a happy accident", like often happen to unconscious native digital students which act out of cultivated awareness – Lynn continues. It is the case of a generation of architects which found the way to transfer the information of their drawings into to the frame space structures system in order to build their projects. Each of them found in digital architecture a way to "complete" their intellectual impulse toward construction. Eisenman, following his theoretical impulse, did not disdain "the idea of genetic based architecture" and fed up a generation of younger practitioner and thinkers (included Greg Lynn), and built the uncompleted Santiago de Compostela project. Hadid (a graduate in mathematics before architecture) wished to escape the condemnation of being a paper architect. Gehry got the artistic intuition (the most advanced intellectual stage) to select the right tools for his formal research and was able to make business with them too.

Archaeology or pre-consciousness?

It is highly significant that Greg Lynn, who has been a young collaborator of Eisenman and Gehry during the eighties and nineties, and one of the first hour prophets of morphogenetic design (see his Embryologic House Project), has recently written a book entitled *Archaeology of the Digital* linked to an exhibition at the Canadian Center for Architecture which he has curated with a retrospective programmatic understanding – and that leaves us very much thinking if we "pre-consciously" remember the early work by Peter Eisenman "Cities of Artificial Excavation" although Eisenman commented that archaeological is better than antiquarian. Lynn can be considered together with Hani Rashid and Lise Anne Couture, Ben van Berkel and Caroline Bos, Farshid Moussavi and Alejandro Zaera Polo, Jesse Reiser and Nanako Umomoto, Patrik Schumacher, a generation of academics strongly engaged with digital architecture which have made of that domain their avant-garde language, intellectual tool and generational strategy to enter the architectural business (academic and professional) which Antonino Saggio with the IT Revolution in Architecture and other books [8] had been monitoring in the last twenty years for the Italian and non-Italian scholars and readers. Many of these architects delivered an early or late book manifesto. In particular Patrik Schumacher [5, 6] had been ambitiously theorizing in 2010 with *The Autopoiesis of Architecture*, Farshid Moussavi wrote the book *The Function of Ornament* [9] in 2006, and van Berkel and Bos published in 1999 *Move (3 volumes)* [10].

In *Archaeology of the Digital* Lynn analyses four architectures of the eighties classifying them by the following categories: 1 Design using procedural processing; 2. Transforming robotic structures; 3. Structural optimization and expression; 4. Optimization of the digital medium. Besides the book the exhibition at CCA curated by Lynn in Canada included additional drawings, physical models and the interviews to the four authors delving into "the genesis and establishment of digital

⁴ Greg Lynn - Conversation about Digital Archaeology, Architectural Association 2013-10-10; <https://www.youtube.com/watch?v=U2ujuFMYkzA&t=4077s>

tools for design conceptualization, visualization, and production at the end of the 1980s and the beginning of the 1990s. Conceived as an object-based investigation of four pivotal projects that established distinct directions in architecture's use of digital tools, the book highlights the dialogue between computer sciences, architecture and engineering that was at the core of these experiments". The four architectures that Greg Lynn selected are discussed in his book considered also the specific software and methodology, drawing techniques used for their design and eventual construction, probably in search for a "pre-conscious origin of the digital", as Peter Eisenman affirmed in the interview released for the publication and for the exhibition [11].

1) The Peter Eisenman's Biozentrum (1987) (Figure 1) to which Lynn worked as a young architect, was drew by the software Form Z integrating the digital drawings by a craft assessment system (layer copies of the variation of geometries superimposed, produced and mailed by Fed Ex every day from a laboratory located elsewhere). 2) The Chuck Hoberman's Expanding Sphere (1992) (Figure 2), in which the author used his own scripts hand-written – he owns 18 patents. After graduation Hoberman worked at Honeybee robotic and got acquainted with transformable objects. The software Hoberman used was AutoVIZ, a former language from Autocad. Hoberman had been the first in architecture to use the CnC manufacturing and 3d printing. The aluminium components of the Expanding Sphere were CnC cuts and the drawing presented, rarely at that time, by computer hidden line. 3) The Shoji Yoh's unbuilt Odawara Municipal Sports Complex (1990–91) (Figure 3). Shoji came from manufacturing and is the structural engineer that Toyo Ito collaborated with for Sendai Mediatheque. The project uses space frame structure, a parametric construction technique that later on become a standard. It is the only project built by the group with the use of the truss. 4) The Lewis Residence by Frank Gehry (Figure 4), Lyndhurst, Ohio (1989-1995). For this project SOM did the Autodesk 3D model and structural analysis. From that experience on Gehry decided he wanted to manage and control the 3D assessment phase and looked for the right tools to succeed. The physical wooden model was completed by the wax/velvet cloth and pin bars. Then the house was redesigned for further exhibition and for the CCA's exhibition by the CATIA's inelegant spline surfaces. After starting using CATIA, Lynn affirms, Gehry started sketching like the spline surface. CATIA models had to be completely re-drawn in Autocad, but there was a fidelity in the measure and geometries. Referring to the time in which the four projects were conceived Greg Lynn comments that after "five years later it was a total different story". Before there were a lot of correspondences back and forth from the two kind of techniques (digital and analogical): "Either these techniques were abandoned or diverged. At that time the speed of computer was slow. Analogical and digital methods were equal in speed, so that producing physical models and sheet analysis was sometimes more speed than a computer model."

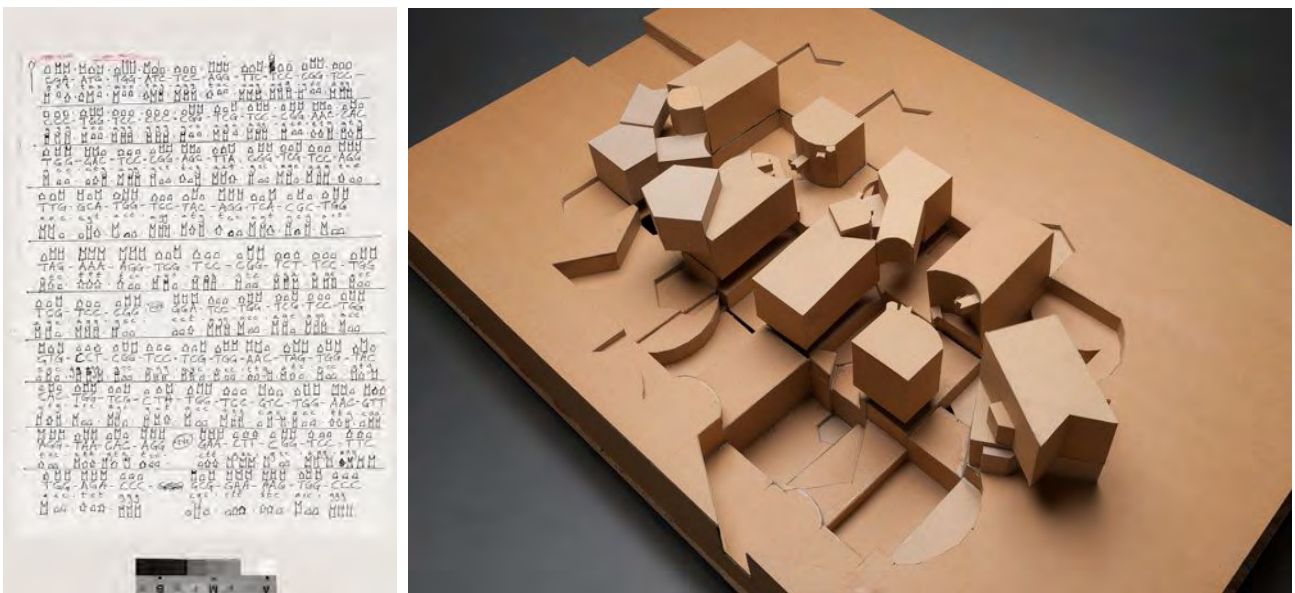


Figure 1. Peter Eisenman's Biozentrum (1987)

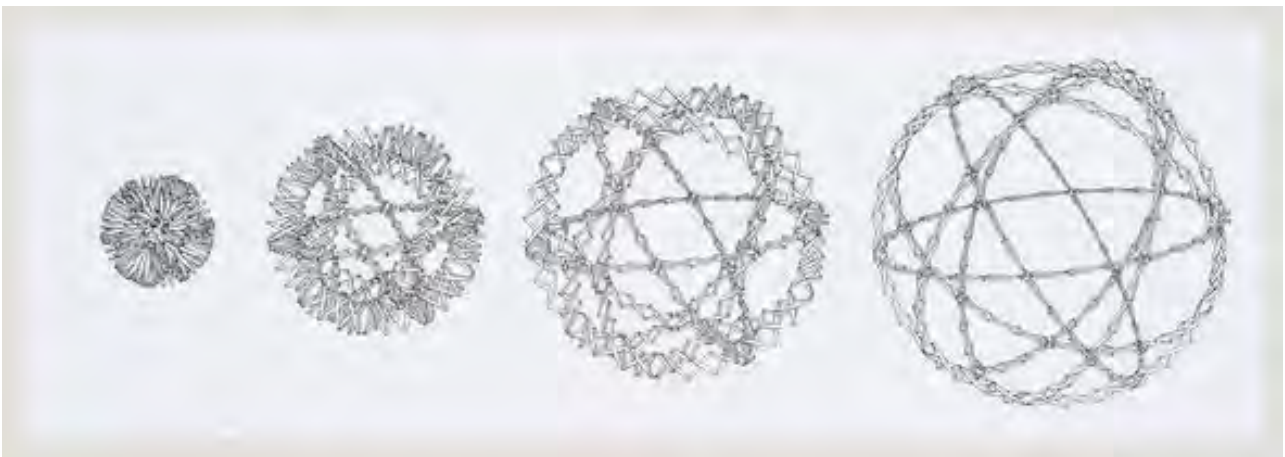
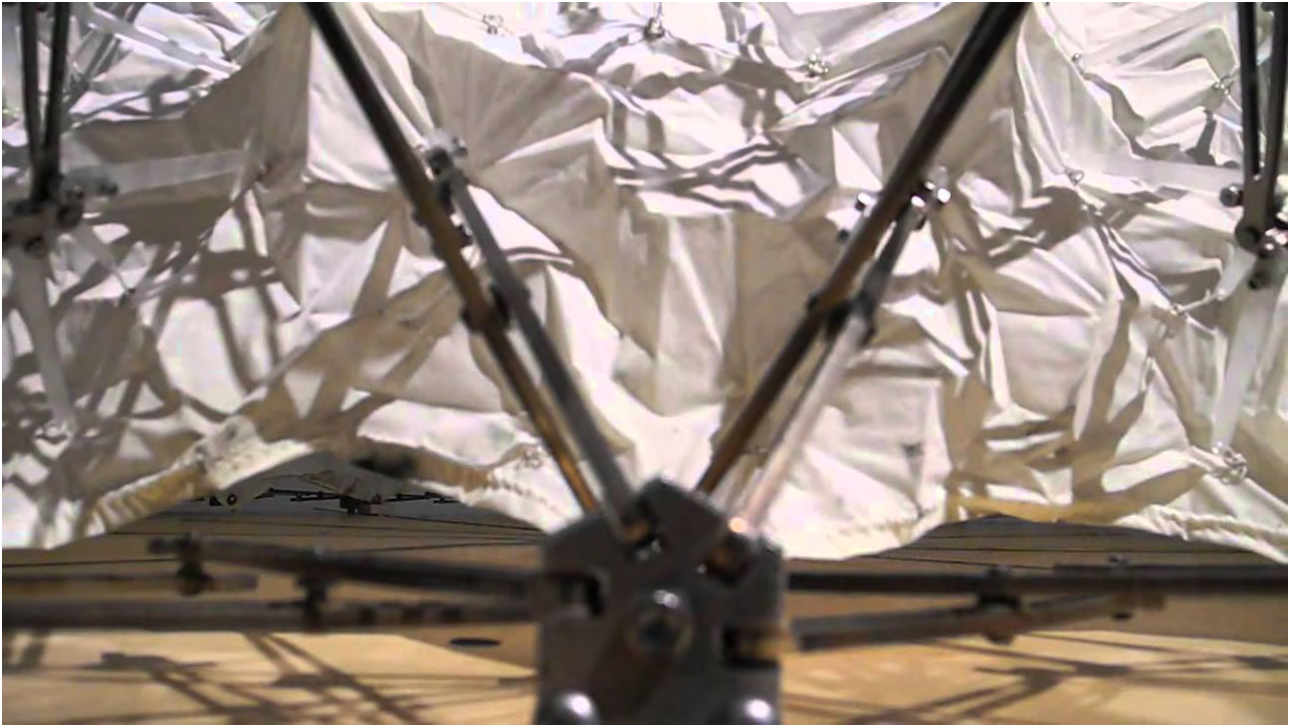


Figure 2. The Chuck Hoberman's Expanding Sphere (1992)

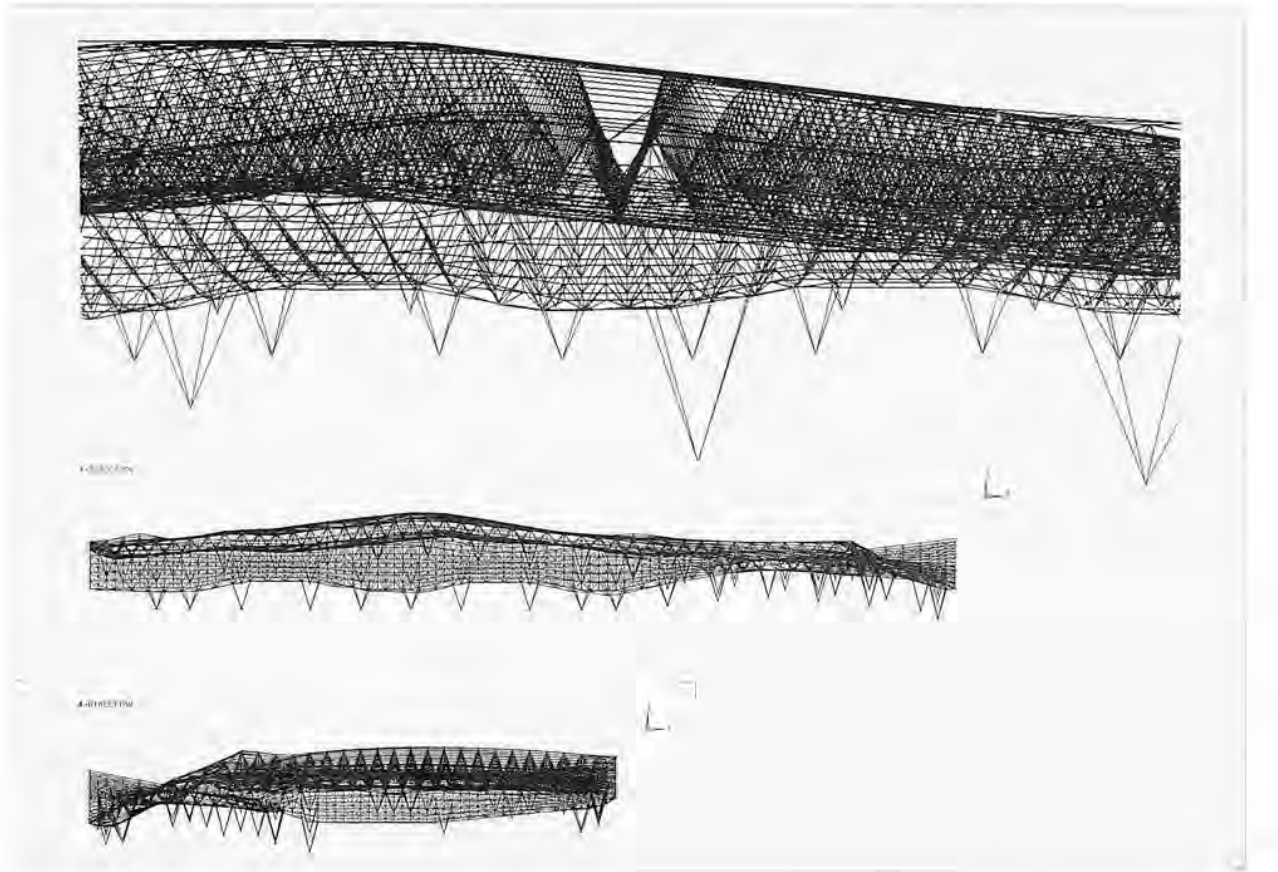


Figure 3. Shoei Yoh's unbuilt Odawara Municipal Sports Complex (1990–91)



**Figure 4. The Lewis Residence by Frank Gehry, Lyndhurst, Ohio (1989-1995).
The pin bars, plexiglass and wax/velvet cloth model and the CATIA spline drawing.**

The awake of historians

The fact that the generation of “young” enthusiasts experimenters of the nineties in the last decade is focused on theoretical works on Archaeology (Greg Lynn, self-consciousness, quoting Eisenman again) or Autopoiesis (Parik Schumacher) and that some historians of architecture have begun to reflect on the phenomenon of digital architecture – some of which establishing that it is in a phase of “Second Turn”– tells us many things. Recently, Mario Carpo (professor of Renaissance and history of cultural technologies) and Antoine Picon (professor of History of Architecture and technology) have been doing a significant work in conducting their researches on digital architecture from an architectural historian point of view and providing focus on different aspects, at least producing a counterpart effect with respect to the mainstream culture.

However some years ago also James Ackerman, a great historian of Renaissance architecture and Palladio's scholar, writing his contribute after fifty years to the proceedings of the Triennale Milan 1951 "La divina proporzione" (with Le Corbusier and Wittkower as chairman) introduced significant arguments. He observes that after the Modulor the "golden section" disappears from the architectural discourse and, at the end of the twentieth century, proportional systems based on plane geometry and straight spaces begin to "give way to the fluid forms of the architects-engineers contemporary of Le Corbusier": Pier Luigi Nervi, Edoardo Torroja, Felix Candela, Buckminster Fuller, Heinz Eisler. The arguments proposed by Ackerman confirm furthermore the reason why Mario Carpo and Peter Eisenman are so involved in the digital architecture discourse. In particular, the passage in which, starting from the definition of Wittkower of Renaissance architecture (Andrea Palladio, Sylvio Belli) as architecture based on number, seems to be significant, Ackerman articulates the reasoning inside and outside the scope of the 1951 conference: "But in the field of musical composition the limits of the simple Pythagorean consonances had been surpassed for more than a century. From the beginning of the 15th century, polyphonic music composers had introduced what theorists followed to define dissonances (as was shown by Robin Evans in the brilliant essay "The Projective Cast: Architecture and its Three Geometries"; it was to demolish the belief that 'ideal forms are ... in themselves, ideally beautiful')." [12] Then, in addition to addressing the work of Frank Gehry, whose "buildings with an evolved formal vocabulary [...] have made the definition of proportion" more complex, and to mention the experiments of Greg Lynn, who promoted a type of planning based on the concept of calculation and topology, the American historian highlights the theoretical and design work of Preston Scott Cohen, in particular, the book *Contested symmetries and other predicaments in architecture*, which reopens to a line of research that supports the survival of the idea of proportion.

In 2011 Mario Carpo published *The Alphabet and the Algorithm* [13] for the Writing Architecture series, a project of Any Corporation, Cynthia Davidson editor and brilliant wife of Peter Eisenman, who certainly appreciated Carpo's highly-qualified consistent engagement with both Renaissance and Digital Design. The synopsis of Carpo's book clarify what in our discourse get us reconnected to the Long's arguments on "imitation" and "modeling": "In *The Alphabet and the Algorithm*, Mario Carpo points to one key practice of modernity: the making of identical copies. Carpo highlights two examples of identity crucial to the shaping of architectural modernity: in the fifteenth century, Leon Battista Alberti's invention of architectural design, according to which a building is an identical copy of the architect's design; and, in the nineteenth and twentieth centuries, the mass production of identical copies from mechanical master models, matrixes, imprints, or molds. The modern power of the identical – Carpo writes – came to an end with the rise of digital technologies. Everything digital is variable. In architecture, this means the end of notational limitations, of mechanical standardization, and of the Albertian, authorial way of building by design".

A significant advice already came from ACADIA members. CAD technology, had already be introduced in the postwar era and adopted in architectural practice beginning in the 1970s, affirmed Yehuda E. Kalay [14] – founding member, and past president, of ACADIA (Association for Computer Aided Design in Architecture), and former co-Editor-in-Chief of Automation in Construction – already in 2004, "CAD [goes] beyond the improvements in drafting, modeling, and rendering for which it is commonly used. Computer-aided architectural design (CAAD) is capable of modeling and manipulating objects (not merely their graphical representations), reasoning about and predicting performance of design solutions, generating new design solutions through algorithmic and other methods, managing vast amounts of information, and taking advantage of opportunities offered by the Internet for collaboration across time and space and for design of the virtual 'space' of the Internet itself."

Moreover Carpo came out recently with another book, *The Second Digital Turn*, whose synopsis shows the further conceptual steps proposed by the Italian historian: "Almost a generation ago, the early software for computer aided design and manufacturing (CAD/CAM) spawned a style of smooth and curving lines and surfaces that gave visible form to the first digital age. [...] this is because the design professions are now coming to terms with a new kind of digital tools they have

adopted – no longer tools for making but tools for thinking.” In March 2018 the author presented the main contents of his late book in a lecture [15] to a non-architects audience (Google's employs) at University College London and used the metaphor of fishes, potatoes and dogs. “Frank Gehry started with a fish!” the historian explained, talking of the fish streamline sculpture by Gehry built along the Barcelona waterfront and of the “spline” working methods of CATIA which impressed the streamline also to other Gehry's projects, from the Bilbao Museum up to the Foundation Cartier. Carpo discussed also about the spread of softwares as Rhino, Maia, Form Z, cheaper than CATIA, in the teaching courses and in the profession and the fact that today computers are faster and cheaper, and that big data and environment are emerging issues to be considered. So that, according to him, the style produced by the software used in the most recent years shifted from fishy (Gehry's CATIA) to doggy (Heydar Aliyev Center, Baku by Zaha Hadid Architects), a less sophisticate streamline developed by a very simple algorithm. Unlike “fishes”, “potatoes” and “dogs” belong to the free-forms category and have no sophisticated mathematics embedded inside their geometries. The reason why Carpo consider the current condition a “second turn” depends on the fact that for architecture the technical bottleneck is always determinant. Architecture is always at the mercy of the tools that make it happens, and his authoritative discourse upholds what I tried to demonstrate in a comparative study between composition in music and architecture [16] too.

The Italian born historian was particularly bright in clarifying during his London talk that architecture is a liberal-art and not a scaffolding art and that even though drawing is a notation system used to eventually build “architects can enter the spirit of the game better than other professionals, since they uses simple software and have the building phase as a feedback loop, although computing is not their game but your (google's people) game...”

Geometries, families of solutions, fabrication

However geometry is at the base of every discourse. Also Mario Carpo recalls that his main arguments are related to the shift of quantification from algebra to calculus. Euclidean geometry, a quantification and proportional system, was typical of Greek and Roman architecture, while numbers and arithmetic are the proportional system specific of the neoclassical architecture. Then, Carpo remind us that with Pierre Étienne Bézier there is the culmination of western mathematic, after the work of Descartes, Leibniz, Newton which were able to notate conics but not complex geometries like Gehry's “fishes” – Jean Louis Cohen has started to study Frank Gehry from an architectural historian point of view [17]. Other relevant contributes to the question of contemporary geometry, fabrication and construction are elaborated by ETH Zürich research teams as he Block Research Group (BRG), also established at ETH Zürich hosts led by Philippe Block and Tom Van Mele, whose experiments were hosted at the Venice Biennale 2016 (Figure 5). Their research focuses on the analysis of masonry structures, graphical analysis and design methods, computational form finding and structural design, discrete element assemblies, and fabrication and construction technologies, geometry based approach. The Block Research Group of ETH Zürich works are somehow on the track already traced by Pier Luigi Nervi and Sergio Musmeci (Figure 6). Another significant contribute from the ETH Zürich is the book *Architectural Geometry* [18] edited by Helmut Pottmann, Andreas Asperl, Michael Hofer, Axel Kilian and published by Bentley Institute Press – Bentley is the company releasing Microstation, a software used, for example by Forster and Partners. Shajay Bhooshan, graduate from AA London and Lead Designer, Zaha Hadid Architects Computation and design group, is completing his PhD with Block Research Group office and has been a bridge for experimenting collaboration between ETH and ZHA office. There is another interesting contribute coming from Southern USA: Ronald Rael and Virginia San Fratello and their practice Emerging Objects (Figure 7) together with their teaching activity at UC Berkeley and Arizona University demonstrate the impact of computation on a different kind of architectural and construction tradition and materials as clay and hearth [19, 20].

Ali Farzaneh and Elif Erdine chaired an interesting conference [21] at the Architectural Association London within the Symposium “Advancements in Design Computation” on February 2016. They

invited Antoine Picon (GSD Harvard), Achim Menges (Stuttgart University), Francis Aish (Foster and Partners). Picon focused on the importance of the concept of “materiality” – not “materials” – to understand digital architecture, in the sense of Condillac's treatise on the sensations (senses and matter), and of Boullée's relations between human subjects and the material worlds [22, 23]. Achim Menges's [24] – raised within the Emergent Technologies group at the AA London, developed and inspired by the morpho-ecologies of Michael Weinstock – design approach is also focused on “materiality” to what extent you can push the limit of design toward material systems. Menges reveals to be among the most promising experimenter in the future, for the philological clearness in the research of coherence between the use of the material, the structural geometrical conception (derived from the material morphogenetic character) and its fabrication (Figure 8). His work is highly inspired by nature, biology, material behaviour (spider works: pneumatic inflations; lobster shells' cavity and fabrics) and by the possibility to synthesize the capability of a machine to imitate the “modelling” not the form of a spider or a lobster, etc. (Figure 9).



Figure 5. Block Research Group. Above: Construction prototype for ultra-thin concrete roof; Down: vaulted form made up from two layers of a new type of compressed earth blocks, Durabrick, developed by the LafargeHolcim Foundation for sustainable construction.

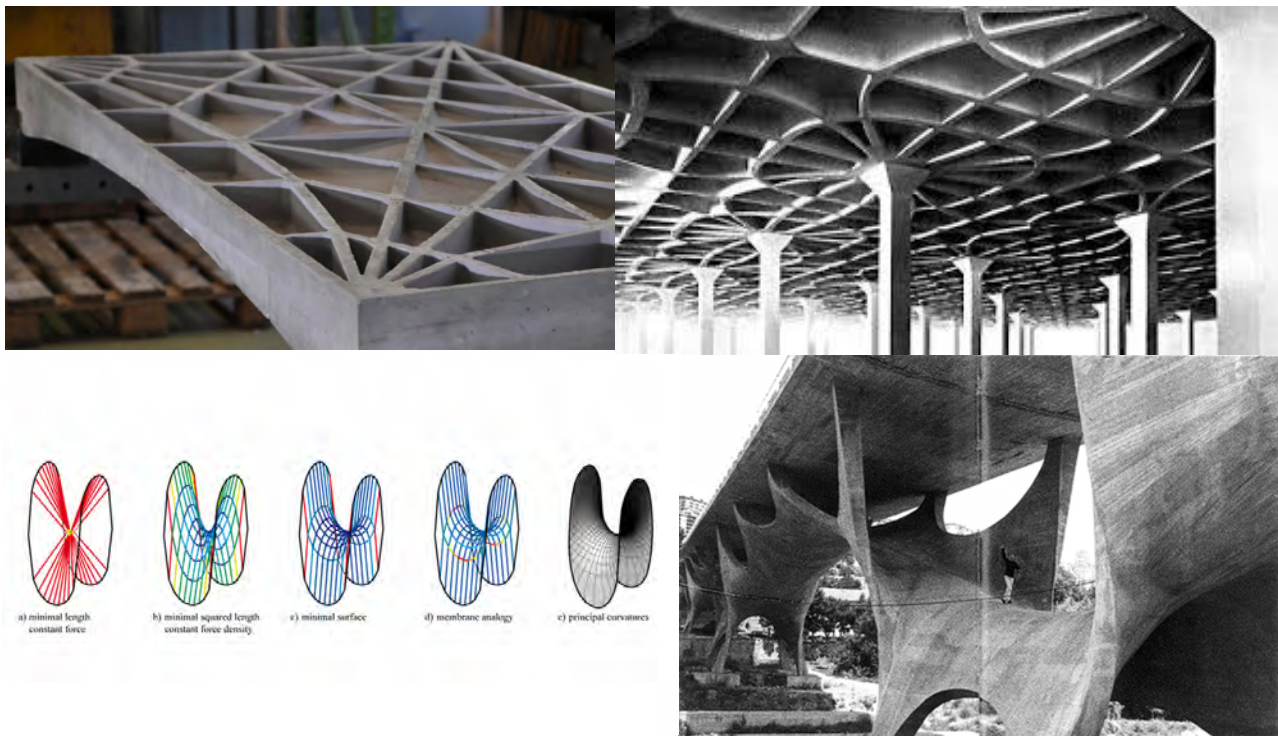


Figure 6.up-left: 3D-printed floor system by Block Research Group; up-right: Industrial warehouse “Lanificio Gatti” (1951-1953) by Pieri Luigi Nervi, structural concrete.

Down-left: Comparing form finding methods by Block Research Group; down-right Ponte sul Basento (1971-1976), by Sergio Musmeci.

In the same conference after the advanced experiments by Menges, Francis Aish from Foster & Partners, presented a number of projects revealing the other side of computation potentiality applied at the top level professional world: to accelerate the design process. In particular the Dresden Railway station project in 1997 and the Beijing New Terminal Airport for the 2008 Olympics are a demonstration of how to shorten the time during the design decision. In Dresden's project the canopy system and form-finding system were integrated arriving to a solution in less than one hour. The roof of Beijing Airport represents a low-tech intense project: 3,5 years to design and built the entire airport and 1,5 years to complete the roof by a space frame structure system with 5,000 workers (Figure 10).

To conclude, Picon and Carpo have a similar opinion toward the impact of Artificial Intelligence in architecture. The French historian says: “in twenty years from now we will have much more sophisticated software that we will enable to produce a lot of solutions (families of solutions) [25] and there will be a problem of choice, which is also one of the fundamental way to understand information [21]. The Italian historian maintains: “computational design is already a fascinating testing ground for Artificial Intelligence since the stuff we do are simple and cheap and software we use are elementary and we do physical things, so the feedback loop, the verification stage, is faster probably than in other professions.” [15]

References

- [1] Carpo M. (2017), *The Second Digital Turn: Design Beyond Intelligence*, The MIT Press 2017.
- [2] Longo G., *Comment and Reflections*, in Anna Irene Del Monaco (edited by) *Becoming Architect in the XXI Century*, Casa Editrice La Sapienza 2006, p. 256.
- [3] Longo G. (2009), *Critique of Computational Reason in the Natural Sciences*, In *Fundamental Concepts in Computer Science* (E. Gelenbe and J.-P. Kahane, eds.), Imperial College Press/World Scientific, 2009, Text originally written in Italian as *Lezione Galileana*, Pisa, 25 ottobre 2006. A French version is also downloadable: <http://www.di.ens.fr/users/longo>
- [4] Menges A., *Integrative Design Computation. Integrating material behaviour and robotic manufacturing processes*

- in computational design for performative wood constructions, Acadia Proceedings 2011. Menges quoted Howard, R. 1998. *Computing in construction: Pioneers and the future*. Oxford/ Woburn: Butterworth-Heinemann.
- Lynn G. (ed.) (2013), *Archaeology of the Digital*, Sternberg-press/Canadian Centre for Architecture.
- [5] Schumacher, P. (2017), *From Typology to Topology: Social, Spatial, and Structural*, Architectural Journal, No. 590, Source journal for Chinese scientific and technical papers and citations; Sponsor: The Architectural Society of China, Chief editor: Cui Kai.
- [6] Schumacher, P. (2010), *The Autopoiesis of Architecture, Volume 1-2, A New Framework for Architecture*, John Wiley & Sons.
- [7] Antoine Picon, Lecture "Digital Culture in Architecture", Harvard GSD, <https://www.youtube.com/watch?v=yfTpWttoo6g>
- [8] Saggio N. (2007), *Introduzione alla rivoluzione informatica in Architettura*, Carrocci
- [9] Moussavi F., Kubo M. (2006), *The Function of Ornament*, Actar.
- [10] van Berkel B., Bos C. (1999), *Move (3 volumes)*, Goose Press.
- [11] The Foundations of Digital Architecture: Peter Eisenman, CCA Channel, May 2013. <https://www.youtube.com/watch?v=hKCrepGoiX4>
- [12] Ackerman J. (2007), *Ricordi della Nona Triennale De Divina Proporzione*, in Anna Cimoli, Fulvio Irace, *La divina proporzione. Triennale 1951*, Electa, p. 33.
- [13] Carpo M. (2011), *The Alphabet and the Algorithm*, MIT Press.
- [14] Mario Carpo: The Second Digital Turn | Talks at Google, 21 March 2018.
- [15] Kalay, Y.E. (2004), *Architecture's New Media, Principles, Theories, and Methods of Computer-Aided Design*, MIT Press.
- [16] Del Monaco A.I. (2017), *Osservazioni sulle Corrispondenza fra la composizione in Musica e in Architettura*, Nuova Cultura.
- [17] Cohen, J.L. (2003), Frank Gehry, Architect, Guggenheim Museum Publications.
- [18] Helmut Pottmann, Andreas Asperl, Michael Hofer, Axel Kilian (2017), *Architectural Geometry*, Bentley Institute Press.
- [19] Rael R. (2009), *Earth Architecture*, Princeton Architectural Press.
- [20] Rael R., San Fratello V. (2018), *Printing Architecture. Innovative recipes for 3D printing*, Princeton Architectural Press.
- [21] Ali Farzaneh & Elif Erdine, Symposium, *Advancements in Design Computation*, Architectural Association; <https://www.aaschool.ac.uk/VIDEO/lecture.php?ID=3376>
- [22] Picon A. (2010), *Digital Culture in Architecture: An Introduction for the Design Professions*, Birkhauser.
- [23] Picon A. (2013), *Ornament: The Politics of Architecture and Subjectivity*, "Architectural Design" May 2013.
- [24] Menges, A. (2010), *Form Generation and Materialization at the Transition from Computer-aided to Computational Design*, "Detail" (English Edition), Vol. 2010 No. 04, pp. 330-335.
- [25] Leyton M. (1992), *Symmetry, Causality, Mind*, MIT Press.



Figure 7. Ronald Rael and Virginia San Fratello, Emerging Objects: “Cabin of Curiosities” (4.500 ceramic tiles) Oakland, 2018; “Bloom Pavillion”, 3D print cement (This structural system was inspired by the *Iglesia Cristo Obrero* of Uruguayan architect and engineer Eladio Dieste and Jefferson’s serpentine brick walls at the University of Virginia, while the precise form was inspired by Richard Serra’s *Torqued Ellipse*.)



Figure 8. Achim Menges. Above: Elytra. Filament Pavilion (Robotic Pavilion) for V&A Museum, London; Down: Component Membrane AA London, 2008.

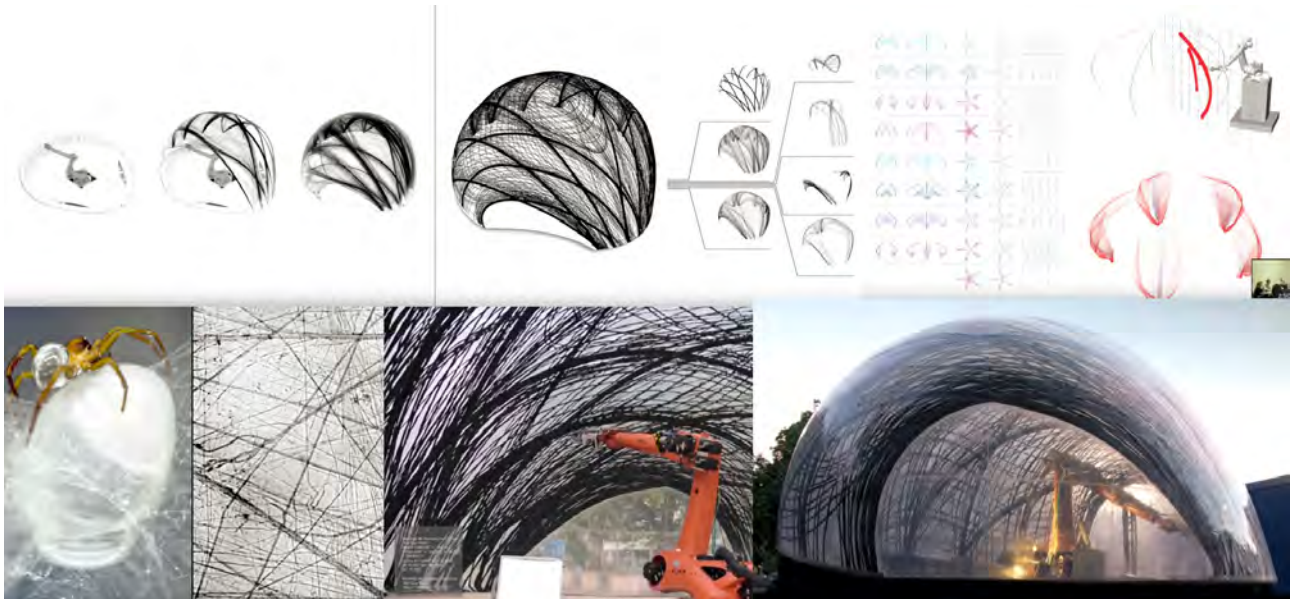


Figure 9. Achim Menges, ICD / ITKE Research Pavilion inspired by the submarine nest of the diver spider.

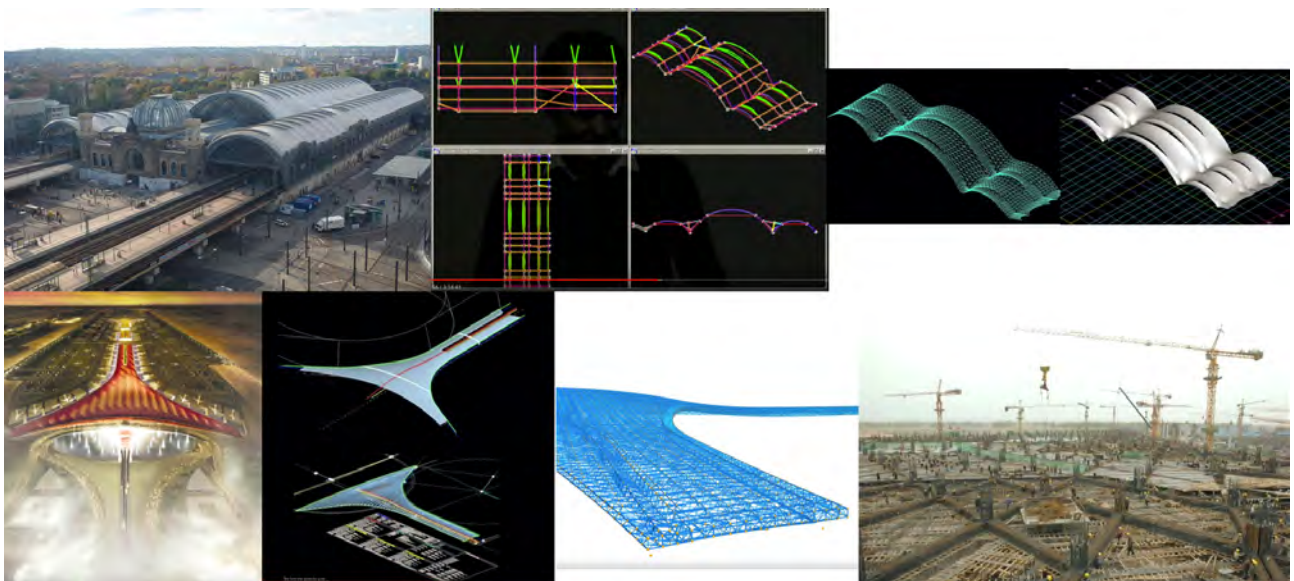


Figure 10. Foster & Partners (Francis Aish): Above: Dresden Railway Station, 1997 Refurbishment; Down: The Beijing Capital International Airport, 2008

Numerical analyses of cement-based piezoelectric smart composites

†*Jan Sladek¹, Pavol Novak², Peter L. Bishay³, and Vladimir Sladek¹

¹Institute of Construction and Architecture, Slovak Academy of Sciences, 84503 Bratislava, Slovakia.

²Faculty of Mechanical Engineering, University of Zilina, Slovakia

³College of Engineering and Computer Science, California State University, Northridge, USA.

*Presenting author: jan.sladek@savba.sk

†Corresponding author: jan.sladek@savba.sk

Abstract

Smart cement-based composite materials have great potential to be used in structural health monitoring (SHM) systems. In this work, the effective thermo-electro-mechanical material properties of cement-based piezoelectric smart composites are characterized. The homogenization techniques are applied on a representative volume element (RVE), where a typical distribution of piezoelectric 3-D particles in the cement matrix is considered. The finite element method (FEM) is used to solve sets of different boundary value problems for the RVE to get the effective thermo-electro-mechanical properties. The effect of the particle volume fraction on the effective composite properties is investigated.

Keywords: 3-D piezoelectric particles, cement matrix, volume fraction, effective material properties, finite element method

Introduction

Reinforced concrete made from cement-based materials is the most popular materials used for larger civil engineering structures. However, this material is not convenient for structural health monitoring (SHM) systems to prevent catastrophic failure of civil structures. To have this possibility for civil structures, cement-based composites with piezoelectric properties need to be developed. Normal mixing and spread techniques were utilized to produce cement-based piezoelectric smart composite [1-4]. Published technical papers on the effect of volume fraction of particles and their size on the properties of the PZT–cement based composites are still very limited in the literature and are mostly based on experimental approaches.

The goal of the present paper is to provide an efficient computational tool for evaluating the effective material properties of PZT–cement based composites. Numerical analyses are performed on a representative volume element (RVE), which contains essential physical geometrical information about the microstructural components represented by the PZT particles in the cement matrix. The computational thermal homogenization, applied to the microscale and mesoscale of concrete sequentially in [5], is extended to the PZT–cement based composites. The finite element model of the RVE is developed to solve boundary value problems with different boundary conditions in order to evaluate the effective thermo-electro-mechanical properties of PZT cement-based composites. These results have not been reported in the literature, according to the best of the authors' knowledge.

Thermal homogenization of smart concrete

Consider a periodic distribution of 3D piezoelectric particles in a cement matrix, and for the numerical simulation, we select a representative volume element (RVE) as shown in Figure 1.

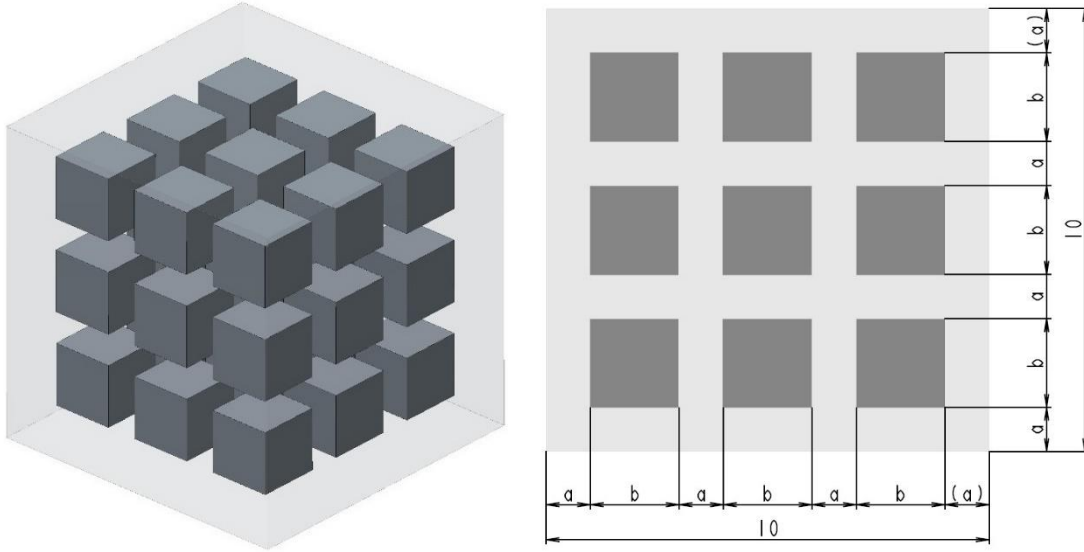


Figure 1. The RVE of a piezoelectric cement composite

The governing equations of the thermo-piezoelectricity consist of Maxwell's equations, the heat conduction equation and the balance of momentum [6]:

$$\sigma_{ij,j}(\mathbf{x}) = 0; \quad D_{i,i}(\mathbf{x}) = 0; \quad \psi_{i,i}(\mathbf{x}) = 0, \quad (1)$$

where σ_{ij} , D_i and ψ_i are the stress tensor, electric displacement vector and heat flux vector, respectively. \mathbf{x} is the position vector. The strain tensor ε_{ij} and the electric field vector E_j are related to the mechanical displacement vector u_i and the scalar electric potential ϕ , respectively, by

$$\varepsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}); \quad E_j = -\phi_{,j}. \quad (2)$$

The constitutive equations express coupling of the mechanical, electrical and thermal fields:

$$\sigma_{ij}(\mathbf{x}) = c_{ijkl}\varepsilon_{kl}(\mathbf{x}) - e_{kij}E_k(\mathbf{x}) - \gamma_{ij}\theta(\mathbf{x}); \quad D_i(\mathbf{x}) = e_{ijk}\varepsilon_{ik}(\mathbf{x}) + h_{ij}E_j(\mathbf{x}) + p_i\theta(\mathbf{x}), \quad (3)$$

where c_{ijkl} , e_{ijk} , h_{ij} and p_i are the elastic, piezoelectric, dielectric and pyroelectric material tensors in a thermal piezoelectric medium, respectively. θ is the temperature difference. The stress-temperature moduli γ_{ij} are expressed as functions of the elastic stiffness coefficients and the thermal expansion coefficients α_{kl} : $\gamma_{ij} = c_{ijkl}\alpha_{kl}$.

The constitutive equations can be written in matrix form (using the reduced Voigt notation) as:

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{23} & 0 & 0 & 0 \\ c_{13} & c_{23} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{bmatrix} - \begin{bmatrix} e_{11} & 0 & e_{13} \\ 0 & e_{11} & e_{13} \\ e_{13} & e_{13} & e_{33} \\ e_{15} & 0 & 0 \\ 0 & e_{15} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix} - \begin{bmatrix} \gamma_{11} \\ \gamma_{22} \\ \gamma_{33} \\ 0 \\ 0 \\ 0 \end{bmatrix} \theta, \quad (4)$$

$$\begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} e_{11} & 0 & e_{13} & e_{15} & e_{15} & 0 \\ 0 & e_{11} & e_{13} & e_{15} & e_{15} & 0 \\ e_{13} & e_{13} & e_{33} & e_{15} & e_{15} & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{bmatrix} + \begin{bmatrix} h_{11} & 0 & 0 \\ 0 & h_{22} & 0 \\ 0 & 0 & h_{33} \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix} + \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} \theta, \quad (5)$$

where

$$\begin{bmatrix} \gamma_{11} \\ \gamma_{22} \\ \gamma_{33} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{12} & c_{22} & c_{23} \\ c_{13} & c_{23} & c_{33} \end{bmatrix} \begin{bmatrix} \alpha_{11} \\ \alpha_{22} \\ \alpha_{33} \end{bmatrix}. \quad (6)$$

The thermal constitutive equation (Fourier's law) is given by

$$\psi_i(\mathbf{x}) = \kappa_{ij} \beta_j(\mathbf{x}); \quad \text{where} \quad \beta_j = \theta_{,j} \quad (7)$$

with κ_{ij} being the thermal conductivity tensor.

The governing equations (1) can be satisfied in the weak form. The residual form in the analyzed domain V is written as:

$$\int_V (\sigma_{ij,j} w_i + D_{i,i} w_\phi + \psi_{i,i} w_T) dV = 0, \quad (8)$$

where w_i , w_ϕ and w_T are arbitrary weighting functions that are assumed to vanish along the boundaries where essential boundary conditions are specified (Γ_u , Γ_p and Γ_θ respectively).

In FEM, the generalized primary (mechanical displacements, electric potential and temperature) and secondary (strains, electric field and temperature gradients) field variables, $\tilde{\mathbf{u}}$ and $\tilde{\boldsymbol{\varepsilon}}$ respectively, can be approximated over the domain of each finite element V^e in terms of the nodal degrees of freedom, \mathbf{q} , and the corresponding shape functions as

$$\tilde{\mathbf{u}} = \mathbf{N}\mathbf{q}; \quad \tilde{\boldsymbol{\varepsilon}} = \mathbf{B}\mathbf{q}, \quad (9)$$

where \mathbf{N} is the shape functions matrix, \mathbf{B} is a matrix that contains derivatives of the shape functions, and $\mathbf{q} = [u_1^{(1)} \ u_2^{(1)} \ u_3^{(1)} \ \phi^{(1)} \ \theta^{(1)} \ u_1^{(2)} \ \dots]^T$ is the nodal degrees of freedom vector in 3D, with superscript (k) indicating node number. The finite element equation is then given by

$$\mathbf{K}\mathbf{q} = \mathbf{f}_t; \quad \mathbf{K} = \int_{V^e} (\mathbf{B}^T \tilde{\mathbf{C}} \mathbf{B} + \mathbf{B}^T \mathbf{G} \mathbf{N}) dV; \quad \mathbf{f}_t = \int_{\Gamma_n} \mathbf{N}^T \tilde{\mathbf{t}} d\Gamma, \quad (10)$$

where

$$\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\varepsilon} + \mathbf{G}\mathbf{u}; \quad \mathbf{C} = \begin{bmatrix} \mathbf{C} & -\mathbf{e}^T & \mathbf{0} \\ \mathbf{e} & \mathbf{h} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\kappa} \end{bmatrix}; \quad \mathbf{G} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & -\boldsymbol{\gamma} \\ \mathbf{0} & \mathbf{0} & \mathbf{p} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (11)$$

The composite response is determined by the homogenized composite properties or the effective material properties. The effective material coefficients of the piezoelectric cement composite can be computed from the constitutive equations written for the average values of the secondary fields and the conjugated fields on the RVE sample [7].

Results

Consider a sample of Portland cement matrix with embedded PZT-SH particles. The material coefficients of the PZT-SH can be found in [8]. Portland cement matrix properties are as follows: Young's modulus $E = 1.4 \times 10^{10} \text{ Nm}^{-2}$, Poisson's ratio $\nu = 0.2$, heat conductivity $\kappa = 0.29 \text{ W/m deg}$, and thermal expansion coefficient $\alpha = 11.10^{-6} \text{ 1/deg}$. Figure 2 and Figure 3 show the effect of the PZT particle volume fraction in the smart composite on the effective elastic and piezoelectric material properties. The volume fraction of the PZT particles is varied from 3% to 60%. One can observe from the figures that as the volume fraction of the PZT ceramic particles increases, the elastic material properties of the composite increase. The absolute value of all piezoelectric coefficients significantly increases as the volume fraction increases. The red point in Figure 2 (right) indicates the values of the effective c_{13} for volume fraction 30% and sample size of 1.85 mm obtained in [9].

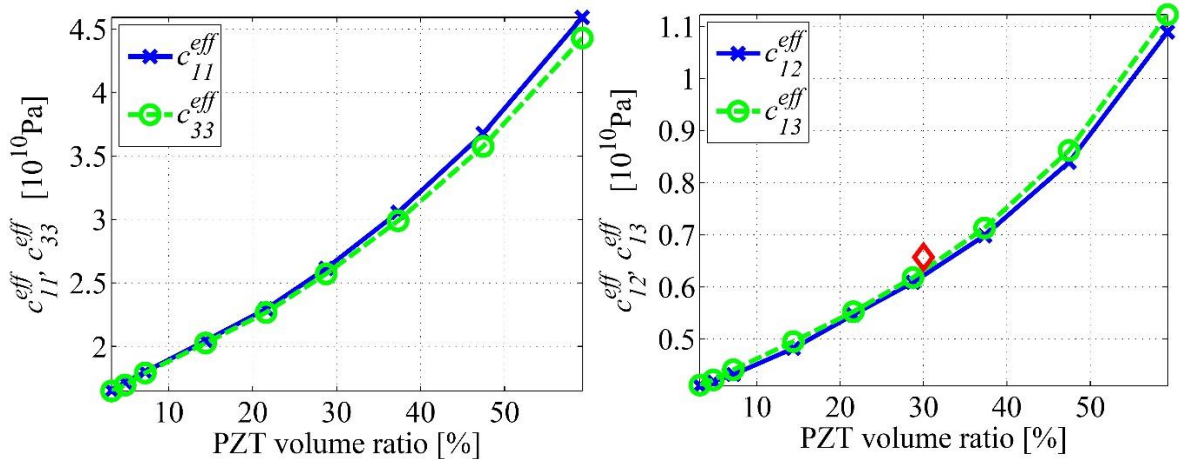


Figure 2. Variation of the effective (left) c_{11} and c_{33} , (right) c_{12} and c_{13} with PZT volume fraction in the smart ceramic composite

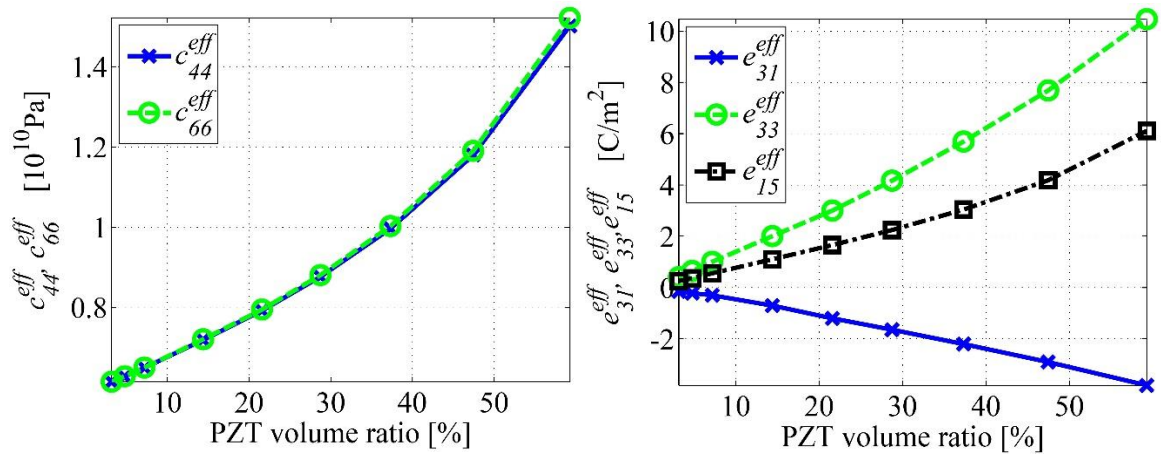


Figure 3. Variation of the effective (left) c_{44} and c_{66} , (right) e_{31} , e_{33} and e_{15} with PZT volume fraction in the smart ceramic composite

Figure 4 shows the effect of the PZT particle volume fraction on the effective dielectric (left) and thermal conductivity (right) material properties. One can observe that as the volume fraction increases, the effective dielectric and thermal conductivity coefficients decrease.

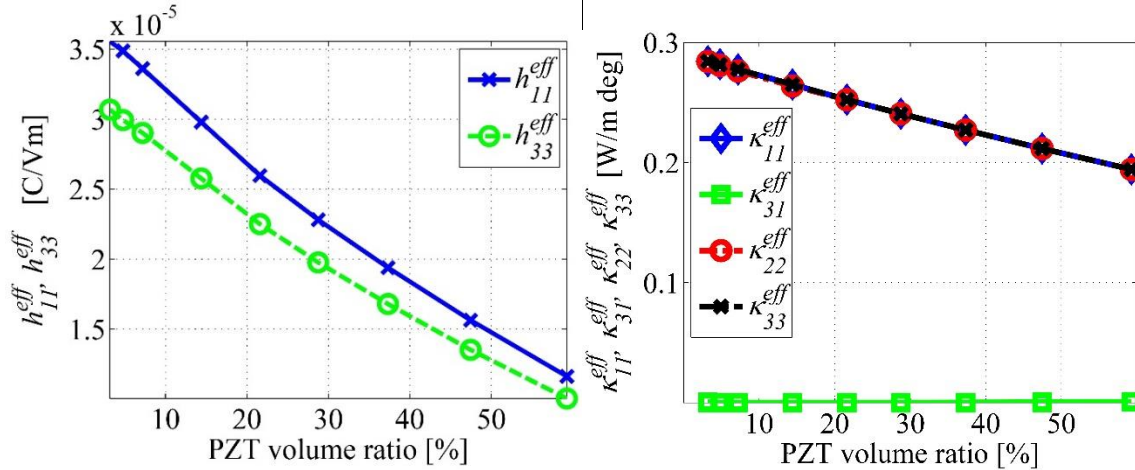


Figure 4. Variation of the effective (left) h_{11} and h_{33} , (right) k_{11} , k_{31} , k_{22} and k_{33} with PZT volume fraction in the smart ceramic composite

Conclusions

A finite element model of an RVE of cement-based piezoelectric composite is developed and applied to evaluate the effective elastic and piezoelectric material properties. The effect of the volume fraction of the piezoelectric particles on the effective thermo-electro-mechanical material properties of the composite are presented.

Acknowledgement

The authors gratefully acknowledge the supports by the Slovak Science and Technology Assistance Agency registered under number APVV-14-0216 and the Slovak Grant Agency VEGA-2/0046/16.

References

- [1] Li, Z., Zhang, D. and Wu, K. (2001) Cement matrix 2-2 piezoelectric composite – part 1. Sensory effect, *Materials and Structures* **34**, 506–512.

- [2] Dong, B. and Li, Z. (2005) Cement-based piezoelectric ceramic smart composites, *Composites Science and Technology* **65**, 1363–1371.
- [3] Chaipanich, A. (2007) Effect of PZT particle size on dielectric and piezoelectric properties of PZT-cement composites, *Current Applied Physics* **7**, 574–577.
- [4] Wang, F., Wang, H., Song, Y. and Sun, H. (2012) High piezoelectricity 0-3 cement-based piezoelectric composites, *Materials Letters* **76**, 208–210.
- [5] Wu, T., Temizer, I. and Wriggers, P. (2013) Computational thermal homogenization of concrete, *Cement & Concrete Composites* **35**, 59–70.
- [6] Mindlin, R.D. (1974) Equations of high frequency vibrations of thermopiezoelectricity problems, *International Journal of Solids and Structures* **10**, 625–637.
- [7] Li, J.Y. and Dunn, M.L. (1998) Micromechanics of magneto-electroelastic composite materials: average fields and effective behavior, *Journal of Intelligent Material Systems and Structures* **9**, 404–416.
- [8] Xing, F., Dong, B. and Li, Z. (2008) Dielectric, piezoelectric, and elastic properties of cement-based piezoelectric ceramic composites, *Journal of American Society* **91**, 2886-2891.
- [9] Saputra, A., Sladek, V., Sladek, J., Song, C. (2017) Micromechanics determination of effective material coefficients of cement-based piezoelectric ceramic composites. *Journal of Intelligent Material Systems and Structures* **29**(5), 845-862.

Wire rope model with elliptic cross sectional outer wires

***Erdönmez C.¹**

¹Department of Science, Turkish Naval Academy, National Defense University, 34940 Istanbul.

*Presenting and Corresponding author: cerdonmez@gmail.com

Abstract

Wire ropes are used in variety of application areas such as cranes, bridges, mining due to their large tensile and in addition easy storage capability. It is inevitable to carry out numerical tests in order to determine rope behaviors in different application fields. Three dimensional solid models without errors are needed for finite element analysis. Therefore, the three-dimensional meshed model must be prepared for analysis. Traditionally, wire ropes made by using circular cross-sectional wires but this behavior is started to change due to the different needs in various application areas. For this reason, it seems that different types of rope models have been designed, such as flattened or swaged rope surfaces. In this article, a mesh model of a wire rope with an elliptical cross-sectional outer wires has been created. This model has a great emphasis in the creation of rope models with using different cross-sectional outer wires. In future studies, other type of cross-sectional outer wires can be used to model new wire ropes using the similar technique.

Keywords: Wire Strand, Wire Rope, Independent Wire Rope Core, meshed helical model

Introduction

Love is described the mechanical theory of thin rods in his treaties named "Theory of Elasticity" [1]. In this book Love mentions the mechanical characteristics of thin rods and gives the general equilibrium equations. After 1970, with the help of computer technology, many studies are published in the literature. Costello's papers on various aspects of wire ropes are shed light to many other researchers to study the different behavior of wire ropes [2]. By using the Finite Element Analysis capability in engineering applications computer models of wire strands and wire ropes are take attention. At the beginning of the seventies, Carlson and Kasper built a simplified model for armored ropes using the finite element method [3]. A finite element model of a simple straight strand based on a Cartesian isoparametric formulation is given by Nawrocki who take into account every possible interwire motion in [4]. A wire rope with an IWRC model, which fully considers the double-helix configuration of individual wires considered by Elata et.al. in [5]. Parametric mathematical equations of single and double helical wires within an IWRC is represented in [6]. Modeling issues of nested helical structure based geometry for numerical analysis and the encountered problems and solution techniques are mentioned in [7]. Lately a geometric model of spiral one or two-layered oval nested wire strands are proposed by Stanova et.al. in [8].

Modeling wire rope structures needs a comprehensive care for Finite Element Analysis (FEA) due to the requirement of error free structured meshed model of the wire rope for FEA. Structure of a wire rope consists of different sized wires with various pitch lengths and helix angles. In addition wires within a wire strand or wire rope named according to its form as straight, single, double or triple. These types of helical shapes are composed by coiling a wire around another or wrapping a strand over another one. But before this wrapping process each wire basically has a cylindrical shape with circular cross sections. Wire ropes are seen in

different structures according to their usage areas. Some new kind of wire ropes by using non-common cross sectional shaped wires such as a strand created by using some techniques such as compacting, swaging, plastic coating and filling or using different shaped outer wires such as elliptic outer wires. The process of compacting or swaging flattens the surface of the outer wires and reforms internal wires of the strand to increase the density of the strand. In compacted wire rope represented in Figure (1) or swaged wire rope, the wires are compacted or reduced in diameter while stranding operation or in a separate operation after stranding. Due to need for mechanical process for compacting or swaging, elliptical outer wires are preferred to use in numerical modeling of wire ropes to find flattened outer surface. The importance of this modeling process is to have a meshed model of the wire rope at the same time without the need for any other tool. Once the necessary parameters such as pitch length, rotational angles, wire radiuses and the type of the wire rope are defined the meshed model of the wire rope is created at the same time which is ready to make FEA.

Each wire within a wire strand or wire rope are straight wires at the beginning of the winding process. Wires within a wire strand consists of a straight center wire wound by 7 single helical wires called as Wire Strand (WS). Single helical wires also composed by using straight wires. Before winding process each wire is commonly has a circular cross sections. In this article a strand, whose outer wire cross section is elliptic, is modeled and meshed. It has a straight center wire which is coiled up with six elliptic shaped outer wires as presented in Figure (2).

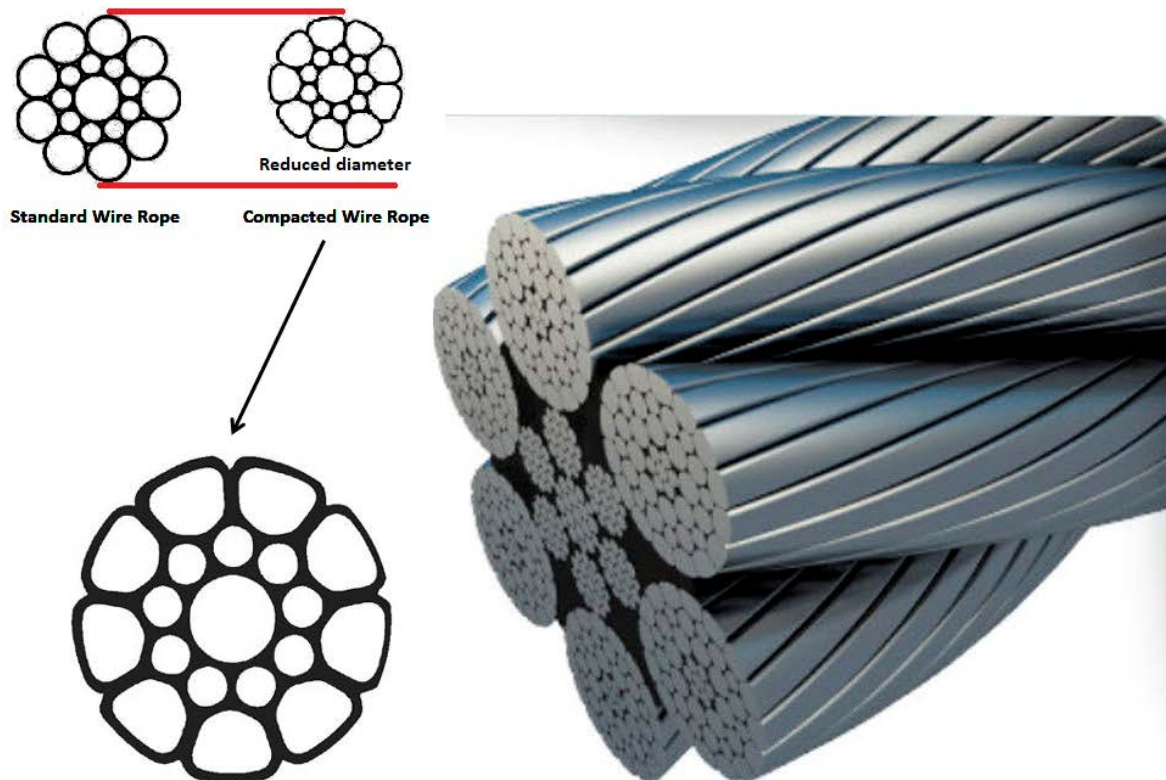


Figure 1: Compacted wire rope.

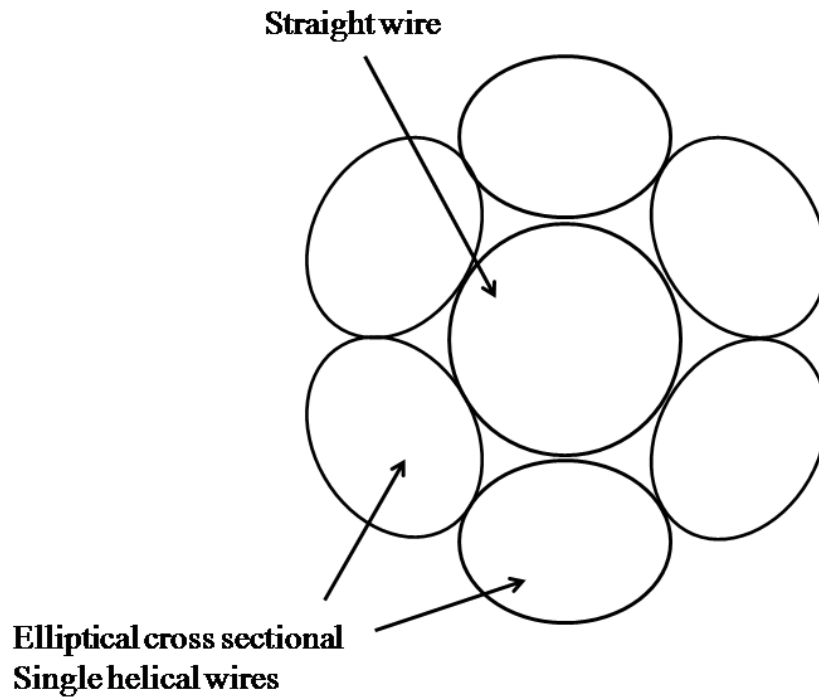


Figure 2: A simple wire strand with elliptic outer wires.

Conclusions

Wire rope modeling with different shaped cross-sectional outer wires is described in this paper. The meshed 3D model of the wire strand is modeled using the parametric equations of helical geometry. The importance of this modeling process is to have a meshed model of the wire rope at the same time without the need for any other tool. In this article a strand, whose outer wire cross section is elliptic, is modeled and meshed. It has a straight center wire which is coiled up with six elliptic shaped outer wires. This modeling issue enables to analyze various shaped wire ropes using the Finite Element Analysis. In addition other type of shapes could be modeled using the described modeling technique to create different type of wire ropes.

References

- [1] Love AEH (1944) A treatise on the mathematical theory of elasticity, 4th ed., New York: Dover Publ., First Am. Print. 1944, Chapter XVIII-XIX, 381-426.
- [2] Costello GA (1990) Theory of wire rope, Berl.: Springer.
- [3] Carlson, A.D., Kasper, R.G., 1973. A structural analysis of multi-conductor cable. Report No. AD-767 963, Naval underwater systems center, distributed by National Technical Information Service.
- [4] Nawrocki, A., Labrosse, M., 2000. A finite element model for simple straight wire rope strands, Computers and Structures 77, 345-359.
- [5] Elata, D., Eshkenazy, R., Weiss, M.P., 2004. The mechanical behavior of a wire rope with an independent wire rope core, International Journal of Solids and Structures 41, 1157-1172.
- [6] Erdönmez, C., İmrak, C.E., (2011). A finite element model for independent wire rope core with double helical geometry subjected to axial loads, SADHANA-Academy Proceedings in Engineering Sciences DOI: 10.1007/s12046-011-0053-1.

- [7] Erdönmez, C., İmrak, C.E., (2011). Modeling Techniques of Nested Helical Structure Based Geometry for Numerical Analysis, *Strojniški vestnik - Journal of Mechanical Engineering* Volume 57(2011)4, 283-292, DOI:10.5545/sv-jme.2009.006
- [8] Eva Stanova, Gabriel Fedorko, Stanislav Kmet, et.al., (2015). Finite element analysis of spiral strands with different shapes subjected to axial loads, *Advances in Engineering Software*, Volume 83, 2015, Pages 45-58, ISSN 0965-9978, <https://doi.org/10.1016/j.advengsoft.2015.01.004>.

Damage assessment by Non-Smooth Contact Dynamics method of the iconic crumbling of the clock tower in Amatrice after 2016 center Italy seismic sequence

***Francesco Clementi¹, Gabriele Milani², Valentina Gazzani¹, †Marina Poiani¹ and Stefano Lenci¹**

¹Department of Civil and Building Engineering, and Architecture (DICEA), Polytechnic University of Marche, Ancona, Italy

²Department of Architecture, Built Environment and Construction Engineering (ABC), Politecnico di Milano, Milan, Italy

*Presenting author: francesco.clementi@univpm.it

†Corresponding author: m.poiani@pm.univpm.it

Abstract

The dynamics of the medieval civic clock tower of Amatrice (Rieti-Italy) has been studied by means of the Non-Smooth Contact Dynamics (NSCD) method, implementing a discrete element numerical model in the LMGC90[®] code. Schematized as a system of rigid blocks, undergoing frictional sliding and plastic impacts, the tower has exhibited a complex dynamic, because of the geometrical non-linearity and the non-smooth nature of the contact laws. Numerical simulations are performed with the aim of comparing the numerical result and the observed damages after the seismic sequence of the Central Italy earthquakes.

Keywords: NSCD method, nonlinear dynamics, masonry towers, damage assessment

Introduction

The seismic events which hit Central Italy on 24th August, 26th and 30th October 2016, and 18th January 2017, have caused casualties and major damage mostly to buildings and architectural heritage of the Italian regions of Marche, Lazio, Abruzzo, and Umbria. The mainshock occurred on August 24th at 3:36 am (local time) with an epicenter close to Accumoli (Rieti province) and with a magnitude $M_w = 6.2$; it was followed, at 4.33 am, by an aftershock with an epicenter close to Norcia (Perugia province) and with a magnitude $M_w = 5.5$. These events caused a total of 299 fatalities, 386 injured and about 4800 homeless [1,2]. Most of the victims were in the areas of Amatrice, Accumoli, and Arquata del Tronto. In these municipalities, heavy damage and collapse of residential buildings were reported.

On 26th October, there were two strong aftershocks, the first at 07:10 pm with M_w 5.6 and the second at 09:18 pm with M_w 6.1. The earthquake of 30th October, which happened at 07:40 am, with a M_w 6.5, is the largest event in terms of released energy occurred in Italy since the M_w 6.9 in 1980 in the event of the Irpinia earthquake.

The events of 26th and 30th October did not cause any victim thanks to the evacuation of people from damaged and vulnerable houses after the previous seismic events. It has also to be considered that the October epicenters are located close to Norcia municipality, where many buildings had been strengthened after the 1997 earthquake. Nevertheless, the impact of the seismic events of 26th and 30th October 2016 and 18th January 2017 was distributed on a larger portion of territory extending northwards in the Marche Region respect to the earthquake of 24th August that had a very destructive impact on a restricted area included in the above-listed municipalities. Many small towns and villages, which have survived to the first earthquake, were heavily damaged during the 30th October earthquake [3–5]. Finally, on

18th January 2017 took place a new sequence of four strong shocks of $M_w=5$, with a maximum equal to $M_w=5.5$, and epicenters located between the municipalities of Montereale, Capitignano and Cagnano Amiterno. All these earthquakes are indicated in Fig. 1 with the relative intensity map.

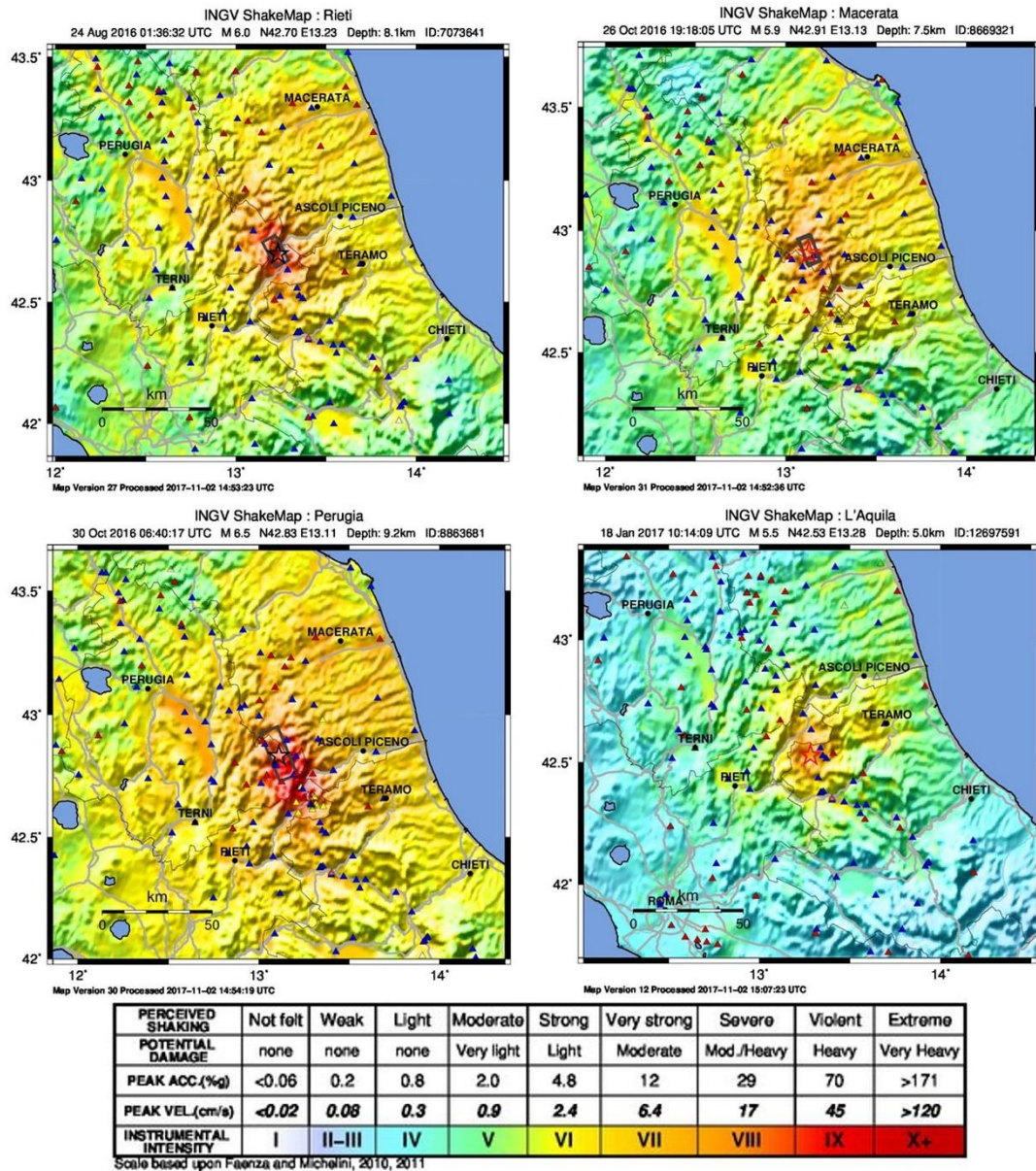


Figure 1. The maps of ground shaking (<http://shakemap.rm.ingv.it/shake/index.html>) of the four main shocks of the central Italy sequence 2016-2017.

Symbol of the damage and destruction done by a long sequence of the strong earthquake of 2016-2017 is the Amatrice civic clock tower which will be investigated in this paper. An advanced numerical model is here utilized to have an insight into the modalities of progressive damage and the behavior of the structure under strong non-linear dynamic excitations, namely the Non-Smooth Contact Dynamic (NSCD) method. A full 3D detailed discretization is adopted. The tower is schematized as a system of rigid blocks, undergoing frictional sliding and perfect plastic impacts. The structure exhibited a complex dynamic behavior, because of the geometrical non-linearity and the non-smooth nature of the contact laws.

From the numerical results, both the role played by the actual geometries and the insufficient resistance of the constituent materials are envisaged, showing a good match with actual crack patterns observed after the seismic sequence. The numerical analyses provide a valuable picture of the actual behaviour of the structure, thus giving useful hints for the reconstruction.

The Amatrice clock tower

Symbol of the city of Amatrice (Fig. 2a), the Civic Tower rises in the city center namely Cacciatori del Tevere square, underlining the crossroads of the two main streets of the city, Via Roma and Corso Umberto I (Fig. 2b). In the opposite side of the tower there is the Town Hall, intentionally positioned to symbolize the centralization of civic power. In the area behind the tower, still on the same square, it overlooks another small treasure of Amatrice i.e. the Church of San Giovanni. There are few historical data about the Civic Tower, its origins are placed back to medieval times, as early as 1293 was mentioned in ancient documents. These sporadic historical informations tell us that the tower was originally connected to the Church of Santa Lucia, demolished in 1684 by the feudal "lord" Alessandro Maria Orsini who wanted to give a wider space to the street of the main course, expanding the square. On this occasion, the base of the tower was reinforced and a small annex was added, leaning against it on two sides, to obtain the main entrance to the belfry.



Figure 2. Amatrice (Rieti, Italy) (a), and location of the civic clock tower inside the historic center (b).

The last and probably the only consolidation intervention was carried out on the tower dates back to the early 80s. Already requested in the previous decades it was solicited in 1979, when, following the earthquake of the Alta Valnerina (central Italy), significant damage was noticed to the tower. In 1985 the original bell of 1494 was replaced as the latter, because it had undergone a crack during the restoration phases. The original one is preserved in Madonna di Porta Ferrata Church and a lighter bell has been inserted in the tower so as to avoid the high oscillation of the tower as in the past.

Geometric survey

The civic tower of Amatrice has a rectangular plan of 4.00 m x 5.30 m and a height of about 25 m. At the base, there is a small overhang leaning only two walls, to north elevation with a depth of 1.5 m, and to east elevation with a depth of 0.60 m.

The annex houses the staircase leading to the entrance of upper floors. In its vertical development, there are three distinct areas, identified by decorative frames and folds that externally show the reduction of the wall thickness. The first floor is located at about 9 m in height and it is composed of smoothed stones on the outer side, falling 15 centimeters in the

wall thickness, while the second frame is located at a height of about 19 meters and it marks the passageway from the tower to the belfry. In its highest part, there is the bell which develops longitudinally for just over 5 m. It consists of 4 regular piers, with dimension $0.90 \times 0.80 \text{ m}^2$, on which rest round arches that make up the single-hole openings on the four sides of the tower. The belfry has a single symmetry, so the openings of the belfry are the same on the opposite sides (see Fig. 3). Finally, the tower presents a pavilion wooden roof.

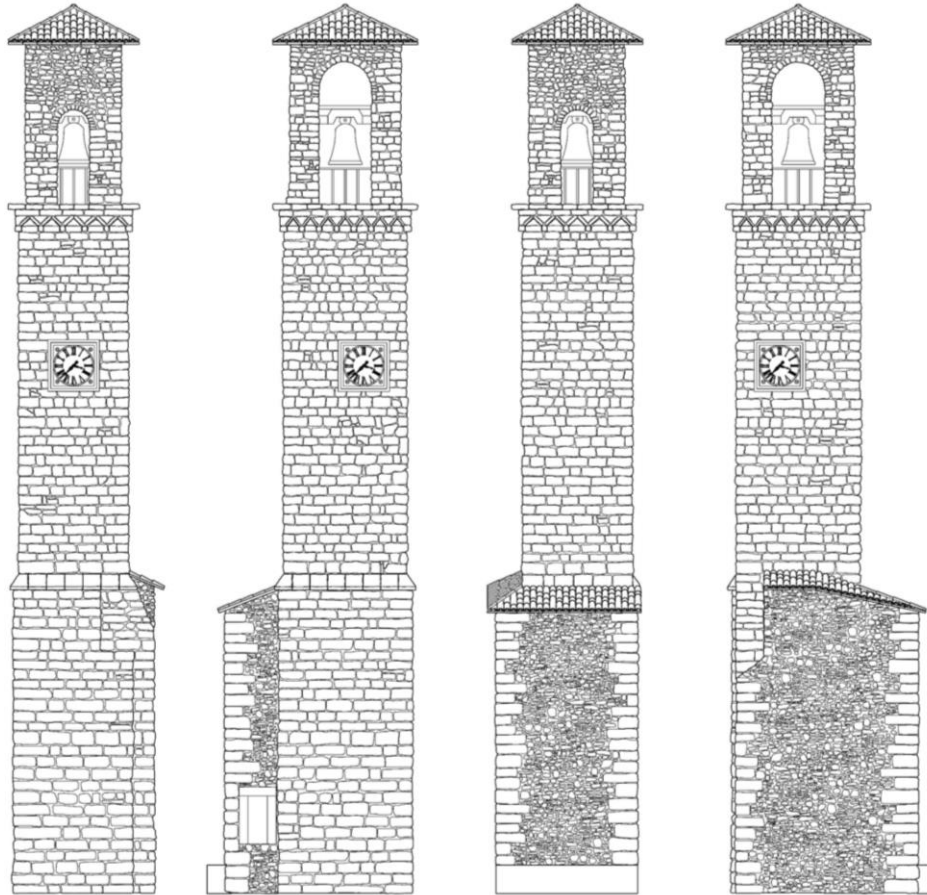


Figure 3. Main fronts of the civic tower of Amatrice.

The model

In this section the principal peculiarities of the Non-Smooth Contact Dynamics method, the modelling simplifications on which it is based, and its aptitude to reproduce the dynamics of large three-dimensional ancient masonry structures are highlighted. As the deepening of the NSCD theory goes beyond our purposes, an exhaustive description of the method is reported in [6,7]. The problem parameters and the seismic excitation applied to the base of the civic tower are also briefly reported.

Non-smooth Contact Dynamic method

The dynamics of a system of rigid bodies is governed by the equation of motion and by the frictional contact conditions. To describe the frictional contact laws, we must introduce some basic definitions. In the following, the notation adopted in [6] is used (scalars, vectors and tensors are explicitly declared, and italic letters are used for all of them). Given two arbitrary bodies B_i and B_j , let P_i and P_j (Fig. 4a) be the points of possible contact on the boundaries of B_i and B_j , respectively, and let n be the outer unit vector, orthogonal to the boundary of B_i in

P_i . We define $g = (P_j - P_i) \cdot n$ the gap between P_i and P_j (a dot means scalar product), $(\dot{u}_n; \dot{u}_t)$ the normal and tangential velocities of P_j with respect to P_i , and (r_n, r_t) the normal and tangential reactive forces of B_i on B_j .

The contact conditions are:

1. The Signorini's law of impenetrability (Fig. 4b)

$$g \geq 0, r_n \geq 0, gr_n = 0, \quad (1)$$

which, in the case of contact $g = 0$, is equivalent to the following Kuhn–Tucker conditions [7]

$$\dot{u}_n \geq 0, r_n \geq 0, \dot{u}_n r_n = 0, \quad (2)$$

written in term of relative normal velocity.

2. The dry-friction Coulomb's law (Fig. 4c), that governs the behavior in the tangential direction

$$|r_n| \leq \mu r_n; \begin{cases} r_t < \mu r_n \rightarrow \dot{u}_t = 0 \\ |r_t| = \mu r_n \rightarrow \dot{u}_t = -\lambda \frac{r_t}{|r_t|}, \end{cases} \quad (3)$$

with μ the friction coefficient, and λ an arbitrary positive real number.

If q is the vector of the system configuration parameters (unknown translations and rotations of each body), and p is the global vector of reaction forces, the equation of motion can be written as follows

$$M\ddot{q} = f(q, \dot{q}, t) + p, \quad (4)$$

where M is the mass matrix, and f is the vector of external forces.

The local pairs $(\dot{u}_n; \dot{u}_t)$ and (r_n, r_t) , characteristic of each contact, are related to the global vectors \dot{q} and p , respectively, through linear maps which depend on q (see [1] for details). Since the contact laws (1) - (3) are non-smooth, velocities \dot{q} and reactions p are discontinuous functions of time. They belong to the set of bounded variation functions, i.e. functions which, at each time, have finite left and right limits. Since the accelerations are not defined when the velocities are discontinuous, Eq. (4) is reformulated in integral form (see [6,7] for the mathematical details), and solved numerically using a time-stepping approach, alternative to the event-driven method (we point out that time-stepping approaches are more appropriate than event-driven schemes when problems with many contacts are handled). The time is discretized into time intervals, and, within each time interval $[t_i, t_{i+1}]$, the equation of motion is integrated over the interval as follows

$$M(\dot{q}_{i+1} - \dot{q}_i) = \int_{t_i}^{t_{i+1}} f(q, \dot{q}, t) dt + \bar{p}_{i+1},$$

$$q_{i+1} = q_i + \int_{t_i}^{t_{i+1}} \dot{q}(t) dt. \quad (5)$$

Where \bar{p}_{i+1} is the impulse in $[t_i; t_{i+1}]$. The primary variables of the problem are the velocity vector \dot{q}_{i+1} and the impulse vector \bar{p}_{i+1} at the instant t_{i+1} . In the NSCD method, the integrals in (5) are evaluated by means of an implicit time integrator. The overall set of global Eq. (5) and

local contact relations (1) and (2), where the reactions are approximated by the average impulses in $[t_i; t_{i+1}]$, are condensed at the contact local level, and then they are solved by means of a non-linear Gauss-Seidel by block method.

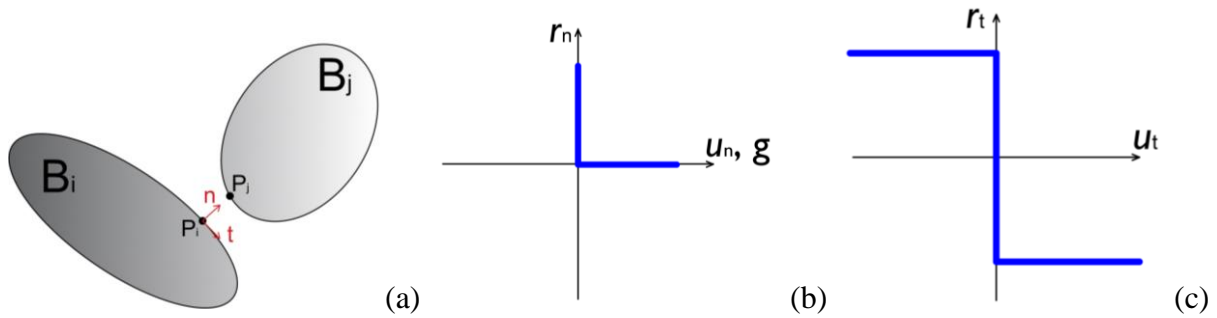


Figure 4. The interaction between two bodies (a) Signorini's law (b) Coulomb's law (c)

The numerical model, which implements the NSCD method, is based on some modelling simplifications which deserve some comments in perspective of its application to ancient buildings. Regarding the contacts between bodies, the model does not account for elasto-plastic impacts governed by restitution laws on velocities (Newton law) or impulses (Poisson law) [8], or energetic impact laws [9], although originally the Contact Dynamics method proposed by Moreau [7] considered the Newton law.

The relations (1) imply perfectly plastic impact, i.e., the Newton law with restitution coefficient equal to zero. A perfectly plastic impact law makes impossible to describe, for instance, bouncing phenomena, and, furthermore, overestimates the energy dissipated during impacts. However, in the case of systems of bricks or stones, the restitution coefficient has low values, and thus bouncing phenomena are secondary, and they can be neglected. More sophisticated impact laws would bring to more accurate models, such as that proposed in [10] for the dynamics of masonry arches, but they drastically increase the model complexity and result impracticable for large systems with many impacts, like that considered here.

Furthermore, the deformability of blocks is neglected. This is a reasonable approximation since the expected operating compressive stresses at the base of the masonry walls of the tower are reasonably low. Deformable blocks have been considered in [11,12] for two-dimensional systems. Since deformability drastically increases the computational complexity, practically it cannot be applied to large three-dimensional structures, like the tower of this study. On the other hand, simplified two-dimensional schemes rule out a crucial aspect of the dynamics of box-shaped structures such as houses, churches, and towers that is, the interaction between adjacent walls laying on different planes (for instance the façade and the longitudinal walls in churches), which mutually exchange considerable inertia forces.

Since we are interested in the dynamical interactions between different parts of the civic clock tower, we consider three-dimensional schemes but we neglect blocks deformability. It follows that the numerical results obtained depict an overall picture of the tower dynamics and describe the failure mechanisms of the whole tower, due to blocks rocking and sliding, but, obviously, they do not give a description of the stresses and strain distributions within each block.

Since experimental data are not available, the friction coefficients were selected from standard values reported in the literature. The values of μ range from 0.3 to 1.2, according to different combinations of units and mortars [13]. As a first attempt, we assume the value $\mu = 0.5$ for the interface block/block, and $\mu = 0.9$ for the interface block/foundation in order to observe, mainly, the dynamics of the tower without interaction with the foundation. Furthermore, it is important to underline that, in real old masonry buildings, the degradation of the mortar over

time contributes to deteriorate the friction coefficient confirming the hypothesis of the first attempt.

Finally, it is important to observe that damping, a fundamental aspect of continuum models, is not considered here, and only friction and perfectly plastic impacts contribute dissipating energy.

Discretization scheme and analysis settings

The geometrical complexity of Amatrice civic clock tower requires some geometrical simplifications, being impossible to reproduce the real layout of the masonry walls, made by brick fragments and ashlar of small size.

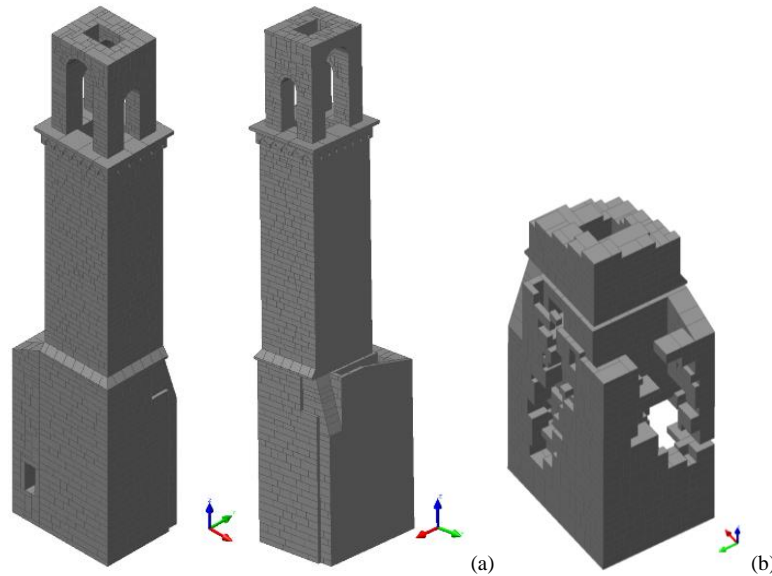


Figure 5. Blocks discretization for the geometrical scheme of the tower in LMGC90[®] (b) view inside the thickness of the walls in LMGC90[®].

With the final aim of confirming what occurred following the seismic shocks, it was decided to use a punctual mapping of the masonry as it is possible to see from Fig. 2 with the discrete approach using the code LMGC90[®]. The size of the blocks (Fig. 5a) used is directly taken up by the reliefs of the facades, while the internal wall texture has been hypothesized avoiding the addition of several transversal connections (i.e., *diatoni*) between the two-external leaf. Only in the presence of very small and irregular ashlar, typically at the top of the belfry and in the annex walls, we have used larger dimensions than relief, simply merging up to five (small) adjacent blocks. As it is possible to observe in Fig. 5b, where the numerical model is reported, the two-leaf masonry is modelled at the best of possibilities. Obviously, the rounded geometry of the blocks has been regularized using straight vertical and horizontal surfaces to avoid further computational burdens. Finally, the numerical model is composed of 2899 rigid blocks with different geometries.

Concerning the seismic loading, the accelerations of Amatrice (Italy) of the 24th August 2016 earthquake have been considered. During that earthquake in Amatrice (AMT) a Peak Ground Acceleration (PGA) of 850.804 cm/s² and a Peak Ground Velocity (PGV) of 43.549 cm/s have been registered (see the website: <http://itaca.mi.ingv.it>). In the numerical simulations, accelerations are applied to FE model and velocities are applied to DE model at the base where the tower is laid. The three velocities components in the three main coordinate directions are determined by direct integration of the accelerations in a time interval of 28 s,

during which the maximum amplitudes are attained, without the use of correlation method (Fig. 6). In the simulations, the time step $dt = 0.005$ s has been used.

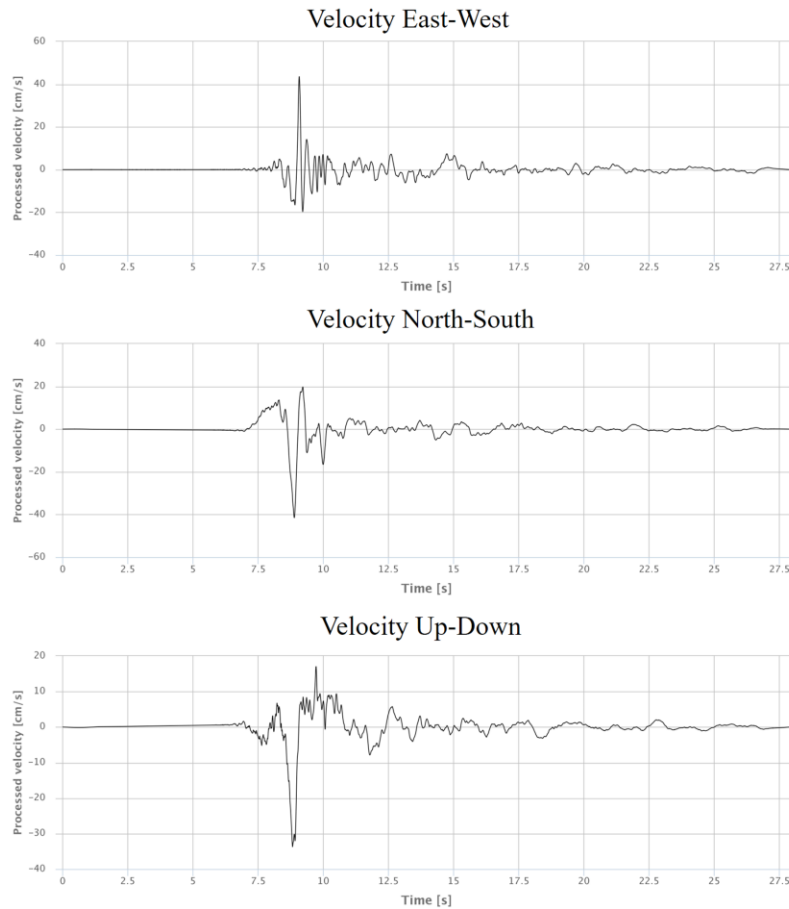


Figure 6. Velocities of the shock of 24th August 2016 applied to the foundation in the three coordinate directions.

Preliminary results of numerical analyses

The first numerical results are reported in Fig. 7 for different time steps. With the above data, the tower collapses at the bell cell, where the overturning mechanisms is favored by the presence of non-regular and small-sized materials. The bell blocks motion occurs in the first 10 s of the seismic excitation, during which the largest acceleration peaks are attained. For time instants larger than 12.5 s, the tower stays at rest, since the acceleration peaks of the seismic excitation are not sufficiently large to activate other blocks sliding. Further damage can also be read from Fig. 7a in the enlargement area at the base of the tower, where there are significant cracks along the perimeter walls, without showing any collapses.

This agrees with the real damage shown by the tower following the shock of 24th August 2016, after that the bell tower was seriously damaged in correspondence of the bell cell (Fig. 7b). The non-activation of the two mechanisms - instead numerically obtained - are certainly due to the presence of an interaction with the roofs which, to a certain extent, have prevented the start of a tilting mechanism of the bell tower columns and of the perimetral wall at the base of the tower.

Then the subsequent earthquakes have led to the collapse of the bell cell already seriously compromised (Fig. 7b).

As can be observed from the Fig. 7a, the NSCD method totally differs from a continuum approach, and it stands as complementary to this latter. While the NSCD accounts for an

accurate description of the motions induced by the inertial masses, a continuum approach describes stress and strain distributions [3,14–16]. The combination of both two methods can bring to a complete comprehension of the mechanical response of such complex structures to seismic loadings, but it will be left for future works.

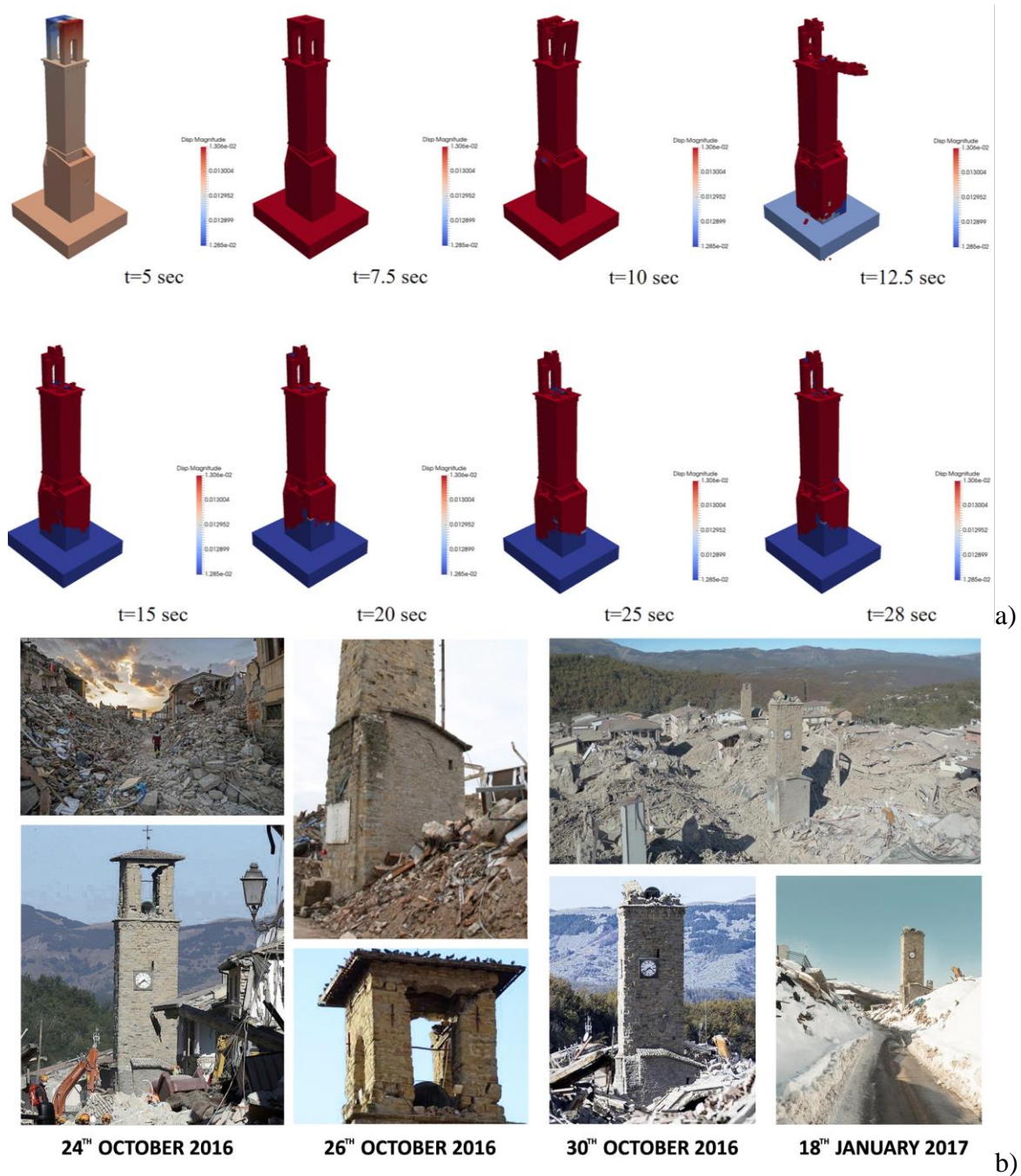


Figure 7. Numerical analysis damage with the NSCD method for the friction coefficient $\mu=0.5$ at different time steps (a) the increment of real damages after the seismic sequences (b).

Conclusions

The Non-Smooth Contact Dynamics method, implemented in the LMGC90[®], has revealed a powerful tool to explore the complex dynamical behavior of an ancient brick-made building, as the civic clock tower of Amatrice (Rieti, Italy). Indeed, it combines modelling simplicity

and great predictive capabilities. Its simplicity comes from the following fundamental simplifying assumptions: (i) block rigidity (often a feasible assumption for masonry structures); (ii) simple contact laws, such as the Signorini's law, which supposes that impacts between blocks are purely plastic, and the dry friction Coulomb's law for the tangential relative motions between blocks; (iii) absence of any damping, according to which the kinematic energy of the system is dissipated only by impacts and friction. As a result, the mechanics are governed by only one material parameter, the friction coefficient. This is an advantage for modelling ancient building, for which the determination of the mechanical properties (like the material moduli) is always uncertain and variable. Despite its simplicity, the model can predict a large variety of behaviors.

The numerical results have given a deep insight into the seismic vulnerability of the considered civic clock tower, pointing out the same portions most damaged during the earthquake of 24th August 2016. In fact, the belfry and the enlargement at the base are the most damaged portions following the main shock.

References

- [1] Italian Department of Civil Protection, 2016 Central Italy earthquake, (2017). http://www.protezionecivile.gov.it/jcms/it/terremoto_centro_italia_2016.wp.
- [2] Fiorentino, G., Forte, A., Pagano, E., Sabetta, F., Baggio, C., Lavorato, D., Nuti, C. and Santini, S. (2018) Damage patterns in the town of Amatrice after August 24th 2016 Central Italy earthquakes, *Bulletin of Earthquake Engineering* **16**, 1399–1423.
- [3] Clementi, F., Pierdicca, A., Formisano, A., Catinari, F. and Lenci, S. (2017) Numerical model upgrading of a historical masonry building damaged during the 2016 Italian earthquakes: the case study of the Podestà palace in Montelupone (Italy), *Journal of Civil Structural Health Monitoring* **7**, 703–717.
- [4] Clementi, F., Quagliarini, E., Monni, F., Giordano, E. and Lenci, S. (2017) Cultural Heritage and Earthquake: The Case Study of “Santa Maria della Carità” in Ascoli Piceno, *The Open Civil Engineering Journal* **11**, 1079–1105.
- [5] Clementi, F., Nespeca, A. and Lenci, S. (2016) Seismic behavior of an Italian Renaissance Sanctuary: Damage assessment by numerical modelling, *Proceedings of the International Conference of Computational Methods in Sciences and Engineering 2016 - ICCMSE 2016*, Athens, Greece, 2016.
- [6] Jean, M. (1999) The non-smooth contact dynamics method, *Computer Methods in Applied Mechanics and Engineering* **177**, 235–257.
- [7] Moreau, J.J. (1988) Unilateral Contact and Dry Friction in Finite Freedom Dynamics, *Proceedings of the Nonsmooth Mechanics and Applications*, Springer Vienna, Vienna, 1988.
- [8] Pfeiffer, F. and Glocker, C. (1996) Multibody Dyn. with Unilateral Contacts, Wiley-VCH Verlag GmbH, Weinheim, Germany, 307–314.
- [9] Nordmark, A., Dankowicz, H. and Champneys A. (2009) Discontinuity-induced bifurcations in systems with impacts and friction: Discontinuities in the impact law, *International Journal of Non-Linear Mechanics* **44**, 1011–1023.
- [10] De Lorenzis, L. DeJong, M. and Ochsendorf, J. (2007) Failure of masonry arches under impulse base motion, *Earthquake Engineering & Structural Dynamics* **36**, 2119–2136.
- [11] Chetouane, B., Dubois, F., Vinches, M. and Bohatier, C. (2005) NSCD discrete element method for modelling masonry structures, *International Journal for Numerical Methods in Engineering* **64**, 65–94.
- [12] Rafiee, A., Vinches M. and Bohatier C. (2008) Modelling and analysis of the Nimes arena and the Arles aqueduct subjected to a seismic loading, using the Non-Smooth Contact Dynamics method, *Engineering Structures* **30**, 3457–3467.
- [13] Vasconcelos, G. and Lourenço, P.B. (2009) Experimental characterization of stone masonry in shear and compression, *Construction and Building Materials* **23**, 3337–3345.
- [14] Clementi, F., Gazzani, V., Poiani, M. and Lenci, S. (2016) Assessment of seismic behaviour of heritage masonry buildings using numerical modelling, *Journal of Building Engineering*, 29–47.
- [15] Clementi, F., Gazzani, V., Poiani, M., Mezzapelle, P.A. and Lenci, S. (2017) Seismic Assessment of a Monumental Building through Nonlinear Analyses of a 3D Solid Model, *Journal of Earthquake Engineering*, 1–27.
- [16] Clementi, F., Pierdicca, A., Milani, G., Gazzani, V., Poiani, M. and Lenci, S. (2018) Numerical model upgrading of ancient bell towers monitored with a wired sensors network, *Proceedings of the 10th International Masonry Conference*, Milano, Italy, 2018.

New computational algorithms for the Limit Analysis of large-scale 3D truss-frame structures

***Rosalba Ferrari¹, Giuseppe Cocchetti^{1,2}, and **Egidio Rizzi¹**

¹ Università degli Studi di Bergamo, Dipartimento di Ingegneria e Scienze Applicate
viale G. Marconi 5, I-24044 Dalmine (BG), Italy

² Politecnico di Milano, Dipartimento di Ingegneria Civile e Ambientale
piazza L. da Vinci 32, I-20133 Milano, Italy

*Presenting author: rosalba.ferrari@unibg.it

**Corresponding author: egidio.rizzi@unibg.it

Abstract

Two new computational algorithms for the Limit Analysis (LA) of large-scale 3D truss-frame structures recently proposed by the authors are reconsidered and adapted for a comparison prediction of the elastoplastic response of a strategic beautiful historic infrastructure, namely the Paderno d'Adda bridge, a riveted wrought iron railway viaduct that was built in northern Italy in 1889. The first LA algorithm traces a fully exact evolutive piece-wise linear elastoplastic response of the structure, up to collapse, by reconstructing the true sequence of activation of made-available plastic joints (as a generalization of plastic hinges), in the true spirit of LA. The second LA algorithm develops an independent kinematic iterative approach apt to directly determine the plastic collapse state, in terms of collapse load multiplier and plastic mechanism, based on the upper-bound theorem of LA. Specifically, the marvelous doubly-built-in parabolic arch of the bridge is analyzed, under a static loading configuration at try-out stage, and its elastoplastic response is investigated, in terms of evolutive load-displacement curve, collapse load multiplier and plastic collapse mechanism. The two LA algorithms are found to much effectively run and perform, despite for the rather large size of the computational model, with a number of dofs in the order of four thousand, by achieving good corresponding matches in terms of the estimate of the load-bearing capacity and of the collapse characteristics of the arch substructure, showing this to constitute a well-set structural element. Moreover, the direct kinematic method displays a rather dramatic performance, in truly precipitating from above on the collapse load multiplier and rapidly adjusting to the collapse mode, in very few iterations, by a considerable saving of computational time, with respect to the complete evolutive elastoplastic analysis. This shall open up the way for a further adoption of such advanced LA tools, with LA regaining a new momentum within the optimization analysis of structural design and form-finding problems.

Keywords: Limit Analysis (LA), evolutive elastoplastic response, kinematic (upper-bound) theorem, collapse load multiplier, plastic collapse mechanism, truss-frame structures, historic construction

1. Introduction

Limit Analysis (LA) constitutes by today a well-established and consolidated discipline, for evaluating consistent bounds on the collapse (limit) loads acting on engineering structures characterized by a mechanical behavior that may be idealized as perfectly plastic, subjected to constant permanent loads and increasing live loads.

LA may be considered as a milestone in the more recent history of structural mechanics and provides a rather powerful tool for structural design and assessment purposes, within a wide variety of engineering situations. It has acquired its rational formulation thanks to the contribution by Drucker et al. [1], who have formulated and demonstrated the fundamental theorems of LA for a continuum. More recent consolidated contributions, such as those by Massonet and Save [2], Kaliszky [3], Lubliner [4], Jirasek and Bazant [5], and many others, as those quoted in Maier et al. [6], have further made the theory and the methods of LA rather fundamental in various applications of mechanics of solids and structures, becoming by now classical references on the topic. LA of frames has also been revived in recent years, with several interesting applications (see e.g. Cocchetti and Maier [7], Tangaramvong and Tin-Loi [8], Lògò et al. [9], Skordeli and Bisbos [10], Bleyer and Buhan [11], Nikolaou et al. [12]), despite, basically being employed at an academic level, somehow still far from implementations within the engineering profession.

A new impetus toward the application of the LA discipline within structural engineering has been recently targeted by the authors by two different proposals, respectively in Ferrari et al. [13,14] and in Ferrari et al. [15], with reference to the structural elastoplastic analysis of large-scale 3D truss-frame structures. The effort was motivated by the goal of performing an effective LA of the collapse state of a strategic and historic truss-frame infrastructure, as described below. In former contribution Ferrari et al. [13,14], a computational algorithm for tracing the “exact” piece-wise linear evolutive elastoplastic response of the structure up to collapse in the true spirit of LA has been formulated and implemented. In latter attempt Ferrari et al. [15], a kinematic upper-bound LA direct method of computational analysis has been proposed, allowing to iteratively converge on the collapse state in terms of collapse load multiplier and plastic mechanism.

In the work herein presented, such two new approaches are employed toward the LA analysis of a large truss-frame substructure displaying a rather considerable complexity, namely a 3D model of the box-formed doubly-built-in parabolic arch of the Paderno d’Adda bridge (1889), a beautiful monumental infrastructure in the local territory. This roughly involves 1,050 beam finite elements with potentially active plastic joints and more than 4,000 degrees of freedom.

In the paper, it is shown that the kinematic algorithm rapidly converges onto the collapse state, as also traced by the full evolutive elastoplastic analysis, with a kinematic load multiplier that quickly precipitates from above on the sought collapse load amplifier and a plastic mechanism that rapidly adjusts to the true collapse one. The kinematic algorithm has proven to display a saving of more than 96% of the computational time employed for the evolutive elastoplastic analysis, which additionally reconstructs the whole sequence of activation of the plastic joints and traces a rather ductile piece-wise linear load-displacement curve, showing the arch to constitute a beautiful well-designed structural element, in terms of load-carrying capacity up to the limit state of plastic collapse.

The present paper is structured as follows. Section 2 briefly introduces the two developed LA algorithms. Section 3 illustrates the results achieved from the numerical analyses conducted on the truss-frame parabolic arch of the Paderno d’Adda bridge. Brief comments on the various computational aspects and the effectiveness of the performed simulations are concisely pointed out in the Conclusions.

2. Limit Analysis formulation

The two computational formulations herein adopted are based on the same underlying hypotheses, as briefly listed here. Perfectly-plastic joints have been located at the two edges of each beam finite element composing the model structure; a piece-wise linear yield domain,

specifically an uncoupled Rankine-type boxed-form yield domain in the space of the static variables (i.e. beam internal actions), is chosen. Linear kinematics is assumed, namely second-order non-linear geometrical effects are ruled out and equilibrium is enforced in the initial (unstressed) configuration of the structure. Permanent (i.e. gravity) loads and a set of (basic) loads amplified by a common (“load multiplier”) factor are considered.

Both formulations have been implemented within MATLAB[®], with specific computer programs exploiting 3D beam finite elements. Exact time integration is one of the main characteristics of the evolutive elastoplastic formulation (see e.g. Ferrari et al. [13,14]). The kinematic algorithm of Limit Analysis, discussed in Ferrari et al. [15], and prodromal references quoted therein, is based on a kinematic upper-bound approach and is below newly applied to a large-scale 3D truss-frame structure, for comparison purposes with the former evolutive elastoplastic algorithm.

2.1. Evolutive elastoplastic approach

The evolutive LA elastoplastic approach proposed in Ferrari et al. [13,14] relies on an iterative procedure in which the global elastoplastic matrix of the structure is iteratively updated, based on the plastic modes that become active during the increment of the applied live loads. In this way, the characteristic non-linear (piece-wise linear) load/displacement response curve is estimated, up to collapse.

A synthesis of the main characteristics of the computational algorithm can be outlined as follows. In the approach, the kinematic constraints on the structural system are imposed by enforcing the stationary condition of Lagrangian function \mathcal{H} :

$$\mathcal{H}(\mathbf{U}, \mathbf{R}) = \frac{1}{2} \mathbf{U}^T \mathbf{K} \mathbf{U} - \mathbf{F}^T \mathbf{U} - \mathbf{R}^T (\mathbf{C} \mathbf{U} - \mathbf{V}) \quad (1)$$

where vector \mathbf{U} collects all the nodal dofs of the system, matrix \mathbf{K} is obtained by a classical assembly procedure of the finite element stiffness matrices, and vector \mathbf{F} is created as the assembly of the equivalent nodal force vectors, including only the active loads. In Eq. (1), constraint equations $\mathbf{C} \mathbf{U} = \mathbf{V}$ are imposed: matrix \mathbf{C} depends only on geometrical quantities, vector \mathbf{V} collects possible given values of the kinematic quantities, vector \mathbf{R} gathers the so-called *Lagrangian multipliers*, which may be interpreted, from a mechanical point of view, as the reactions supplied by the constraints.

The solution of the structural system is derived from Eq. (1), and by applying a Gaussian elimination to the system of the constraint equations, in order to eliminate those, among the latter, that are just a repetition of others. If this occurs, the elimination procedure leads to a reduced system, from which the solution in terms of (independent) displacements (\mathbf{u}) can be symbolically given by the following expression:

$$\mathbf{u} = \mathbf{K}_{uu}^{-1} (\mathbf{F}_D + \mathbf{F}_u) \quad (2)$$

where \mathbf{K}_{uu} is a reduced stiffness matrix and \mathbf{F}_D and \mathbf{F}_u are reduced known-term vectors, derived through appropriate mathematical steps, from vector \mathbf{V} and vector \mathbf{F} , respectively (see Ferrari et al. [13] for the details).

The evolutive approach is based on an incremental procedure that can be outlined as follows:

- (i) At the beginning of the analysis, the structure is not under the effect of loading or of pre-existing stress states; then, there is no plasticization in the structure (no yield plane is active). Afterwards, the following quantity is set acting on the structure: $\Delta \mathbf{F}_{Li} = \Delta \alpha \mathbf{F}$, where $\Delta \mathbf{F}_{Li}$ represents the load increment, determined from \mathbf{F} through load factor $\Delta \alpha$,

which is always positive and can be assumed as a chronological variable in the increment of \mathbf{F} ;

- (ii) Given $\Delta \mathbf{F}_{Li}$, also quantity $\Delta \mathbf{F}_u$ can be determined. Then, substituting the latter into Eq. (2), the incremental solution in terms of displacements $\Delta \mathbf{u}$ is thus provided. It is worth to note that at the beginning of the procedure matrix \mathbf{K}_{uu} is totally elastic;
- (iii) Being $\Delta \mathbf{u}$ strictly related to the increment of the static internal variables of each finite element ($\Delta \mathbf{N}$), the latter can also be computed. Incremental solutions $\Delta \mathbf{u}$ and $\Delta \mathbf{N}$ do not take into account that some modes may be possibly activated during the step; on the other hand, it is possible that the extent of incremental load $\Delta \mathbf{F}_{Li}$ is not enough to determine the reaching of at least one internal static variable in a section on the corresponding yield limit. Through a comparison between each static internal variable at the beginning of the step and the corresponding yield limit, the procedure calculates for each mode, among all non-activated ones, the specific *scale factor* (β) of incremental load $\Delta \mathbf{F}_{Li}$ leading to possible activations. The minimum among all such estimated factors, $\gamma = \min\{\beta\}$, allows to get the *exact* extent of the current step to achieve new activation(s);
- (iv) Through coefficient γ , the original incremental solution is proportionally rescaled, and the static and kinematic quantities of the structure at the end of the step are updated;
- (v) At the end of the step some yield planes (at least one) are now activated and stiffness matrix \mathbf{K}_{uu} becomes elastoplastic. Its determination is computed through a convenient Gaussian elimination procedure, which represents a peculiar feature of the implemented evolutive approach (see Ferrari et al. [13] for a comprehensive description of the computational formulation).

The whole procedure assumes an iterative configuration in the repetition of steps (ii)-(v). It stops when the collapse of the structure is reached, namely when the minimum eigenvalue of global (updated) matrix \mathbf{K}_{uu} calculated within step (ii) vanishes (or it is below a given numerical tolerance) and the eigenvector is everywhere associated to a positive load dissipation (Ferrari et al. [13]).

2.2. Kinematic upper-bound direct method

A synopsis of the principal features of the computational algorithm outlined in Ferrari et al. [15] is recalled below.

The direct LA kinematic method is distinguished from other kinematic approaches due to its alternative and elegant formulation, which provides a very convenient procedure for the LA of truss-frames. According to the upper-bound (kinematic) theorem, the *kinematic load multiplier* is defined as follows:

$$\mu_k = \frac{\dot{L}_i - \dot{L}_e^g}{\dot{L}_e^0} \quad (3)$$

where \dot{L}_i represents the internal power dissipation of the structure, \dot{L}_e^g and \dot{L}_e^0 represent the power of the base live loads and the power produced by the permanent loads applied to the structure, respectively. A corollary of the upper-bound theorem states that the collapse load multiplier is the minimum among the kinematic load multipliers defined for whole set K of the kinematically admissible mechanisms, namely $\mu_c = \min_K \{\mu_k\}$. Being \dot{L}_i and \dot{L}_e homogeneous functions of order one with respect to the velocity field, $\dot{L}_e^0 = 1$ can always be

set for a kinematically admissible mechanism (M); thus, the collapse load multiplier can be obtained as the solution to the following constrained optimization problem:

$$\mu_c = \min_{M \in C} \left\{ \dot{L}_i(M) - \dot{L}_e^g(M) \mid \dot{L}_e^0(M) = 1 \right\} \quad (4)$$

If the minimization is limited to a subset of compatible mechanisms $\bar{C} \subseteq C$, it results:

$$\mu_c \leq \mu_k = \min_{M \in \bar{C}} \left\{ \dot{L}_i(M) - \dot{L}_e^g(M) \mid \dot{L}_e^0(M) = 1 \right\} \quad (5)$$

For instance, with reference to the structure discretization, this can happen when the collapse mechanism requires an active plastic joint within one or more beam elements, besides the plastic joints at the beam ends.

The procedure proposed by Ferrari et al. [15] exploits a convenient reshaping of the internal power dissipation description of a beam element, in order to adapt an existing efficient approach originally presented by Zhang et al. [16], for the LA of continua, within the classical FEM analysis based on beam element discretization. In particular, the procedure adopts a *quadratic form* of the internal power dissipation of a beam element, whose expression turns out to be defined as follows:

$$\dot{L}_i(\dot{\mathbf{q}}) = \dot{\mathbf{q}}^T \mathbf{S}(\dot{\mathbf{q}}) \dot{\mathbf{q}} \quad (6)$$

where vector $\dot{\mathbf{q}}$ lists the nodal velocities of the structure and matrix $\mathbf{S}(\dot{\mathbf{q}})$ is a function of: (a) the generalized plastic strains of all the beams that form the structure (the axial and the transverse relative rotations and the axial relative displacement, for each joint); (b) the corresponding internal actions activating the plastic strains (torque, two bending moments and axial force, for each joint). Matrix $\mathbf{S}(\dot{\mathbf{q}})$ thus governs the global internal power dissipation of the structure; it is symmetric, highly sparse (even narrow-banded with a proper numbering of dofs) and positive definite (or positive semidefinite when non-dissipative rigid-body motions are allowed by constraints), namely the same properties characterizing the classical global stiffness matrix for an elastic FEM frame analysis.

The direct LA kinematic method is based on an iterative procedure that can be outlined as follows:

- (i) At the beginning of the analysis, vector $\dot{\mathbf{q}}_0$ is set up by a vector of random numbers and normalized in order to guarantee that $\dot{L}_e^0(\dot{\mathbf{q}}_0) = 1$;
- (ii) Let one now consider iteration $n+1$. Vector of nodal velocities $\dot{\mathbf{q}}_n$ relevant to the n -th iteration (still normalized in order to guarantee that $\dot{L}_e^0(\dot{\mathbf{q}}_n) = 1$) and corresponding load amplification factor μ_k^n are assumed to be known, the last one computed through Eq. (3). Then, the iterative process leads to generate a new mechanism, governed by new vector $\dot{\mathbf{q}}_{n+1}$ that is obtained as the solution of the quadratic constrained minimization problem in Eqs. (5)-(6), in which the internal power dissipation, given by Eq. (6), is computed through coefficient matrix $\mathbf{S}_n = \mathbf{S}(\dot{\mathbf{q}}_n)$, related to previous iteration n . The resulting constrained minimization problem can be solved by enforcing the stationarity condition on a Lagrange function, in this way leading to a set of two equations, each one to be a function of an unassigned variable, namely $\dot{\mathbf{q}}$ and λ , respectively, being the latter a *Lagrangian multiplier*;

- (iii) Ones variables $\dot{\mathbf{q}}_{n+1}$ and λ_{n+1} are calculated, according to Eq. (3) a new kinematic load multiplier (μ_k^{n+1}) follows, as well as coefficient matrix \mathbf{S}_{n+1} governing the global internal power dissipation of the structure.

The whole procedure assumes an iterative configuration in the repetition of steps (ii)-(iii). It stops when: (a) the number of “inactive” modes has to remain constant for 10 iterations; (b) the relative change of the load multiplier from a step to another is lower than 10^{-3} .

In Ferrari et al. [15] a comprehensive description of the computational formulation is provided, along with the proof of convergence of the iterative procedure and some interesting remarks concerning the influence of computing round-off errors on the effectiveness of the algorithm. The latter turns out to be essential in order to deal with 3D large-scale macro-structures, endowed with a large number of dofs and associated potentially active plastic joints, as the one herein presented as a benchmark structure.

3. Large-scale structural numerical simulations

3.1. Benchmark structure

The benchmark structure considered in the present paper is represented by the parabolic arch of the Paderno d’Adda bridge, an iron railway viaduct that crosses the Adda river between Paderno d’Adda (LC) and Calusco d’Adda (BG) at a height of approximately 85 m from water. It allows for connecting the provinces of Lecco and Bergamo, near Milano, in Lombardia, northern Italy (Figure 1). The viaduct, co-aged of most celebrated Tour Eiffel, and rather similar to that in contemporary structural features, was quickly constructed between 1887 and 1889 and is one of the very first great iron constructions designed through the practical application of the so-called “Theory of the Ellipse of Elasticity” (Ferrari and Rizzi [17]), a graphic-analytic method of structural analysis that was developed in the 19th century. The bridge is made of a wrought iron material; the structural elements are interconnected by riveted joints. It is composed of: an elegant and robust doubly-built-in parabolic arch of about 150 m of horizontal span and 37.5 m of vertical rise; vertical bearing piers with a height up to 31.5 m; an upper continuous beam of 266 m of length. The main upper continuous beam, 5 m wide and 6.25 m high, is formed by a metallic truss-frame supported by nine bearings, four of which are supported by the arch. Wide documentations on the historic bridge are available in Ferrari and Rizzi [17], Ferrari et al. [14] and references quoted therein.

For the finalities of the present paper, the homemade FEM model of the arch of the bridge has been extracted from a preassembled FEM model of the bridge, earlier implemented within commercial FEM code ABAQUS®. It has been constituted by assembling a 3D truss-frame with beam elements, mutually built-in at the nodes. The 3D truss-frame of the arch consists of two planar parabolic trusses laying into two inclined planes (of an angle $\alpha \approx \pm 8.63^\circ$ to the vertical), symmetrically located with respect to the vertical longitudinal plane of the viaduct. The inclined planes are placed at a relative distance of 5.096 m at the keystone. A single arch profile is considered in each inclined plane, with an arch body that accounts for the true presence of two secondary twin inclined arches, on the two sides of such an inclined plane. The truss nodes are linked to each other through a reticular system that corresponds to the true bracing system of the arch. The arch presents some additional reinforcing plates between the vertical bars in each of the secondary twin inclined arches, placed at the locations of the arch/bearing connections. These plates have not been explicitly represented in the FEM model

of the structure. To comply with this, the cross sections of the limiting vertical bars contouring the stiffening plates have been endowed with larger geometrical characteristics. The FEM model of the arch of the bridge is comprised of 1051 beam elements and 342 nodes and is endowed with the following material properties, characteristic of a wrought iron material (SNOS [18], Nascè et al. [19]): Young's modulus $E=17 \cdot 10^6 \text{ tons/m}^2 \approx 170 \text{ GPa}$; Poisson's ratio $\nu=0.3$ (corresponding shear modulus $G=6.54 \cdot 10^6 \text{ tons/m}^2 \approx 65.4 \text{ GPa}$); mass density $\rho=7.7 \text{ tons/m}^3$. Despite that yield limit characteristics of the material could be set to those that may be typically reported for an iron material (Nascè et al. [19]), the yield limit assumed in the present computations has been conservatively taken coincident with the allowable working stress imposed by the builder at design stage (SNOS [18]), namely $\sigma_y=6.00 \text{ kg/mm}^2 \approx 60 \text{ MPa}$ (and $\tau_y=\sigma_y^{1/3}=3.46 \text{ kg/mm}^2 \approx 34.6 \text{ MPa}$). The FEM model of the arch of the Paderno d'Adda bridge is depicted in Figure 2, which forms the basis for the elastoplastic model with potentially active plastic joints located at the extremes of each beam finite element.



Figure 1: Contemporary view of the Paderno d'Adda bridge (1889) from the right bank.

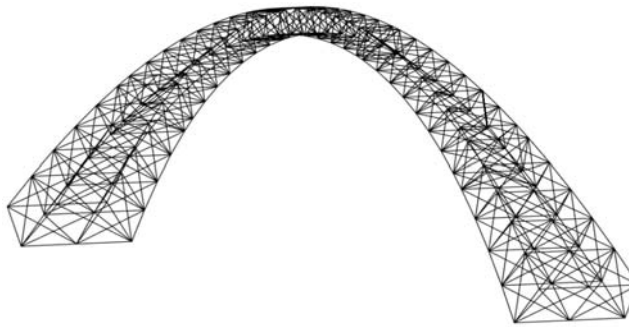


Figure 2: FEM model of the arch of the Paderno d'Adda bridge.

3.2. Loading configuration

The loads considered in the performed numerical simulations have been set with reference to the static try-out railway loading conditions on the Paderno d'Adda bridge (SNOS [18], Nascè et al. [19]). These tests were carried-out using locomotives with tender, each of 83 tons of weight, corresponding to a uniformly distributed load on the beam of about $q=5.1 \text{ tons/m}$. Such distributed live load leads to elastic reactions on the arch, at its interfaces with the piers and the upper beam, equal to $P_1=163.5 \text{ tons}$, $P_2=171.3 \text{ tons}$, $P_3=168.7 \text{ tons}$, $P_4=171.3 \text{ tons}$, for

a total live load of $Q=674.8$ tons just on the arch (Figure 3). These values of P_i come from the elastic solution of a nine-bearing continuous beam, with firm supports, and are here assumed to be representative of the load distribution directly lying on the arch, when a uniform load is acting on the above upper continuous beam (Ferrari et al. [13]).

In the numerical simulations, only above-mentioned live loads P_i and associated total load Q have been considered to be affected by a load multiplier. The self-weight of the structure has instead been treated as a pre-imposed permanent load; namely, this has not been affected by the load multiplier. In particular, the self-weight of the upper continuous beam and the piers on the arch has been considered through loads directly applied to the arch at the arch/beam and arch/piers interfaces. To determine such permanent loads, the upper truss structure has been considered as a continuous beam supported by nine bearings, subjected to a uniformly-distributed load representing the weight of the upper beam and set equal to 6.61 tons/m (SNOS [18]). Therefore, vertical reaction forces at the bearings located at the level of the arch/piers and arch/beam interfaces have been calculated. The two reactions at the level of the arch/beam interfaces (P_2 and P_3 in Figure 3) have then been directly applied to the arch of the bridge, meanwhile the two reactions at the level of the arch/piers interfaces (P_1 and P_4 in Figure 3) have been incremented by the weight of the piers resting on the arch, given equal to 22 tons (SNOS [18]). Permanent resultants P_1 - P_4 applied to the arch, due only to the self-weight of the upper continuous beam and the piers on the arch (not affected by the load multiplier), therefore result equal to $P_1=233.9$ tons, $P_2=222.0$ tons, $P_3=218.7$ tons, $P_4=244.0$ tons.

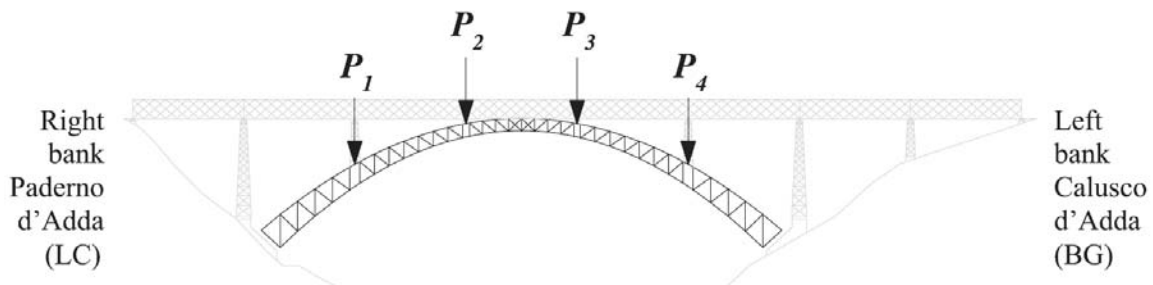


Figure 3: Scheme of the considered static load configuration (view from down-stream), with loads P_i applied on the arch structure at the arch/piers and arch/beam interfaces.

3.3. Numerical results

The results in terms of elastoplastic load-carrying capacity of the bridge's arch obtained with the above-discussed (Section 2) numerical algorithms of Limit Analysis are consistently shown in Table 1. Moreover, the mechanical elastoplastic response of the arch substructure is depicted in Figures 4-6.

Table 1: Collapse load multipliers for the benchmark structure.

Method	Collapse load multiplier	Limit state
Evolutionary	4.17	First null eigenvalue of stiffness matrix
Kinematic	4.22	Formation of a collapse mechanism

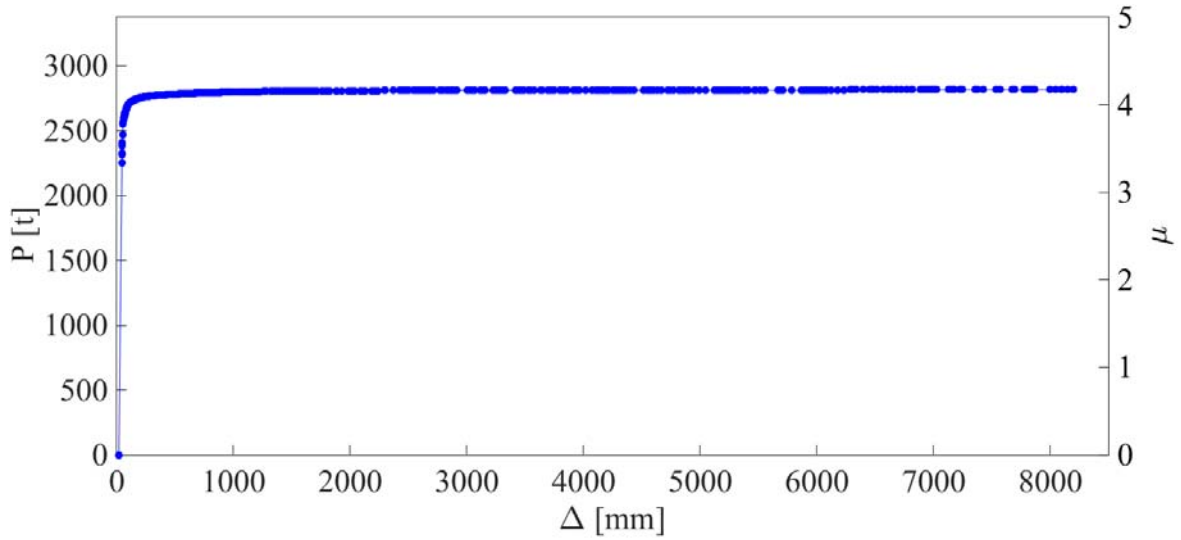


Figure 4: Representation of the characteristic piece-wise linear load-displacement response curve.

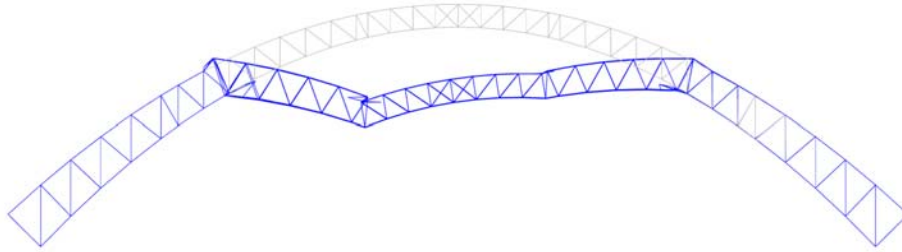


Figure 5: Collapse mechanism of the arch of the Paderno d'Adda bridge by the kinematic algorithm.

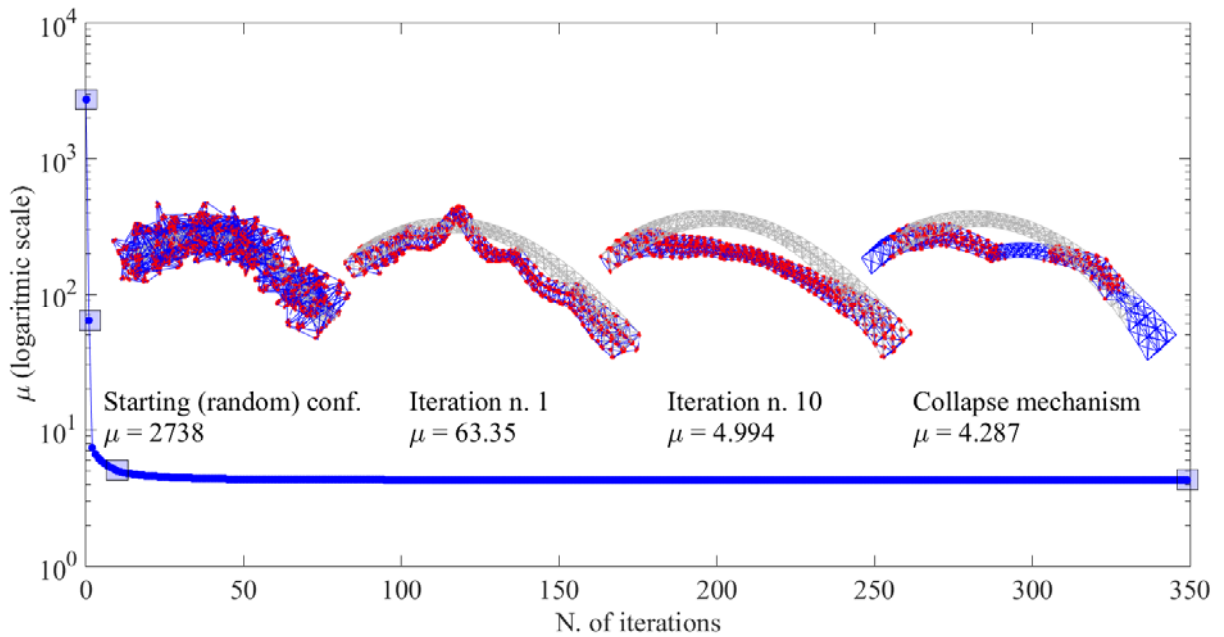


Figure 6: Collapse load multiplier and plastic mechanism estimation along the kinematic iterative procedure.

The numerical evaluation of the collapse load multiplier reported in Table 1 shows a minor discrepancy between the upper-bound value obtained by the kinematic method and the estimated value achieved by the evolutive analysis (a static evaluation, in essence, and as set at the very earliest singularity of the tangent stiffness matrix). This difference (likely forming a fork delimiting the real numerical collapse load multiplier) may be due to round-off errors and adopted numerical tolerances within the kinematic approach, possibly being linked to the diffused amount of very small plastic deformations in several deactivated joints, along the iterations. This may require further inspections and refinements of the kinematic algorithm, in order to handle such large-scale 3D truss-frame structures.

Figure 4 reports the characteristic piece-wise linear load/displacement response curve of the arch obtained from the evolutive algorithm (Section 2.1), showing a considerable global structural ductility. In particular, the horizontal axis depicts the vertical displacement (Δ) of the node that, in the end, has shown the maximum displacement at incipient collapse; the left vertical axis reports total amplified live load $P=\mu Q$, where μ is the load multiplier related to the incremental solution of the non-linear elastoplastic analysis; the right vertical axis also depicts load multiplier μ . Therefore, in the plot, the response curve for the loading configuration can be read on the left or on the right axis, equivalently. The end point of the $P-\Delta$ curve refers to the so-estimated “exact” collapse of the structure in terms of first vanishing minimum eigenvalue of the tangent stiffness matrix of the structure (Section 2.1). Notice that in this plot the non-zero initial displacement (at $P=0$) is due to self-weight only, namely to a pre-imposed permanent load not affected by the load multiplier. In this case, the computing time to achieve the collapse solution was about 1100 s. The algorithm has been implemented and run as a non-compiled code within MATLAB®, under a Windows 10 operating system, on a Dell laptop endowed with an Intel Core i7-6500U Processor, clock at 2.50 GHz and 16 GB RAM.

Figure 5 also depicts the deformed configuration of the structure at incipient collapse (namely, the plastic collapse mechanism), as obtained by the kinematic LA algorithm (Section 2.2), which is rather consistent with that derived from the evolutive program (Section 2.1).

Figure 6 further illustrates the results related to the kinematic algorithm (Section 2.2). It displays the collapse load multiplier computed during the iterative procedure, and the associated estimated collapse mechanism along the iterations. In the picture, it is possible to appreciate how the implemented algorithm is capable to quickly and easily achieve convergence. In fact, the collapse mechanism is almost already achieved at the tenth iteration, and in already about 40 iterations the kinematic load multiplier precipitates on the collapse one; after that, the load multiplier turns out to be almost flat, at an increasing number of iterations. In this case, the computing time to achieve the collapse solution by the direct kinematic method (in about 350 iterations) was nearly 37 s, on the same computer platform as described above, thus with a saving of more than 96% of the computational time employed for the corresponding complete evolutive elastoplastic approach (1100 s).

Conclusions

The present paper has attempted a comprehensive computational elastoplastic structural analysis, truly loyal to the principles of Limit Analysis, in the context of large-scale 3D truss-frame structures, with reference to the elastoplastic response of a strategic historic infrastructure within the local territory (Paderno d’Adda bridge, 1889), in order to investigate its potential performance in terms of LA limit states, with focus on the supporting parabolic arch.

Two approaches have been employed, compared and cross-connected, in view of deciphering the resources of the fundamental arch-bearing mechanism of the iron bridge, toward the limit state of plastic collapse, as follows: a reconstruction of the complete evolutive response up to collapse, with the sequence of activation of the plastic joints in the various members of the boxed-form arch structure and the associated tracing of the global load-displacement curve; a direct determination of the collapse characteristics, in terms of collapse load multiplier and plastic mechanism.

Both LA methods of analysis consistently run and effectively perform, despite for the rather considerable number of dofs considered within the computation (more than 4,000). A true match is displayed for the predictions of plastic collapse, with the kinematic iterative algorithm showing an impressive performance in really precipitating from above onto the collapse load multiplier, and with a plastic mechanism that rapidly adjusts, in very few iterations, to the collapse mode.

As for the interpretation of the achieved results on such a crucial and beautiful historic infrastructure, the arch is revealed by both methods to constitute a well-set fundamental structural element, with considerable resources in terms of plastic collapse, testifying the mastering of the methods of structural design and execution by the art of metallic carpentry at the time, despite for being directly conceived on the basis of (graphic-analytic) elastic methods of analysis only (see Ferrari and Rizzi [17]). This brings really good news, at present stage, in terms of debating possible future destinations of the monumental infrastructure.

These new LA computational methodologies may open up the way for LA to regaining a new, considerable momentum in the structural analysis of large-scale structures, of both a new design concept and historic, to be preserved, ones, and with new challenging targets in terms of structural optimization, as for structural design or form-finding quests, given the prompt availability of the mechanical calculation, within a possible optimization loop that may consider variable structural characteristics and trace (optimize) the corresponding variation of the elastoplastic response and the attached collapse features.

References

- [1] Drucker DC, Prager W, Greenberg HJ (1952) Extended limit design theorems for continuous media. *Quarterly of Applied Mathematics*, 9(4):381–389.
- [2] Massonet CH, Save M (1978) In: Nelissen, B. (Ed.), *Calcul Plastique des Constructions*. Angleur, Liège, Belgium.
- [3] Kaliszky S (1989) *Plasticity: Theory and Engineering Applications*, Elsevier, Amsterdam, Netherlands.
- [4] Lubliner J (1990) *Plasticity Theory*, McMillan Publisher, New York.
- [5] Jirasek M, Bazant ZP (2002) *Inelastic Analysis of Structures*, Wiley, Chichester, U.K.
- [6] Maier G, Carvelli V, Cocchetti G (2000) On direct methods for shakedown and limit analysis. *European Journal of Mechanics - A/Solids*, 19(Special Issue - 4th EUROMECH Solid Mechanics Conference, Plenary Lectures, Metz, France, June 26-30, 2000), S79–S100.
- [7] Cocchetti G, Maier G (2003) Elastic-plastic and limit-state analyses of frames with softening plastic-hinge models by mathematical programming. *International Journal of Solids and Structures*, 40(25):7219–7244.
- [8] Tangaramvong, S., Tin-Loi, F. (2007) A complementarity approach for elastoplastic analysis of strain softening frames under combined bending and axial force. *Engineering Structures*, 29(5), 742–753.
- [9] Lògò J, Kaliszky S, Hjiat M, Movahredi Rad M (2008) Plastic Limit and Shakedown Analysis of Elastoplastic Steel Frames with Semirigid Connections. In Proc. of *Design, Fabrication and Economy of Welded Structures*, Eds. Károly Jármai and József Farkas, Miskolc, Hungary, April 24-26, 2008, pp. 237–244.
- [10] Skordeli M-AA, Bisbos CD (2010) Limit and shakedown analysis of 3D steel frames via approximate ellipsoidal yield surfaces. *Engineering Structures*, 32(6):1556–1567.

- [11] Bleyer J, Buhan P (2013) Yield surface approximation for lower and upper bound yield design of 3D composite frame structures. *Computers and Structures*, 129(December 2013):86–98.
- [12] Nikolaou KD, Georgiadis K, Bisbos CD (2016) Lower bound limit analysis of 2D steel frames with foundation-structure interaction. *Engineering Structures*, 118(1 July 2016):41–54.
- [13] Ferrari R, Cocchetti G, Rizzi E (2016) Limit Analysis of a historical iron arch bridge. Formulation and computational implementation. *Computers and Structures*, 175(2016):184–196, doi: 10.1016/j.compstruc.2016.05.007.
- [14] Ferrari R, Cocchetti G, Rizzi E (2017) Computational elastoplastic Limit Analysis of the Paderno d'Adda bridge (Italy, 1889). *Archives of Civil and Mechanical Engineering*, 18(1):291–310, doi:10.1016/j.acme.2017.05.002.
- [15] Ferrari R, Cocchetti G, Rizzi E (2018) Effective iterative algorithm for the Limit Analysis of truss-frame structures by a kinematic approach. *Computers and Structures*, 197(15 February 2018):28–41, doi:10.1016/j.compstruc.2017.11.018.
- [16] Zhang P, Lu M, Hwang K (1991) A mathematical programming algorithm for limit analysis. *Acta Mechanica Sinica* 7(3):267–74.
- [17] Ferrari R., Rizzi E. (2008) On the theory of the ellipse of elasticity as a natural discretisation method in the design of Paderno d'Adda Bridge (Italy). Chapter 66 in *Structural Analysis of Historic Construction – Preserving Safety and Significance*, In Proc. of 6th International Conference on Structural Analysis of Historic Construction (SAHC08), D. D'Ayala and E. Fodde (Eds.), Bath, UK, July 2–4, 2008, CRC Press, Taylor & Francis Group, London, print ISBN: 978-0-415-46872-5, pp. 583–591; eBook ISBN: 978-1-4398-2822-9, doi:10.1201-9781439828229.ch66.
- [18] Società Nazionale delle Officine di Savigliano (1889). *Viadotto di Paderno sull'Adda (Ferrovia Ponte S. Pietro-Seregno)*. Torino: Tip. e Lit. Camilla e Bertolero.
- [19] Nascè V, Zorgno AM, Bertolini C, Carbone VI, Pistone G, Roccati R (1984). Il ponte di Paderno: storia e struttura - Conservazione dell'architettura in ferro. *Restauro*, Anno XIII, n. 73-74, 215 pages.

Surface discretization based on bionic patterns in search of structural optimization

†Ewelina Gawell¹, *Anna Nowak², Wiesław Rokicki³ and Anna Stefańska⁴

¹ Department of Structure Design, Construction and Technical Infrastructure, Faculty of Architecture, Warsaw University of Technology, Poland

² Department of Structure Design, Construction and Technical Infrastructure, Faculty of Architecture, Warsaw University of Technology, Poland

³ Department of Structure Design, Construction and Technical Infrastructure, Faculty of Architecture, Warsaw University of Technology, Poland

⁴ Department of Structure Design, Construction and Technical Infrastructure, Faculty of Architecture, Warsaw University of Technology, Poland

*Presenting author: anna.patrycja.nowak@gmail.com

†Corresponding author: gawellewelina@gmail.com

Abstract

The division of surface becomes an important element in the pragmatics of creating structural elevations that simultaneously function as an external barrier. The discretization of freely formed surfaces seems to be particularly interesting. Most of the existing solutions are concerned with the shaping of elevation panels. However, in structural searches a greater interest is placed in algorithmic generation of surface divisions, where the use of appropriate method becomes a significant factor in surface discretization. This enables the pursuit of synergistic solutions which optimize both the architectural and structural parameters. As a result one can expect to obtain optimized structural surfaces shaped according to the chosen criteria. In this regard, nature provides an interesting field of research for effective patterns. Biological structures shaped under the influence of acting loads are optimized through evolution on the basis of, among others, minimisation of material and energy consumption. A number of forms observed in the natural world provide interesting examples for observing morphogenetic processes. The implementation of biological patterns in the design of load-bearing structures affects the design of structural forms in contemporary architecture in an interesting way. The article presents the results of model tests conducted on selected grid systems based on structures found in nature.

Keywords: bionic, structural surfaces, structural optimization, discretization of surface, architecture optimization

Introduction

In the design of modern architecture, mathematical methods and algorithms play an increasingly important role, enabling, among others, imitation of patterns taken from nature. The pursuit of synergistic solutions becomes an important element of bionic exploration, especially in the interdisciplinary design environment. In the search for structural solutions in architecture, shape optimization as well as surface discretization methods are interesting research areas. Contemporary digital modeling tools provide new opportunities for the rationalization of technical solutions through the design of the tectonics of architectural forms – thus realizing the idea of cooperation between architecture and structural design. This paper analyzes different variants of canopies' surface division with the use of bionic patterns. The research was aimed at finding effective solutions for the design of structural surfaces.

Methods of structural form's surface discretization

In adopting the method of surface discretization of structural forms it is important to know regular tessellations¹, which are used to determine the division of geometrically complex surfaces and to obtain planar elements. Tessellations can be achieved as a result of geometric transformations of surfaces such as translations, rotation and reflection, and glide reflection - when the surface or space is filled with a repeating geometric motif. The characteristic roof surface of the Sydney Opera House is an example of polygonal tessellation. The spectacular project by Jorn Utzon (created in the years 1956-1973) was the first in which segmental approximation of spherical surfaces with different radii was applied. Structural divisions occurring in nature are characterized by a more complex structure, and what is important, it is possible to find interesting patterns in terms of rationalization of material consumption. The forms observed in the natural world constitute an interesting model of morphogenetic processes - biological development, the aim of which is to create a structure in the context of existing environmental conditions, material, adaptation to live loads, etc. An important factor in the development of organic structures is the minimization of material and energy consumption, which is another significant analogy to the concept of sustainable development in architecture. Parametric modeling, which appeared alongside the possibility of implementing bionic patterns in the field load-bearing structures' design, simultaneously allows to generate diverse variants with the same boundary conditions and to verify the solutions on the basis of the adopted criteria. The Finite Element Method (FEM), which is an advanced numerical method of solving systems of differential equations, uses digital tools in the design and discretization of the surface. Equations of this type most often describe the phenomena and processes known in nature, and in recent years, along with the development of biomimicric trends, are used more and more often in architectural design. In the case of grid structures, FEM divides the structure into finite elements by means of nodal points - in the newly created record, certain geometrical, physical and mechanical features (so-called shape functions) are simultaneously assigned, thus creating a digital model of the structure [3]. Interesting examples of surface optimization using FEM can be found in the projects of the Italian architectural studio Studio Fuksas (Massimilian and Doriana Fuksas). One of the most interesting objects in which the designers used digital triangulation is the 'My Zeil' Shopping Center in Frankfurt located in the PalaisQuartier complex (which has the additional functions of an office and a hotel). The curvilinear surface constituting both the roofing of the building and its frontal elevation was subject to tessellation. From the main entrance to the shopping center there is a characteristic "dent" in the elevation, whose geometry, through the shape of an irregular tube, transforms into a waving roof of commercial spaces. An organic roof with an area of approximately 6000m² was made of a triangular steel mesh (made up of around 3,200 triangles) filled with glass and metal panels. The Knippers Helbig office from Stuttgart was the author of the structure. Another example of the use of digital tools supporting the optimal creation of grid structures by imitating nature can be seen in the Dynamic Relaxation Method. By using Newton's second law, Dynamic Relaxation generates a catenary model for a given point grid - it is possible to search for the optimal number of nodes, and to find the most effective position. Examples of the use of Dynamic Relaxation in surface discretization can be found in Foster & Partners' projects, among others: the

¹ Gawell E., Nowak A., Rokicki W., „Aperiodic tessellations in shaping the structural surfaces in the contemporary architecture”, The Journal Biuletyn of Polish Society for Geometry and Engineering Graphics, Volume 26 (2014), p. 47-54, Publishing House.: Polskie Towarzystwo Geometrii i Grafiki Inżynierskiej (Polish Society of Geometry and Engineering Graphics)

"So far, 17 periodic tessellations have been described, usually composed of regular polygons such as: equilateral triangle, a square, a hexagon, an octagon and a decagon."

Smithsonian Institution in Washington, or at the *British Museum's Great Court* in London. The more irregular divisions built on the basis of structural patterns found in nature are Voronoi diagrams, describing the optimal division of the surface into cells (or planes on convex polygons). The characteristic, polygonal Voronoi fields are created for the indicated points, taking into account their location and value (eg size). One Voronoi cell is created around each point as a polygon adhering to neighboring cells, and the points around which the diagram is created are exactly in the center of gravity of each cell². In addition, the Voronoi Diagram is a dual graph for Delaunay triangulation - in effect, the Voronoi polygon vertices are the centers of circles described on the triangles that form the triangular grid. As a result of the algorithm based on the Voronoi diagram it is possible to create various solutions in the field of surface discretization. Such a method of surface division was used for the *Landesgartenschau Exhibition Hall* project, which was implemented as a research project in the Institute for Computational Design, the Institute of Building Structures and Structural Design and the Institute of Engineering Geodesy at the University of Stuttgart. The design was shaped like the skeleton of a sea urchin, and the structural surface divisions were optimized using numerical analyzes.

Own research for selected grid models

The search for the geometrical shape of structural grids is an important factor in the design of architecture, both reflecting the intended artistic effect as well as rationalizing technical solutions. From the point of view of the solid's tectonics the surface grid division affects the visual reception of the form. Determining the metrics of the distribution is crucial for the realization of the curves describing the form such as in the case of free-form architecture. Increasing or decreasing the density of the grid is fundamental to structural optimization – too many elements cause larger deformations and increase stress due to own weight etc. Obtaining effective structural grid divisions should be addressed interdisciplinarily when analyzing paradigms from two disciplines: architecture and construction. Such phenomena can be observed naturally in the development processes of living organisms. Thanks to the algorithmization of tools in the design process such a pattern can be adopted in the creation of eco-efficient solutions for contemporary architecture. A recently popularized example can be observed in the use of Delaunay triangulation in the construction of grid structures (by using multi-variant solutions for the indicated boundary parameters). Choosing one grid system becomes a difficult task, requiring the ability to rationalize technical solutions in search for artistic effects. The results of the grid structure analysis are described below its construction was modeled on a bionic model. It assumed that the catenary was set on rectangular plan with the 30.0m x 26.0m dimensions, with three supports – the minimal number of supports guaranteeing stability of the structure while reducing any unnecessary geometry (Fig. 1). The supports were set with the assumption that the proportions between the individual support and cantilever spans were fixed. The supports were positioned in such a way that the lines passing through them are always in the 1:3:1 ratio, so the cantilevers are 6.0 m long and the distance between them is 18.0 m (according to the scheme – Fig. 2). Due to the varying length of the bars, the roof surface was discretized and the divisions were based on Delaunay triangulation. (Fig. 3). The idea of search for the curvilinear forms was based on a catenary model, by carrying out transformations in the third dimension for each of the metric variants. The adopted five-step curvature of the surface was dependent on the variable height in the

² On the basis of the Voronoi diagram, two-dimensional space is divided so that for a given set of n points, the plane is divided into n areas in such a way that each point in any area is closer to a specific point from the set of n points than from the other $n-1$ points. Voronoi cells, being an intersection of half-planes, are convex polygons whose collection breaks the two-dimensional Euclidean space creating an optimal net built of nodal points.

proportion of the curvature height to the support span equal to 1/8, 1/6, 1/5, 1/4, and 1/3 (Fig. 3). Models were generated in the Rhinoceros program with the Grasshopper and Kangaroo2 plugins. Determining the basic permutation for planar systems required an assumption regarding the divisions (due to the maximum bar length) – the transformations of the curvilinear variants occurred only in the third dimension, while maintaining the XY axis division of the structure. Additionally, for all examined systems, the location and the fixing of supports (restraints) were assumed to be constant. In addition, homogeneous THEX profiles (hexagonal tubes), TRON (round tubes) and TREC (rectangular tubes) made of S335 steel were used. Due to the initial assumptions, it was assumed that the analyzed structure variants will be homogenous, hence each construction bar will have the same cross-section. To obtain comparable results, the cross-section database was limited to closed sections. Because all variants have a similar geometry (after projecting onto a flat plane), the maximum permissible deformation of elements equal to 17.9cm was assumed. Due to the fact that the geometry of the analyzed structures is a model study, the analyzes were carried out by taking into account the own mass of the structure, wind and snow load according to EC, and the assumed load of 1.0 kPa for the coating material.

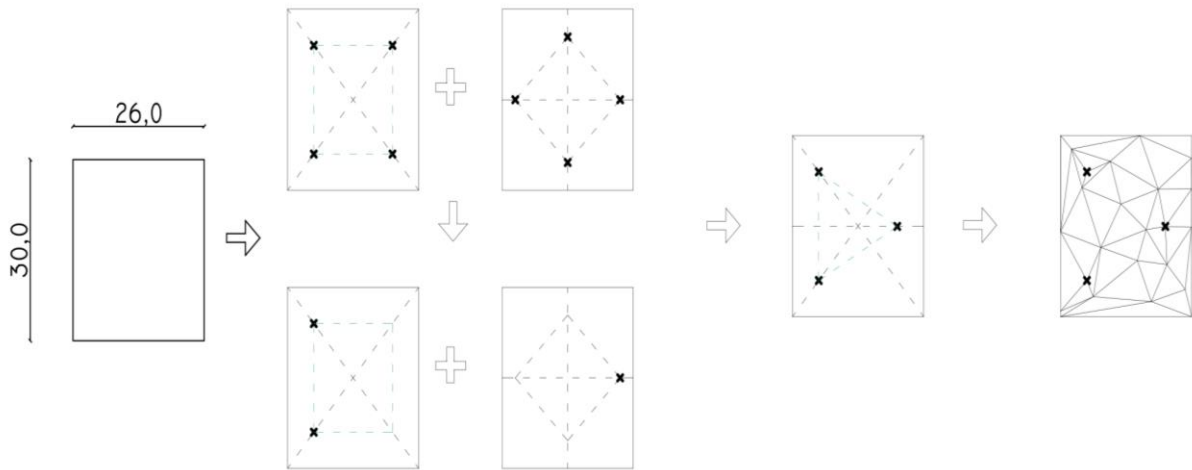


Fig. 1. The principle of setting the roofing supports

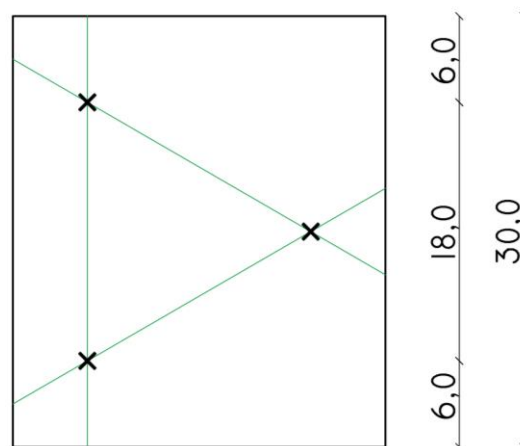


Fig. 2. Scheme of supports' placement – top down view of the model

Planar grid structures – discretization of the Surface

The use of triangular divisions allowed for the creation of a geometrical model, divided into the simplest elements connected at the nodes. Computer aided engineering calculations were used for the structural analyzes using the finite element analysis. Defining the grid metric was the first step of research. We analyzed 3 random planar variant solutions with various degrees of grid density (Fig.3):

- **Variant 1** – maximum bar length = 3,5m
- **Variant 2** – maximum bar length = 4,0m
- **Variant 3** – maximum bar length = 4,5m

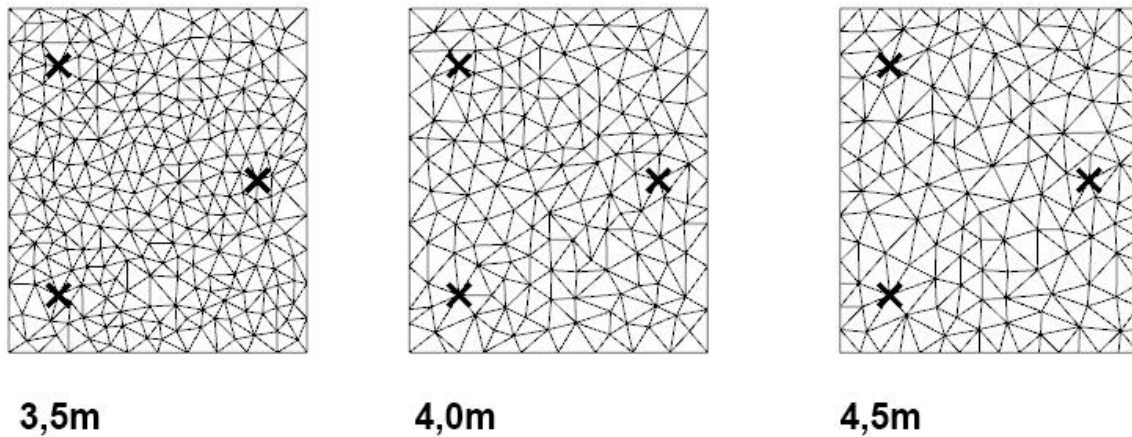


Fig. 3. Juxtaposition of the planar variants determining the grid metrics for curvilinear divisions

The planar systems achieved poor computational results and were thus only used to guide further research. From the 3 analyzed planar variants, the one with the bar length not exceeding 4,5m proved to be the most effective. Thus, further geometrical deformations were made by using this metric.

Space grid structures – shape optimization

In the search for structural efficiency due to the accepted minimum mass criterion, the surface was transformed using Dynamic Relaxation in the Grasshopper / Kangaroo2 plugin. Spatial models were created as a result of moving the center point of the plane on the Z axis (the center of gravity between the supports), while maintaining a constant horizontal projection of structural divisions. In the analyzed variants, the ratio of height to support span was assumed as constant:

- **Variant 4** – 1/8 ratio
- **Variant 5** – 1/6 ratio
- **Variant 6** – 1/5 ratio
- **Variant 7** – 1/4 ratio
- **Variant 8** – 1/3 ratio

By creating catenary models with different heights for the selected planar layout (with a maximum bar length of 4.5m), 5 curvilinear solutions were obtained (Fig.4). It was noticed that as the proportion between the height and the span of curvature between the supports increased, the total mass of the structure decreased (Tab.1). Although the curvature of the surface resulted in an increase of the total length of the bars, the mass of the most effective

Variant 8 (1/3) was smaller by 45% in relation to the heaviest Variant 4 (1/8). In addition, the bar cross-section has also decreased (the difference between the two extremes variants is 33%), making the structure more attractive in terms of aesthetics, which affects the reception of architecture (Fig. 6a, c). Due to the popularization of formative building technologies, such as the printing of construction and building elements, the parameter calculated as the ratio of total weight to the square meter area [kg/m^2] is significant (Tab.1). Curving the surface up to a 1/3 ratio would make it possible to create twice as effective structures as by using the 1/8 ratio.

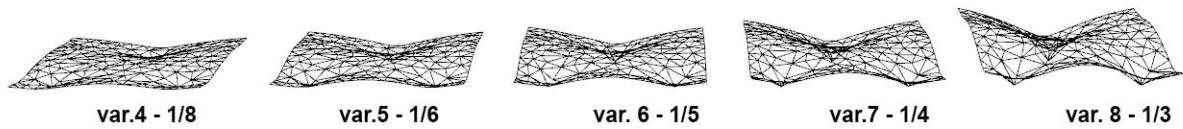


Fig. 4. Analyzed variants of space grid structures

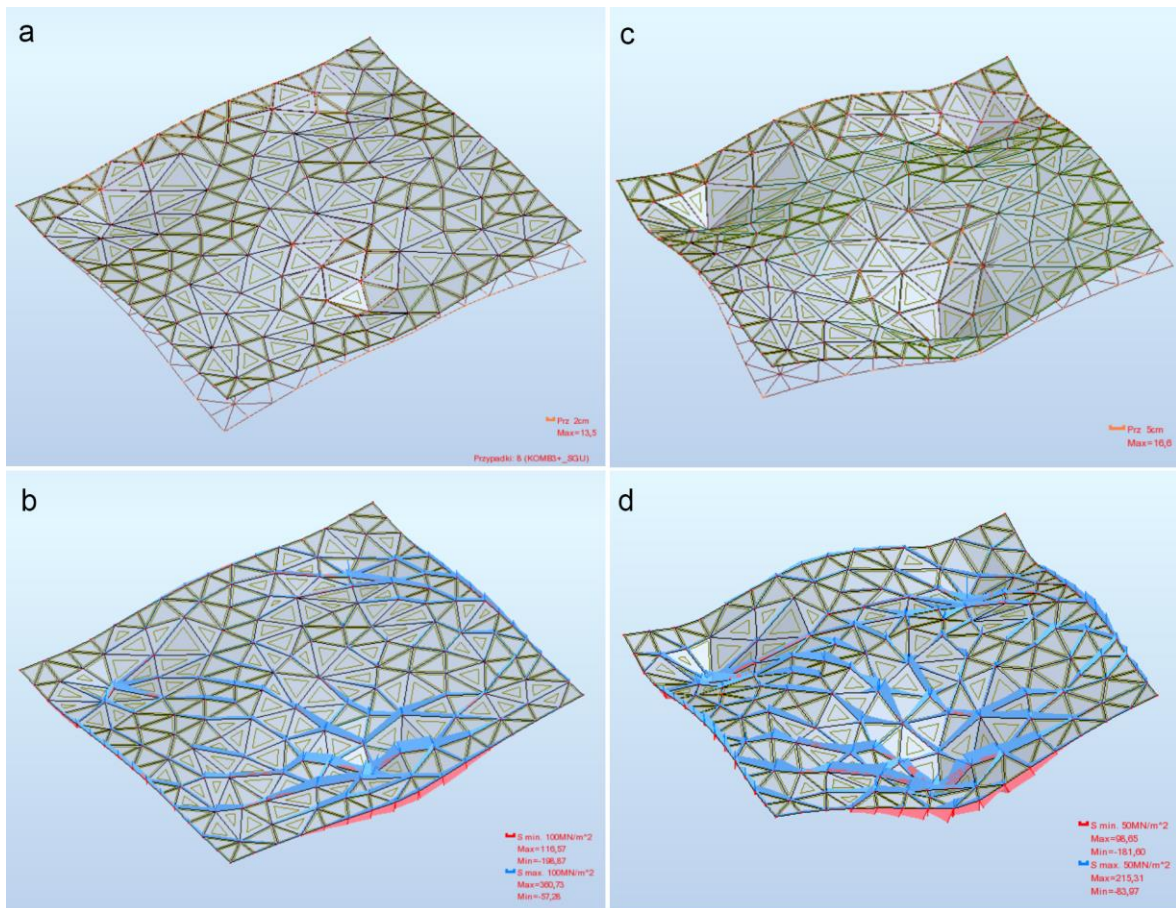


Fig. 5. Analysis for two extreme variants; **Variant 4** (1/8): **a** - deformation chart; **b** - stress diagram; **Variant 8** (1/3): **c** - deformation chart; **d** - stress diagram

Tab. 1. Calculation results for spatial structures when the projected bars do not exceed 4,5m in length.

Types of structure		Profile				
	length of single planarized bar [m]	height of centre point/supports span	Type/Dimensions[mm]	max deformation [cm]	total length of elements [mm]	total weight [kg]
var. 4	4,5	1/8	TRON 406x8	13,5	1160	91 185
var. 5	4,5	1/6	TRON 355x8	13,7	1166,98	80 030
var. 6	4,5	1/5	TRON 323x7,1	15,6	1173,14	65 075
var. 7	4,5	1/4	TRON 323x6,3	14,4	1184,44	58 446
var. 8	4,5	1/3	TRON 273x6,3	16,6	1207,01	50 014

According to the assumptions adopted in the study, the maximum permissible deformation of 17.9 cm was met for all the variants. In each of the systems, the largest bending occurred on the cantilevers. However, differences in the work of individual bars can be seen in the stress diagram - curvature of the surface caused an increase in the number of compressed rods in the support zones forming structures similar to arboreal supports [2]. At the same time, it is worth mentioning that the curvilinear deformations of the roof also cause changes in the reactions at the fixed supports.

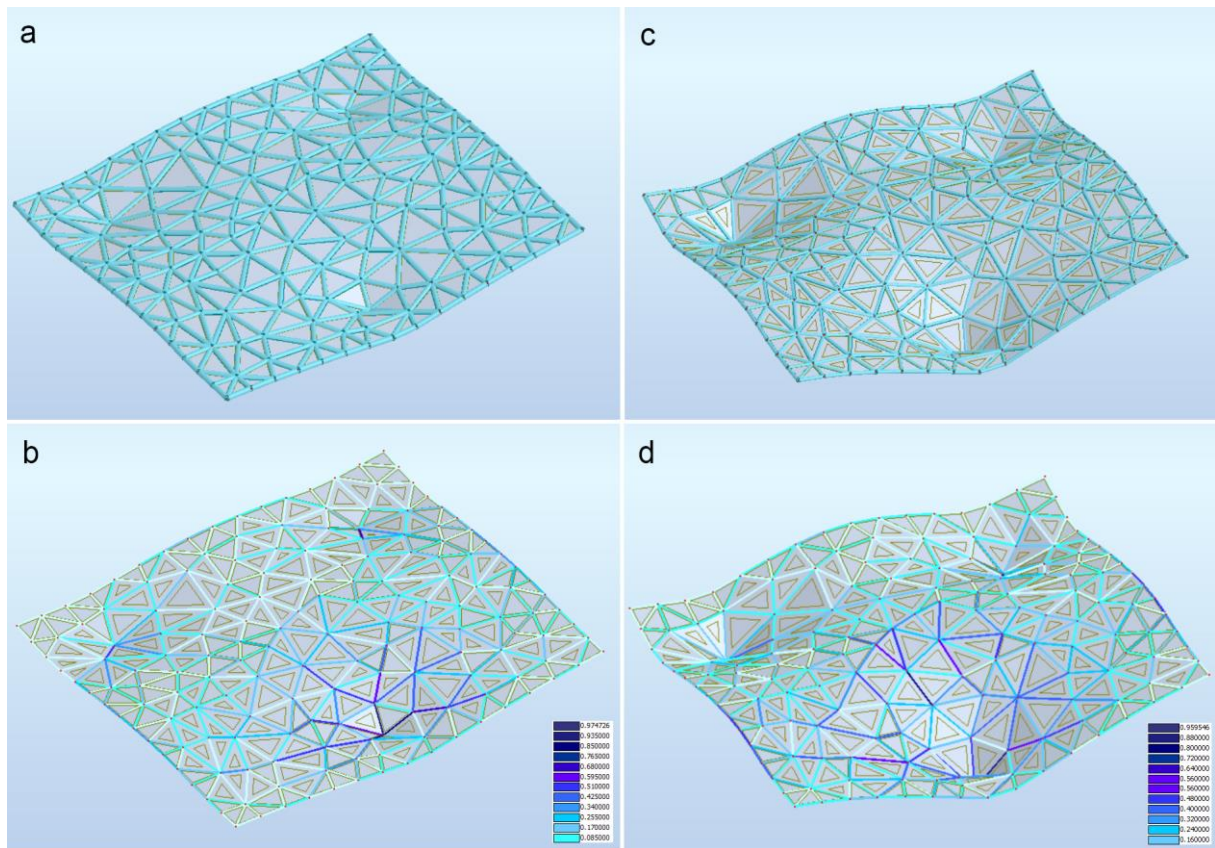


Fig. 6. Analysis of two extreme variants; **Variant 4** (1/8): **a** - shape of the roof together with the sketch of the profile; **b** – graph of limit state ratio for individual bars; **Variant 8** (1/3): **c** - shape of the roof together with the sketch of the profile; **d** – graph of limit state ratio for individual bars

As a result of approximation with the third degree polynomial, the least squares method was used for the given five variants. The curve determined in this way convinces us that for the assumed boundary conditions the optimum was not found. However, continuing the optimization process by increasing the degree of surface curvature on the one hand will lead to a change in the structure of the system (which may be desirable), and on the other will affect the aesthetics and functionality of the architectural form. So in the pursuit of optimal grid structures in architecture, making rational decisions through compromise, also in the context of designing the structure and determining the direction of optimization, is an important element.

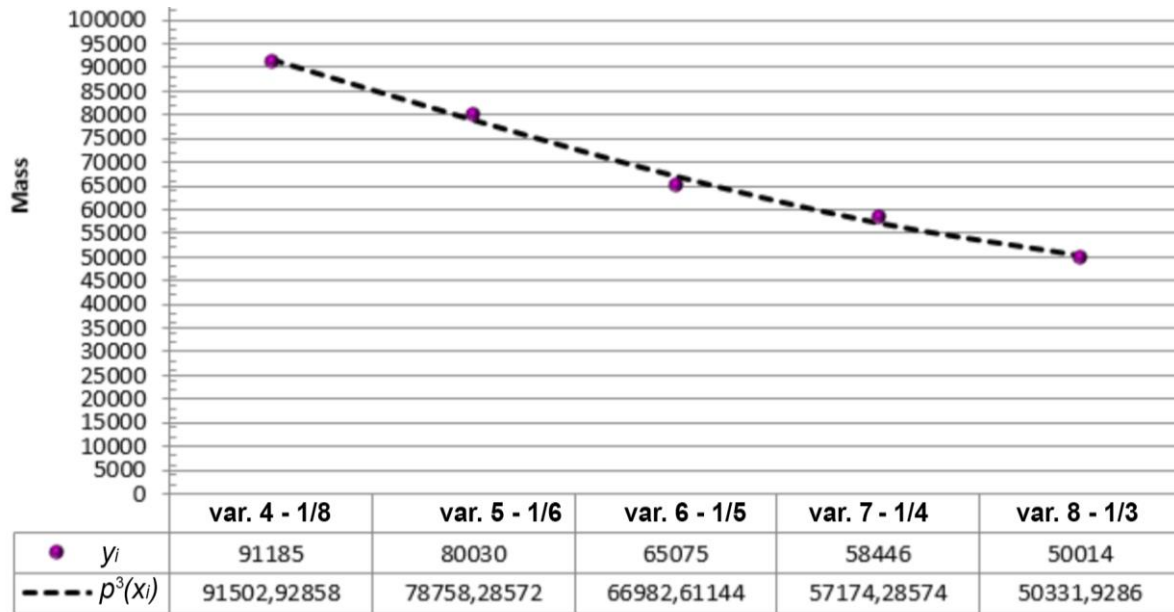


Fig. 7. Polynomial approximation diagram for $p^3(x_i)$ and y_i values (mass) for individual curvilinear variants

Summary

Nowadays we are observing a growing interest in bionic solutions. Digital tools supporting architectural design are more often equipped with methods of surface discretization using laws governing the processes occurring in nature. Biomimicry in architecture fosters the development of digital technologies, thanks to which it is possible to map chaotic systems, while the ways of reproducing the undeniable beauty of nature are gradually discovered and described by means of discrete mathematics.

Structural optimization is an integral part of shaping of the architectural vision, and the development of knowledge and the standardization of generative design methods have a significant impact on the way the architectural form is designed - its shape and the structure of the building's surface. Due to dynamically developing digital processes of design optimization the designers are not only aware of its logical use in bionic structures but also are able to use generative modeling methods to determine rational criteria for more detailed searches.

Delaunay triangulation is one of the bionic methods of surface division which finds application in architecture in the design of shell elements such as roofs, elevations, walls, etc. In architectural and construction optimization Delaunay triangulation can be used as an

alternative way of search for rational, and at the same time, tectonically innovative structural solutions. An important element of the optimization of the structural surfaces is the determination of the metric describing the density of the grid - so that the structure on one hand meets the visual effect intended by the architect, and on the other hand it remains a rational solution in the context of construction and building logic.

References

- [1] Gawell Ewelina, Nowak Anna, Rokicki Wiesław, „*Aperiodic tessellations in shaping the structural surfaces in the contemporary architecture*”, The Journal Biuletyn of Polish Society for Geometry and Engineering Graphics, Volume 26 (2014), Publishing House.: Polskie Towarzystwo Geometrii i Grafiki Inżynierskiej (Polish Society of Geometry and Engineering Graphics)
- [2] Md Rian Iasef, Sassone Mario, *Tree-inspired dendriforms and fractal-like branching structures in architecture: A brief historical overview*, “Science Direct”, < <https://doi.org/10.1016/j.foar.2014.03.006> >, [access: 28.06.2018]
- [3] Sieczkowski Jeremi Michał, „Podstawy komputerowego modelowania konstrukcji budowlanych”, Publishing House: Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2001, ISBN 83-7085-573-3
- [4] Tarczewski Romuald, *Topologia form strukturalnych; naturalne i tworzone przez człowieka prototypy form konstrukcyjnych w architekturze [Topology of structural forms; natural and manmade prototypes of structural forms in architecture]*, Publishing House: Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2011, ISBN 978-83-7493-660-6

A numerical study of thermal impact of forest fires on buildings

†*Valeriy Perminov¹

¹Department of Control and Diagnostics, Tomsk Polytechnic University, Russia.

*Presenting author: perminov@tpu.ru

†Corresponding author: perminov@tpu.ru

Abstract

The mathematical modeling of forest fires actions on buildings and structures have been carried out to study the effects of fire intensity and wind speed on possibility of ignition of buildings. The crown forest fire is introduced as a heat and mass source defined by the empirical values of average crown fire temperature and vertical gas velocity at the top crown surface dependent on fire intensity. The hydrodynamic and thermal interactions between plume, wind flow and building are analyzed. The approach to modeling is based on the use of standard non-stationary three-dimensional conservation equations for turbulent flow in a multiphase reacting medium that are solved numerically under the input conditions characteristic of a large forest fires.

Keywords: Computation Control volume, Crown fire, Fire spread, Forest fire, Mathematical model, Ignition, Building

Introduction

The protection of buildings and structures in a community from destruction by fire is a very important concern. This paper addresses the development of a mathematical model for impact of wildfires with buildings. The forest fire is a very complicated phenomenon. At present, fire services can forecast the danger rating of, or the specific weather elements relating to, forest fire. There is need to understand and predict forest fire initiation, behavior and impact of fire on the buildings and constructions. This paper's purposes are the improvement of knowledge on the fundamental physical mechanisms that control forest fire behavior. A great deal of work has been done on the theoretical problem of forest fires. The first accepted method for prediction of crown fires was given by Rothelmal [1] and Van Wagner [2]. The semi-empirical models [1-2] allow to obtain a quite good data of the forest fire rate of spread as a function of fuel bulk and moisture, wind velocity and the terrain slope. But these models use data for particular cases and do not give results for general fire conditions. Also crown fires initiation and hazard have been studied and modeled in detail (eg: Alexander [3], Xanthopoulos, [4], Van Wagner, [5], Cruz [6], Albini [7], Scott, J. H. and Reinhardt, E. D. [8]. The discussion of the problem of modeling forest fires is provided by Grishin [9]. A mathematical model of forest fires was obtained by Grishin [9] based on an analysis of known and original experimental data and using concepts and methods from reactive media mechanics. The physical two-phase models used in [10] may be considered as a development and extension of the formulation proposed by Grishin [9]. However, the study of crown fires initiation and spread [9,10] has been limited mainly to cases studied of forest fires propagation without take into account the mutual interaction of crown forest fires with different obstacles (roads, glades and etc.), buildings and constructions. In this paper, the impacts of crown forest fires on buildings are studied. The dangerous distances between forest

and buildings are calculated in cases when the buildings will be ignited under the influence of forest fires.

1. Physical and mathematical model

It is assumed that the forest during a fire can be modeled as 1) a multi-phase, multistoried, spatially heterogeneous medium; 2) in the fire zone the forest is a porous-dispersed, two-temperature, single-velocity, reactive medium; 3) the forest canopy is supposed to be non-deformed medium (trunks, large branches, small twigs and needles), which affects only the magnitude of the force of resistance in the equation of conservation of momentum in the gas phase, i.e., the medium is assumed to be quasi-solid (almost non-deformable during wind gusts); 4) let there be a so-called “ventilated” forest massif, in which the volume of fractions of condensed forest fuel phases, consisting of dry organic matter, water in liquid state, solid pyrolysis products, and ash, can be neglected compared to the volume fraction of gas phase (components of air and gaseous pyrolysis products); 5) the flow has a developed turbulent nature and molecular transfer is neglected; 6) gaseous phase density doesn't depend on the pressure because of the low velocities of the flow in comparison with the velocity of the sound. Let the point $x_1, x_2, x_3 = 0$ is situated at the center of the surface forest fire source at the height of the roughness level, axis Ox_1 directed parallel to the Earth's surface to the right in the direction of the unperturbed wind speed, axis Ox_2 directed perpendicular to Ox_1 and axis Ox_3 directed upward (Fig. 1). The building is situated on the right part of the picture.

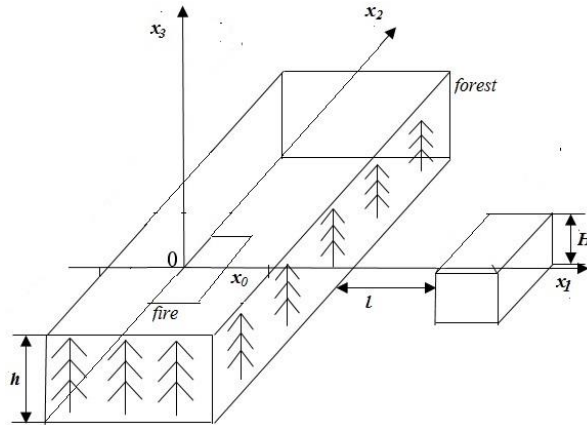


Figure 1. The scheme of calculation domain.

The problem formulated above reduces to the solution of the next system of equations:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}(\rho v_j) = Q, \quad j = \overline{1,3}, \quad i = \overline{1,3}; \quad (1)$$

$$\rho \frac{dv_i}{dt} = -\frac{\partial P}{\partial x_i} + \frac{\partial}{\partial x_j}(-\rho \overline{v'_i v'_j}) - \rho s c_d v_i |\vec{v}| - \rho g_i - Q v_i; \quad (2)$$

$$\rho c_p \frac{dT}{dt} = \frac{\partial}{\partial x_j}(-\rho c_p v'_j \overline{T'}) + q_5 R_5 - \alpha_v (T - T_s) + k_g (c U_R - 4\sigma T^4); \quad (3)$$

$$\rho \frac{dc_\alpha}{dt} = \frac{\partial}{\partial x_j}(-\rho v'_j \overline{c'_\alpha}) + R_{5\alpha} - Q c_\alpha, \quad \alpha = 1,2; \quad (4)$$

$$\frac{\partial}{\partial x_j} \left(\frac{c}{3k} \frac{\partial U_R}{\partial x_j} \right) - kcU_R + 4k_s \sigma T_s^4 + 4k_g \sigma T^4 = 0, \quad (5)$$

$$k = k_g + k_s;$$

$$\sum_{i=1}^4 \rho_i c_{pi} \varphi_i \frac{\partial T_s}{\partial t} = q_3 R_3 - q_2 R_2 - k_s (cU_R - 4\sigma T_s^4) + \alpha_v (T - T_s); \quad (6)$$

$$\rho_1 \frac{\partial \varphi_1}{\partial t} = -R_{1s}, \rho_2 \frac{\partial \varphi_2}{\partial t} = -R_{2s}, \rho_3 \frac{\partial \varphi_3}{\partial t} = \alpha_C R_{1s} - \frac{M_C}{M_1} R_{3w}, \rho_4 \frac{\partial \varphi_4}{\partial t} = 0; \quad (7)$$

$$\sum_{\alpha=1}^3 c_\alpha = 1, P_e = \rho RT \sum_{\alpha=1}^3 \frac{c_\alpha}{M_\alpha}, \vec{v} = (v_1, v_2, v_3), \vec{g} = (0, 0, g).$$

The system of equations (1)–(7) must be solved taking into account the initial and boundary conditions:

$$t = 0 : v_1 = 0, v_2 = 0, v_3 = 0, T = T_e, c_\alpha = c_{ae}, T_s = T_{se}, \varphi_i = \varphi_{ie}; \quad (8)$$

$$x_1 = 0 : v_1 = V, v_2 = 0, v_3 = 0, T = T_e, c_\alpha = c_{ae}, -\frac{c}{3k} \frac{\partial U_R}{\partial x_1} + \frac{c}{2} U_R = 0; \quad (9)$$

$$x_1 = x_{1e} : \frac{\partial v_1}{\partial x_1} = 0, \frac{\partial v_2}{\partial x_1} = 0, \frac{\partial v_3}{\partial x_1} = 0, \frac{\partial T}{\partial x_1} = 0, \frac{\partial c_\alpha}{\partial x_1} = 0, \frac{c}{3k} \frac{\partial U_R}{\partial x_1} + \frac{c}{2} U_R = 0; \quad (10)$$

$$x_2 = -x_{2e} : \frac{\partial v_1}{\partial x_2} = 0, \frac{\partial v_2}{\partial x_2} = 0, \frac{\partial v_3}{\partial x_2} = 0, \frac{\partial T}{\partial x_2} = 0, \frac{\partial c_\alpha}{\partial x_2} = 0, -\frac{c}{3k} \frac{\partial U_R}{\partial x_2} + \frac{c}{2} U_R = 0; \quad (11)$$

$$x_2 = x_{2e} : \frac{\partial v_1}{\partial x_2} = 0, \frac{\partial v_2}{\partial x_2} = 0, \frac{\partial v_3}{\partial x_2} = 0, \frac{\partial T}{\partial x_2} = 0, \frac{\partial c_\alpha}{\partial x_2} = 0, \frac{c}{3k} \frac{\partial U_R}{\partial x_2} + \frac{c}{2} U_R = 0; \quad (12)$$

$$x_3 = 0 : v_1 = 0, v_2 = 0, \frac{\partial c_\alpha}{\partial x_3} = 0, -\frac{c}{3k} \frac{\partial U_R}{\partial x_3} + \frac{c}{2} U_R = 0, \quad (13)$$

$$\rho v_3 = \rho_0 \omega_0, T = T_0, |x_1| \leq x_0, |x_2| \leq x_0,$$

$$\rho v_3 = 0, T = T_e, |x_1| > x_0, |x_2| > x_0;$$

$$x_3 = x_{3e} : \frac{\partial v_1}{\partial x_3} = 0, \frac{\partial v_2}{\partial x_3} = 0, \frac{\partial v_3}{\partial x_3} = 0, \frac{\partial T}{\partial x_3} = 0, \frac{\partial c_\alpha}{\partial x_3} = 0, \frac{c}{3k} \frac{\partial U_R}{\partial x_3} + \frac{c}{2} U_R = 0. \quad (14)$$

Here and above $\frac{d}{dt}$ is the symbol of the total (substantial) derivative; α_v is the coefficient of

phase exchange; ρ - density of gas – dispersed phase, t is time; v_i - the velocity components; T, T_s - temperatures of gas and solid phases, U_R - density of radiation energy, k - coefficient of radiation attenuation, P - pressure; c_p - constant pressure specific heat of the gas phase, c_{pi} , ρ_i , φ_i - specific heat, density and volume of fraction of condensed phase (1 – dry organic substance, 2 – moisture, 3 – condensed pyrolysis products, 4 – mineral part of forest fuel), R_i – the mass rates of chemical reactions, q_i – thermal effects of chemical reactions; k_g, k_s - radiation absorption coefficients for gas and condensed phases; T_e - the ambient temperature; c_α - mass concentrations of α - component of gas - dispersed medium, index $\alpha=1,2,3$ where 1 corresponds to the density of oxygen, 2 - to carbon monoxide CO , 3 - to carbon dioxide and inert components of air; R – universal gas constant; M_α, M_C , and M molecular mass of α - components of the gas phase, carbon and air mixture; g is the gravity acceleration; c_d is an empirical coefficient of the resistance of the vegetation, s is the specific surface of the forest fuel in the given forest stratum. To define source terms which characterize inflow (outflow of mass) in a volume unit of the gas-dispersed phase, the following formulae were used for the

rate of formulation of the gas-dispersed mixture Q , outflow of oxygen R_{51} , changing carbon monoxide R_{52} .

$$Q = (1 - \alpha_c)R_1 + R_2 + \frac{M_c}{M_1}R_3, R_{51} = -R_3 - \frac{M_1}{2M_2}R_5,$$

$$R_{52} = \nu_g(1 - \alpha_c)R_1 - R_5, R_{53} = 0.$$

$$R_1 = k_1 \rho_1 \varphi_1 \exp\left(-\frac{E_1}{RT_s}\right), R_2 = k_2 \rho_2 \varphi_2 T_s^{-0.5} \exp\left(-\frac{E_2}{RT_s}\right),$$

$$R_3 = k_3 \rho \varphi_3 s_\sigma c_1 \exp\left(-\frac{E_3}{RT_s}\right), R_5 = k_5 M_2 \left(\frac{c_1 M}{M_1}\right)^{0.25} \frac{c_2 M}{M_2} T_s^{-2.25} \exp\left(-\frac{E_5}{RT_s}\right).$$

The initial values for volume of fractions of condensed phases are determined using the expressions:

$$\varphi_{1e} = \frac{d(1 - \nu_z)}{\rho_1}, \varphi_{2e} = \frac{Wd}{\rho_2}, \varphi_{3e} = \frac{\alpha_c \varphi_{1e} \rho_1}{\rho_3},$$

where d - bulk density for surface layer, ν_z - coefficient of ashes of forest fuel, W - forest fuel moisture content. It is supposed that the optical properties of a medium are independent of radiation wavelength (the assumption that the medium is “grey”), and the so-called diffusion approximation for radiation flux density were used for a mathematical description of radiation transport during forest fires. To close the system (1)–(7), the components of the tensor of turbulent stresses, and the turbulent heat and mass fluxes are determined using the local-equilibrium model of turbulence (Grishin, [9]). The system of equations (1)–(7) contains terms associated with turbulent diffusion, thermal conduction, and convection, and needs to be closed. The components of the tensor of turbulent stresses $\overline{\rho v'_i v'_j}$, as well as the turbulent fluxes of heat and mass $\overline{\rho v'_j c_p T'}$, $\overline{\rho v'_j c'_a}$ are written in terms of the gradients of the average flow properties using the formulas

$$-\overline{\rho v'_i v'_j} = \mu_t \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) - \frac{2}{3} K \delta_{ij},$$

$$-\overline{\rho v'_j c_p T'} = \lambda_t \frac{\partial T}{\partial x_j}, -\overline{\rho v'_j c'_a} = \rho D_t \frac{\partial c_a}{\partial x_j},$$

$$\lambda_t = \mu_t c_p / Pr_t, \rho D_t = \mu_t / Sc_t, \mu_t = c_\mu \rho K^2 / \varepsilon,$$

where μ_t , λ_t , D_t are the coefficients of turbulent viscosity, thermal conductivity, and diffusion, respectively; Pr_t , Sc_t are the turbulent Prandtl and Schmidt numbers, which were assumed to be equal to 1. In dimensional form, the coefficient of dynamic turbulent viscosity is determined using local equilibrium model of turbulence [9]. The length of the mixing path is determined using the formula $l = x_3 k_t / (1 + 2.5 x_3 \sqrt{c_d s / h})$ taking into account the fact that the coefficient of resistance c_d in the space between the ground cover and the forest canopy base is equal to zero, while the constants $k_t = 0.4$ and $h = h_2 - h_1$ (h_2 , h_1 – height of the tree crowns and the height of the crown base). It should be noted that this system of equations describes processes of transfer within the entire region of the forest massif, which includes the space between the underlying surface and the base of the forest canopy, the forest canopy and the

space above it, while the appropriate components of the data base are used to calculate the specific properties of the various forest strata and the near-ground layer of atmosphere. This approach substantially simplifies the technology of solving problems of predicting the state of the medium in the fire zone numerically. The thermodynamic, thermophysical and structural characteristics correspond to the forest fuels in the canopy of a different (for example pine [9]) type of forest.

2. Numerical Solution and Results

The boundary-value problem (1)–(14) is solved numerically. A discrete analog was obtained by means of the control volume method using the SIMPLE like algorithm (Patankar [11]). Difference equations that arise in the course of sampling were resolved by the method of SIP [11]. In order to efficiently solve this problem in a reactive flow the method of splitting according to physical processes was used. The basic idea of this method is based on the information that the physical timescale of the processes is great than chemical. In the first stage, the hydrodynamic pattern of flow and distribution of scalar functions was calculated. Then the system of ordinary differential equations of chemical kinetics obtained as a result of splitting was then integrated. The time step for integrating each function has to be smaller than the characteristic time of physical process to ensure the convergence of the numerical method. The time step was selected automatically. The accuracy of the program was checked by the method of inserted analytical solutions. Analytical expressions for the unknown functions were substituted in (1)–(14) and the closure of the equations were calculated. This was then treated as the source in each equation. Next, with the aid of the algorithm described above, the values of the functions used were inferred with an accuracy of not less than 1%. The effect of the dimensions of the control volumes on the solution was studied by diminishing them. Fields of temperature, velocity, component mass fractions, and volume fractions of phases were obtained numerically. The first stage is related to increasing maximum temperature in the place of ignition with the result that a crown fire source appears. At this process stage over the fire source a thermal wind is formed a zone of heated forest fire pyrolysis products which are mixed with air, float up and penetrate into the crowns of trees. As a result, forest fuels in the tree crowns are heated, moisture evaporates and gaseous and dispersed pyrolysis products are generated. Ignition of gaseous pyrolysis products of the crown occurs at the next stage, and that of gaseous pyrolysis products in the forest canopy occurs at the last stage. At the moment of ignition, the gas combustible products of pyrolysis burn away, and the concentration of oxygen is rapidly reduced. The isotherms of gas phase components moved in the forest canopy by the action of wind. It is concluded that the forest fire begins to spread. The results of the calculation give an opportunity to consider forest fire spread for different wind velocity, canopy bulk densities and moisture forest fuel. It is considered the effect of forest fire front on the building which is situated near from the forest. The influences of wind velocity and distance between forest and building on ignition of building are studied numerically. The results of calculations can be used to evaluate the thermal effects on the building, located near from the forest fires. The temperature fields of crown forest fire at definite moment will be interacted with the obstacle - building (Figure 2 a) and b)) and ignited it. Fig.2. shows temperature fields at the different instants moments of forest fire spread for a wind speed of 15 m/s. During this process, the surface of the wall of the building heats as a result of convection and radiation heat transfer. The wood building will be ignited at definite temperature. It depends on wind velocities, distances from the forest fire to building, the height of building and others parameters. The Figures 3-5 represent the predicted distributions of temperature on the surfaces of the wall of the building as a function of vertical coordinate for the three selected wind speed values and different distances between forest and building. In paper [12] it is showed that the wood will be ignited when its

temperature exceeds 300C. The results of calculations presented on Figures 3-5 show that the surface temperature reach these values at wind velocities more than 6 m/s. The height of building in these calculations was $H=3$ m.

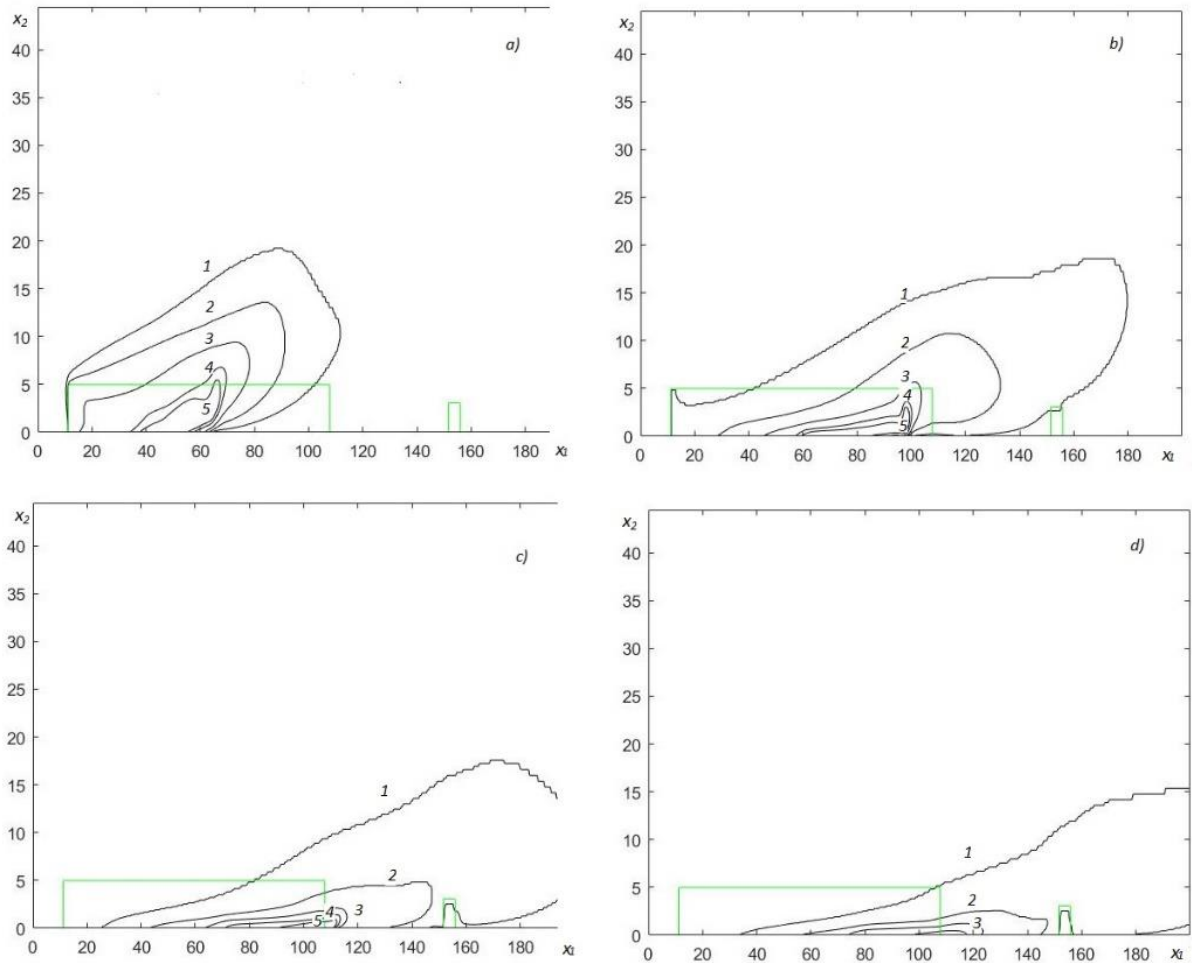


Figure 2. Gas temperature field at a) $t=15$ s, b) $t=20$ s, c) $t=23$ s, d) $t=25$ s for a wind speed of 15 m/s; 1 – 1.2, 2 – 1.5, 3 – 2., 4 – 3., 5 – 4; $\bar{T} = T/T_e$ $T_e = 300$ K.

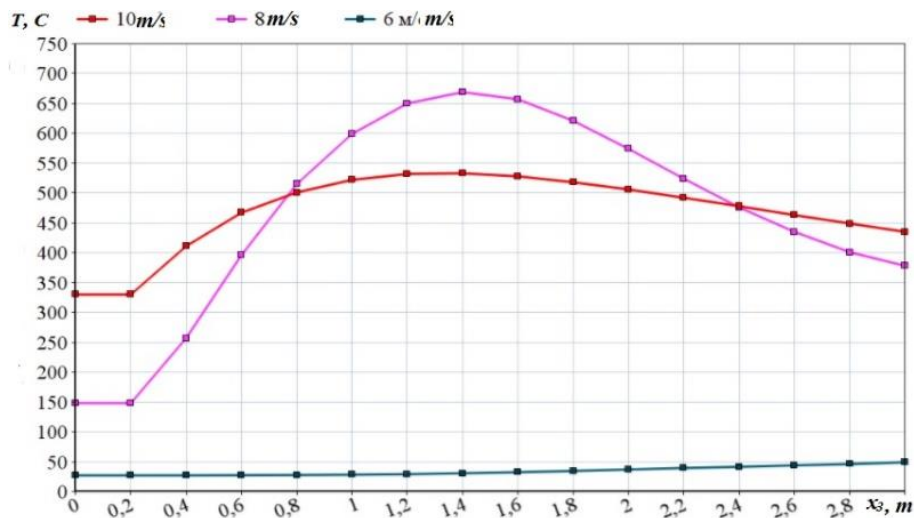


Figure 3. The distribution of temperature on the wall of the building for three wind speed values. The distance between forest and building is 21 m.

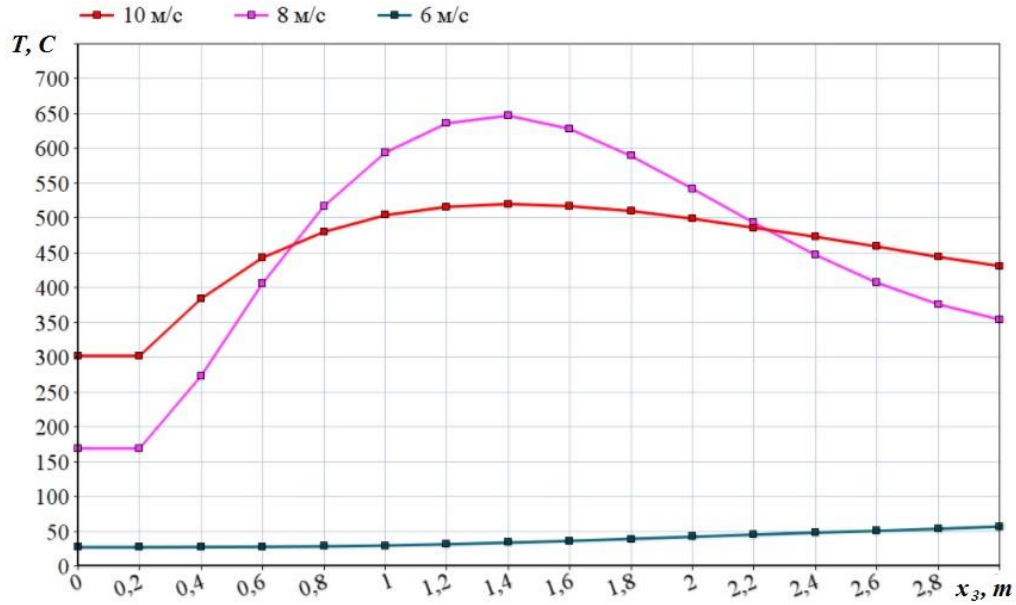


Figure 4. The distribution of temperature on the wall of the building for three wind speed values. The distance between forest and building is 26 m.

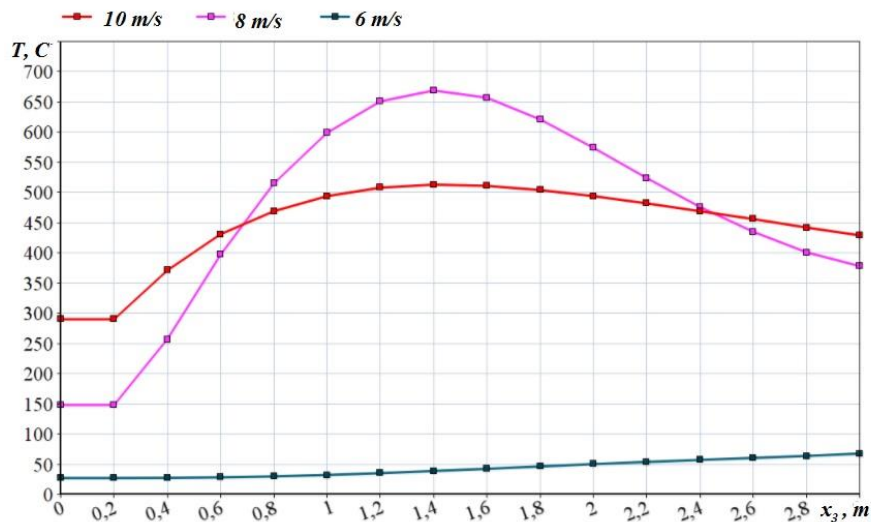


Figure 5. The distribution of temperature on the wall of the building for three wind speed values. The distance between forest and building is 31 m.

As a result of these calculations it was defined maximum safety distances between forest and building when the building would not have been ignited by forest fire (Fig. 6). The wind speed values increase from 6 to 14 m/s. Also, it was studied the influence of the height of building on the value of safety distances. When the height of building changes from 3 to 6 m, the safety distances l also increases (Figure 7) for different values of wind speed.

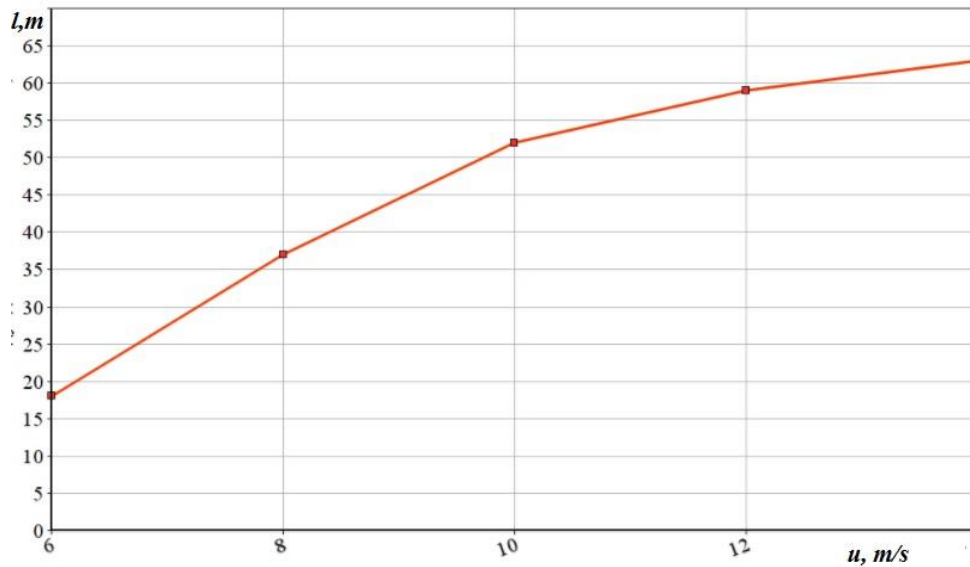


Figure 6. The dependence of safety distances between forest and building as a function of wind speed values; $H=3$ m.

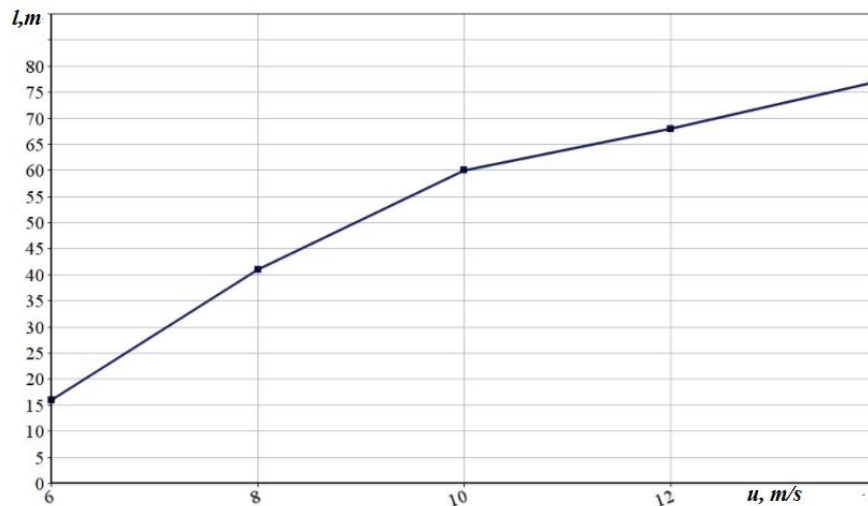


Figure 7. The dependence of safety distances between forest and building as a function of wind speed values. $H=6$ m.

3 Conclusions

A multiphase mathematical model of wind-aided crown forest fires propagating through heterogeneous fuel beds has been performed. It takes into account the hydrodynamic aspects of the flow and uses Arrhenius kinetics to describe the basic physics and chemical processes of thermal decomposition heating, drying, pyrolysis, and combustion. Turbulence and radiation are considered in order to improve the physical insight. It allows to investigate the dynamics of the impact of forest fires on buildings under the influence of various external conditions: a) meteorology conditions (air temperature, wind velocity etc.), b) type (various kinds of forest combustible materials) and their state (load, moisture etc.). The calculations let to get the maximum distance from the fire to the building in which the object possible ignition. It has been found that the effect of increasing the wind speed is to increase the safety distances between forest and building. The increasing of building height is observed also led

to increase the safety distances between forest and building. Specific experiments are also needed to obtain more reliable information on validation of further solution of this problem.

Acknowledgments

The paper was supported from Russian Foundation for Basic Research (project code: № 16-41-700022 p_a) and within the framework of Tomsk Polytechnic University Competitiveness Enhancement Program grant.

References

- [1] Rothermel, R.C. (1972) Predicting behavior and size of crown fires in the Northern Rocky Mountains, Res.Pap. INT-438. Ogden, UT: US Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station.
- [2] Van Wagner, C.E. (1977) Conditions for the Start and Spread of Crown Fire, *Canadian Journal of Forest Research* **7**, 23—34.
- [3] Alexander, M.E. Rothermel, R.C (1972). Predicting behavior and size of crown fires in the Northern Rocky Mountains, Res.Pap. INT-438. Ogden, UT: US Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station.
- [4] Xanthopoulos G. (1999) *Development of a Wildland Crown Fire Initiation Model*, PhD Thesis, University of Montana, Montana.
- [5] Van Wagner, C.E. (1999) Prediction of Crown Fire Behavior in Two Stands of Jack Pine, *Canadian Journal of Forest Research* **23**, 445—449.
- [6] Cruz, M.G., et al. (2002) Predicting Crown Fire Behavior to Support Forest Fire Management Decision-making, *Proceedings of IV International Conference on Forest Fire Research*, Ed. D. X. Viegas. **11** [CD-ROM]. (Millpress). Luso-Coimbra, (Portugal).
- [7] Albini, F.A. et al. (1995) Modeling Ignition and Burning Rate of Large Woody Natural Fuels, *International Journal of Wildland Fire* **5**, 81—91.
- [8] Scott, J.H., et al. (2001) Assessing Crown Fire Potential by Linking Models of Surface and Crown Fire Behavior. USDA Forest Service, Rocky Mountain Forest and Range Experiment Station. Fort Collins: RMRS-RP-29, (USA).
- [9] Grishin, A.M. (1997) *Mathematical Modeling Forest Fire and New Methods Fighting Them*. Publishing House of Tomsk University, Tomsk.
- [10] Morvan, D., Dupuy J.L. (2004): Modeling the Propagation of Wildfire Through a Mediterranean Shrub Using a Multiphase Formulation, *Combustion and Flame* **138**, 199-210.
- [11] Patankar, S.V. (1981) *Numerical Heat Transfer and Fluid Flow*, Hemisphere Publishing Corporation, New York.
- [12] Valendik, E.N., Mathveev, P.M., Safromov, M.A. (1979) Large Forest Fires and Fighting with Them. Science, Moscow (in Russian).

Drag reduction of KCS based on extended FFD method and EGO algorithm

Aiqin Miao^{*}, Decheng Wan[†]

School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

^{*}Presenting author: maq046@163.com

[†]Corresponding author: dcwan@sjtu.edu.cn
<http://dcwan.sjtu.edu.cn>

Abstract

The research presented in this paper is aimed at improve drag performance of a ship based on a flexible hull form modification method and an effective optimization algorithm. To obtain a series of practical new hull forms, a good ship modification method is needed first. Free-form deformation (FFD) method is a good deformation method to be widely applied in many shape design fields, such as aircrafts, Remote Operated Vehicles (ROV), cars and ships. Here, FFD method is extended by adding bending transformation in our in-house ship optimization solver, OPTShip-SJTU. In addition, a better optimization algorithm can greatly reduce computational cost and optimization responsive time. The efficient global optimization (EGO) algorithm has such good properties. It is a Kriging-based global optimization method, making the most of the knowledge of the error of Kriging model to search a cost landscape. In this paper, The KRISO Container Ship (KCS) is used as the initial ship, locally modified in the front half of the ship by FFD method mentioned above. The objective function (the wave-making resistance) is evaluated by the potential theory, Neumann-Michell method. Through the EGO algorithm, the drag of the initial ship is fast and efficiently optimized and the corresponding optimal ship is obtained. To verify the optimal result, the optimal ship is compared in detail with the initial one in the aspects of body lines, pressure distribution of ship surface, wave elevation, etc. It turns out the methods here are well applied to the ship optimization problem.

Keywords: drag performance; ship optimization design; extended FFD; EGO

Introduction

CFD is currently playing an increasingly important role in numerical prediction of ship hydrodynamic performance. The problem of long model test period and high cost is solved to a large extent. The ship hull optimization also began to leave the traditional design mode. A simulation-based design (SBD) mode emerged, that is, Numerical prediction based on CFD was used to evaluate the hydrodynamic performance, and the optimization algorithm was used to minimize the ship's hydrodynamic performance and improved ship hull lines. At present, a large number of scholars are studying such problems and have partly and successfully applied them to engineering practice.

However, ship optimization design requires a large amount of numerical calculation, which greatly increases the time and cost of CFD calculation. It needs to be achieved by means of approximation and parallel techniques. There is no clear direct expression between the hydrodynamic performance of a ship and the deformation parameters of a ship. The approximation technique implicitly expresses the relationship between the design variables and the objective function by mathematical means to construct an approximate model (Wu J. W., 2017; Liu X. Y. 2017), so that there is no need to call the CFD prediction of the hydrodynamic performance in the optimization process. The calculation is directly invoked by the approximation model, which greatly reduces the optimization design time and calculation

cost. Then for only optimization according to the approximate model, there are the following problems: Firstly, the accuracy of the approximate model needs to be accurate enough before optimization, and enough sample points are needed, otherwise the result is unreliable and the optimization fails; at the same time, the obtained optimized solution needs to be calculated again by high-fidelity CFD method or model test to confirm.

This paper adopts an efficient global optimization algorithm, also called the sequential global optimization algorithm based on approximation model. This method makes full use of the approximate model, and combines the estimation results of the existing approximate model with the uncertainty of the approximate model and the optimization algorithm. The high-precision calculation of additional sample points is needed by the above method. It continuously improves the accuracy of the approximate model and continuously explores the optimal value. An EI criterion (Expected Improvement) was proposed by Donald R. Jones in 1998. The sample points are needed by optimizing this criterion, which was successfully applied to the optimization of mathematical functions, especially high-dimensional functions. The convergence speed is much larger than many algorithms such as genetic algorithms. Later, more developments were made (Weihs. C, 2016) from the original single-objective optimization to multi-objective optimization (Seulgi, Y. I. 2014), and It began to be applied to engineering practice, including the design of airfoils, the volume of the oil pump and other optimization design problems (Jeong S, et al, 2015; Yi S, 2015)

In addition to the need for better optimization algorithms, the transformation method for ship lines must be further studied. Free-form deformation is a good and mature method to modify ship lines and it is extended based on the FFD transformation method that has been developed. The object can be achieved bending deformation in the presence of continuity. The method is applied to hull form transformations effectively and efficiently especially for local shapes deformation such as bulbous bow and two-skeg stern.

In the first half of this paper, the FFD method and EGO method will be briefly introduced. In the second of this paper, the methods will be applied to the optimization problems of ship design of KCS.

FFD method

Free-form deformation (FFD) method is a good choice to modify hull form locally. FFD method was first described by Thomas W. Sederberg and Scott R. Parry (1986) and was based on an earlier technique by Alan Barr (1984). Its basic idea of this method is embedding a ship or the region of the ship to be deformed within a parallelepipedical 3D lattice regularly subdivided. Then it can modify the surface shape of a ship by the following relationship.

$$X_{ffd} = \sum_{i=0}^l \sum_{j=0}^m \sum_{k=0}^n B_{i,j}(s) B_{j,k}(t) B_{i,k}(u) Q'_{i,j,k} \quad (1)$$

Wherein $Q'_{i,j,k}$ is the coordinates of the control points on the lattice, while X_{ffd} is the coordinates of the points of the ship surface. B is Bernstein polynomial, l, m, n are the numbers of the control points along the x-axis, y-axis, z-axis direction, respectively. Through changing the movable number, direction and displacement of the control points, the different ship surfaces can be easily obtained. Figure 1 is a sketch of a ship's bulbous bow deformation through the FFD method. Given the control points' rotating angle, the entire lattice is deformed, so that the bulbous bow is bent.

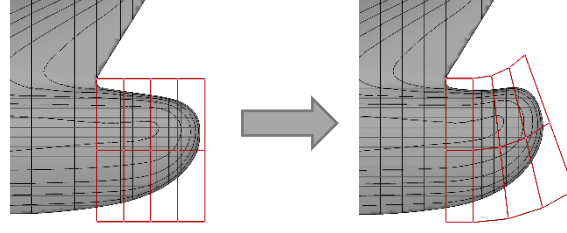


Figure 1 The application of the FFD method

In the paper, two lattices (Fig. 2) are used to modify the shape of the ship's front half part, the small lattice to modify the shape of the bulbous bow including the length, width, and degree of curvature, and the large one to modify the fatness of the ship's front half. A total of 7 design variables related to the shape deformation are involved in the optimization problem. Then, 35 sample hulls are generated through design of experiments.

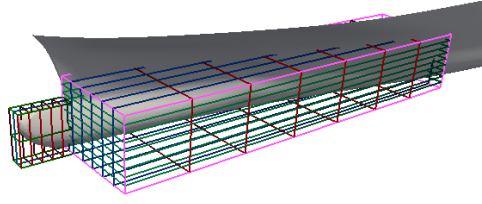


Figure 2 Two lattices used by the FFD method

Kriging Model

Kriging model (Simpson et al., 1994, 2004) is developed from mining and geostatistical applications involving spatially and temporally correlated data. This model combines a global model and a local component:

$$y(x) = f(x) + z(x) \quad (2)$$

where $y(x)$ is the unknown function of interest, $f(x)$ is a known approximation function of x , and $z(x)$ is the realization of a stochastic process with mean zero, variance σ^2 , and non-zero covariance. With $f(x)$ and $z(x)$, the kriging model can build the surrogate model between the input variables and output variables.

The kriging predictor is given by:

$$\hat{y} = \hat{\beta} + \mathbf{r}^T(x) \mathbf{R}^{-1}(\mathbf{y} - \mathbf{f}\hat{\beta}) \quad (3)$$

where y is an n_s -dimensional vector that contains the sample values of the response; \mathbf{f} is a column vector of length n_s that is filled with ones when f is taken as a constant; $\mathbf{r}^T(x)$ is the correlation vector of length n_s between an untried x and the sampled data points $\{x^{(1)}, x^{(2)}, \dots, x^{(n_s)}\}$ and is expressed as:

$$\mathbf{r}^T(x) = [R(x, x^{(1)}), R(x, x^{(2)}), \dots, R(x, x^{(n_s)})]^T \quad (4)$$

Additionally, the Gaussian correlation function is employed in this work:

$$R(x^i, x^j) = \exp \left[- \sum_{k=1}^{n_{dv}} \theta_k |x_k^i - x_k^j|^2 \right] \quad (5)$$

In equation (3), $\hat{\beta}$ is estimated using equation (5):

$$\hat{\beta} = (\mathbf{f}^T \mathbf{R}^{-1} \mathbf{f})^{-1} \mathbf{f}^T \mathbf{R}^{-1} \mathbf{y} \quad (6)$$

The estimate of the variance $\hat{\sigma}^2$, between the underlying global model $\hat{\beta}$ and y is estimated using equation (7):

$$\hat{\sigma}^2 = \left[(\mathbf{y} - \mathbf{f}\hat{\beta})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{f}\hat{\beta}) \right] / n_s \quad (7)$$

where $f(x)$ is assumed to be the constant $\hat{\beta}$. The maximum likelihood estimates for the θ_k in equation (5) used to fit a kriging model are obtained by solving equation (8):

$$\max_{\theta_k > 0} \Phi(\theta_k) = -\left[n_s \ln(\hat{\sigma}^2) + \ln|R| \right] / 2 \quad (8)$$

where both $\hat{\sigma}^2$ and $|R|$ are functions of θ_k . While any value for the θ_k create an interpolative kriging model, the “best” kriging model is found by solving the k-dimensional unconstrained, nonlinear, optimization problem given by equation (8).

the accuracy of the prediction value largely depends on the distance from sample points. Intuitively speaking, the closer point x to the sample point, the more accurate is the prediction \hat{y} . This intuition is expressed as

$$s^2(x) = \hat{\sigma}^2 \left[1 - r^T R^{-1} r + \frac{(1 - R^{-1} r)^2}{1^T R^{-1} 1} \right] \quad (9)$$

where $s^2(x)$ is the mean squared error of the predictor and it indicates the uncertainty at the estimation point. The root mean squared error (RSME) is expressed as $s = \sqrt{s^2(x)}$.

A kriging-based global efficient optimization algorithm

Traditionally, once the surrogate model is constructed, the optimum point can be explored using an arbitrary optimizer on the model. However, it is possible to miss the global optimum because the approximation model includes uncertainty at the predicted point.

In Fig. 3, the solid line is the real shape of objective function. Eight points are selected to construct the kriging model, which is shown as red points. The minimum point on the kriging model is located near $x=16$, whereas, the real global minimum of the objective function is situated near $x=17$. Searching for the global minimum using the present kriging model will not result in the real global minimum near $x=17$. For a robust search of the global optimum, both the predicted value by the kriging model and its uncertainty should be considered at the same time.

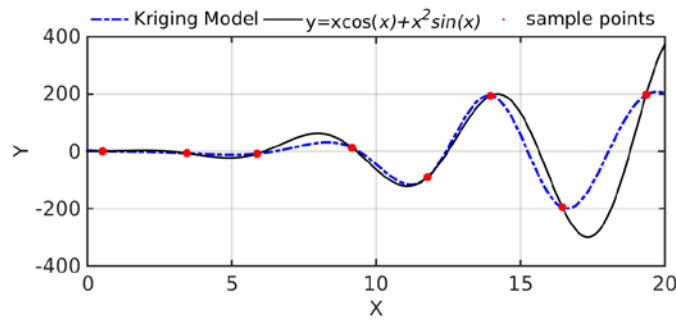


Figure 3. The Kriging model and the real function curve

When Kriging model is built, the mean predicted value and the standard error of the kriging model at any point can be evaluated. Considering the uncertainty of the model, this concept is expressed in the criterion of EI. The EI of minimization problem can be calculated as

$$E[I(x)] = (f_{\min} - \hat{y}) \Phi[(f_{\min} - \hat{y}) / s] + s \phi[(f_{\min} - \hat{y}) / s] \quad (10)$$

Where f_{\min} is the minimum value among n sampled values. Φ and ϕ are the standard distribution and normal density, respectively.

In fact, if we compute the derivative of EI as given in equation (9) with respect to \hat{y} and s , we get several terms that cancel, resulting in the simple following expressions:

$$\partial EI / \partial \hat{y} < 0, \partial EI / \partial s < 0 \quad (11)$$

It turns out that EI is monotonic in \hat{y} and in s . Thus, we see that the EI is larger the lower is \hat{y} and the higher is s . By selecting the maximum EI point as additional sample point through DE algorithm mentioned above, robust exploration of the global optimum and improvement of the model can be achieved simultaneously.

The overview over the whole efficient global optimization (EGO) procedure mentioned before is shown in Figure 4. First, the initial sample points should be chosen by experiment of design uniformly covering the whole design space. Secondly, an ordinary Kriging model is built and used to predict the objective for each design variable. Thirdly, the expected improvement (EI) balancing between regions of the low mean prediction and of high standard error is constructed to select the next point. The choice of the next sample point is the maximization of the EI value. Next, the objective function in the new point is calculated accurately and used to build a new surrogate model with the initial sample points, thus the next iteration is initiated. Finally, when the EI value is very small after n iterations, i.e., $\max(EI) < \Delta s \cdot (\max(y) - \min(y))$, where Δs is the relative stopping tolerance, or reach the maximization of iteration steps, the loop should be stopped.

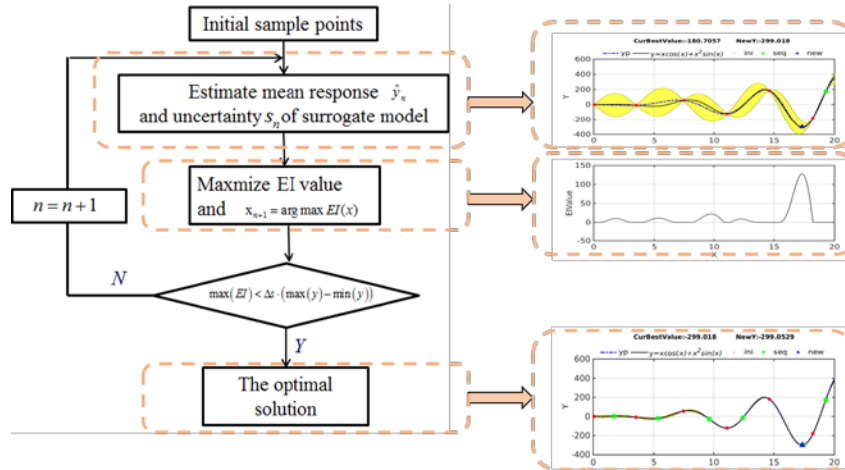


Figure 4. The flow chart of efficient global optimization: on the left, the steps are briefly described; on the right, an example is given (predetermined design points as red dots, the added new points as green squares and the next new point as a blue triangle).

Application of ship optimization design of KCS

In this study, the EGO method was applied to ship hull form design and the optimization of the resistance performance in calm water.

$$f_{obj} = R_w, Fr = 0.26 \quad (12)$$

The design problem is to minimize the wave-making resistance of KCS at the design speed. The extended FFD method mentioned above was applied to modify ship hull form (Fig.5). The design variables are parameters closely related to ship hull form modification. The total 7 design variables in Table 1 are used to define the geometry of ship hull form. The upper and lower bounds of each parameter are determined to avoid unrealistic ship hull geometry.

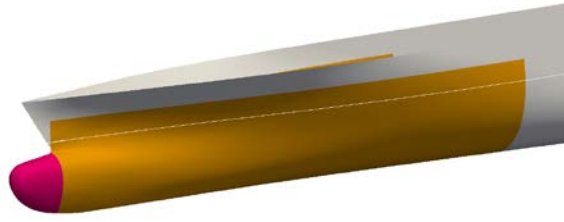


Figure 5. The variable region and the control points distributed on KCS hull by extended FFD method

Table 1. The design variables and their ranges

Design variables	Deformation direction	The bounds
1#	x	[-0.008, 0.01]
2#	y	[-0.01, 0.01]
3#	θ	[-40, 20]
4#	y	[-0.02, 0.015]
5#	y	[-0.015, 0.015]
6#	y	[-0.015, 0.015]
7#	y	[-0.01, 0.01]

At the early stage of optimization design, 49 sample points are spread over the design space and selected by Optimal Latin Hypercube Sampling (OLHS) to obtain a kriging model (Park J. S., 1994). The number of sample points is very important to keep the kriging model accurate in the traditional optimization process. However, in the present kriging model, additional sample points will be added later in the region where the accuracy is not good enough, or the objective function value is lower based on EI evaluation. The wave-making resistance of 49 sample hull form and additional sample points are all evaluated using a potential flow theory, Nuemann-Michell method (Noblesse F., 2013; Wu, J. W., 2016).

The summary of the parameters of EGO method is shown in Table 2. After efficient global optimization search, the total number of sample points reached 79, after adding 30 more sample points (in Fig.6). The objective function converges to the minimum value, 10.5629N, a larger reduction of about 25.13% than the initial value. We just use a PC to finish all the evaluations for about 1.89 hours.

In order to verify the accuracy and efficiency of EGO, we use the traditional optimization method, which firstly enough points are used to construct a surrogate model with sufficient precision (90 sample points are used here), and then the genetic algorithm (GA) is applied to search the optimal value of the surrogate model (in Tab. 3). The optimal ship is obtained through more than 200 iterations, and the wave resistance value reaches 11.1037 N, which is reduced by 21.30%. And the design variables of the optimal hull forms based on two methods are shown in Fig. 7.

The comparison of Table 2 and Table 3 and Figure 6 clearly show the efficiency of the EGO method. It greatly reduces the cost of high-precision numerical calculations. It can be seen from Fig. 7 that the optimized ships obtained by the two methods are different.

Table 2. The parameters of the EGO method

The initial number of sample points	49	The additional number of sample points	30
The number of iterations	30	Optimization time (h)	1.89

Table 3. The parameters of the GA method

The initial number of sample points	90	The number of populations	50
The number of iterations	400	Optimization time (h)	3.12

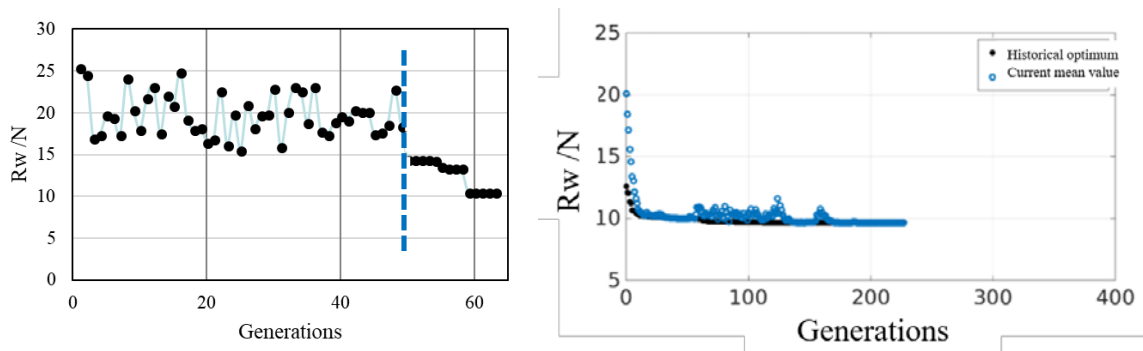


Figure 6. The total numbers of numerical evaluation using the EGO method (left) and the GA method (right) until convergence

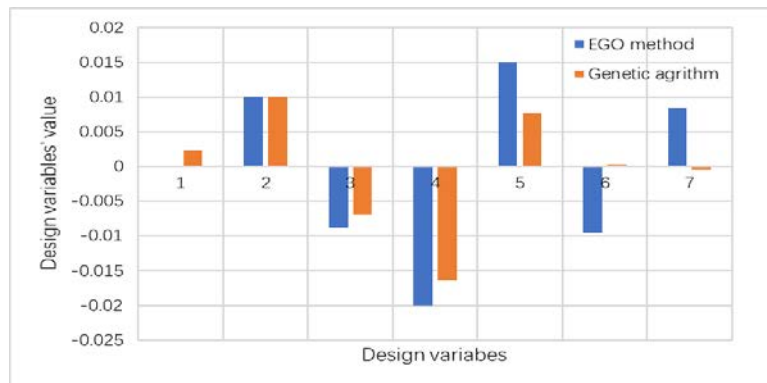


Figure 7. The values of design variables and the optimal solution of the objective function after optimization procedure

The following figures depict the comparisons of body lines and shapes of the initial KCS and the optimal ship based on the two methods. The same changes are as following: first changes

have taken place in the waterline of the optimized ships, the waterline near the inlet has become fatter. In addition, the bilge of the optimal ship has become thinner, the bulbous bow has also become thinner and longer, and has slightly been flattened compared to the original ship. The different changes of the optimized hulls based on the two methods happens at the front shoulder of ships. The optimized ship based on EGO method has a large difference from the initial one, while the optimized ship based on GA has basically not changed.

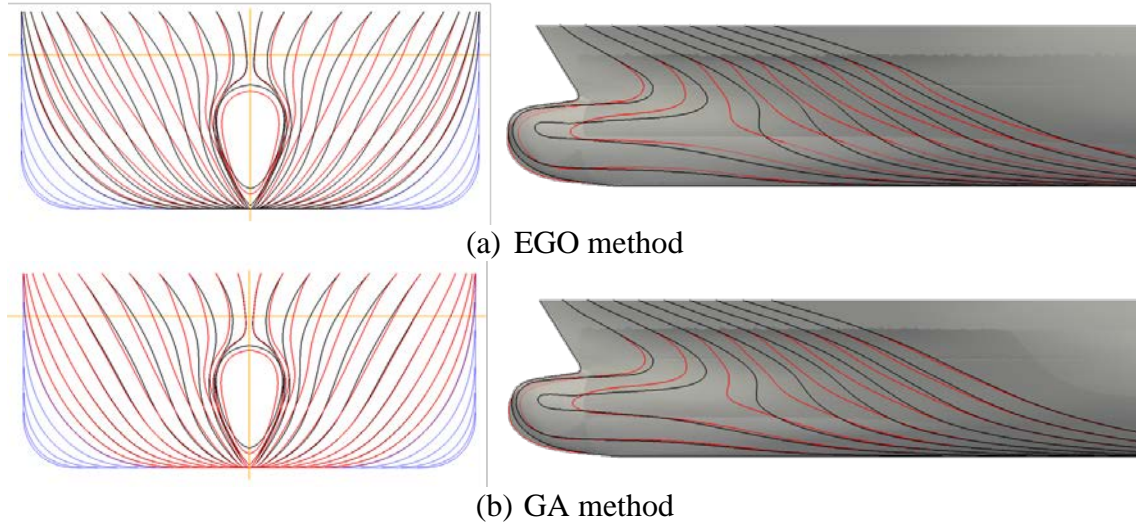


Figure 8. Comparisons of the body lines and shapes between the initial hull and the optimal one

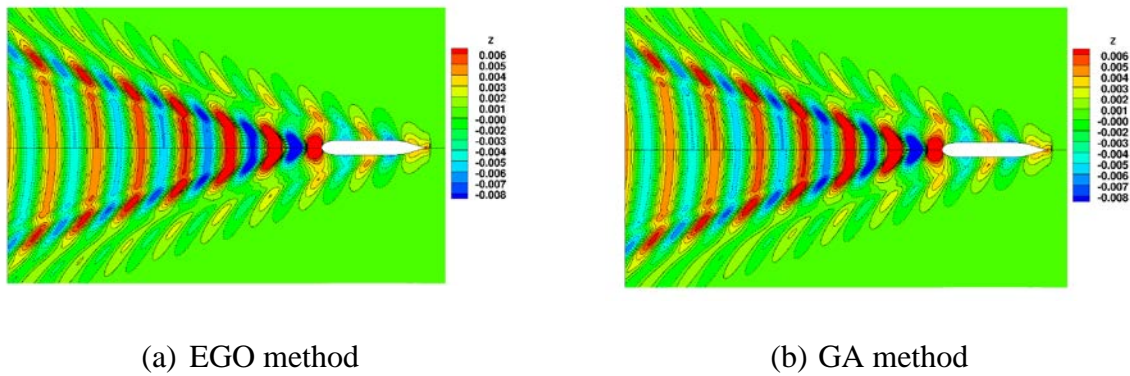


Figure 9. Comparison of free surface elevation between the initial hull and the optimal one

Figure 9 is the wave profile of free surface of the two optimized ships. In contrast, both optimized hulls obtained by two methods generate lower wave than the initial one obviously, thus leads to the reduction of the wave-making resistance of the optimal hull. Also comparing the pressure distribution of the two ships in Fig. 10, the fore parts of the optimal hulls is with lower pressure than the initial one. Here we can conclude that the small variation of the front shoulder of the optimized ships has less effect on the improvement for the wave resistance of this ship.

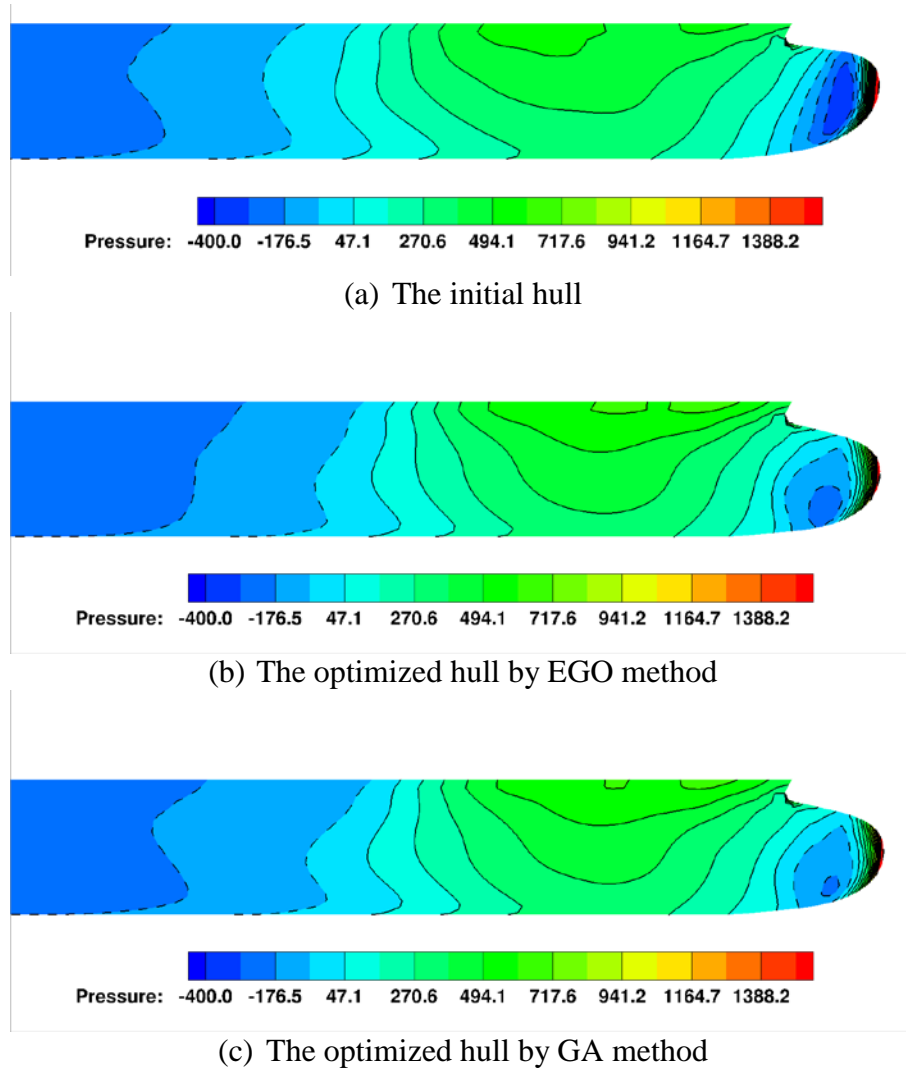


Figure 10. Comparison of pressure distribution between the initial hull and the optimal one

Conclusion and future work

This paper presents a Kriging-based global optimization method, efficient global optimization (EGO), different from the ordinary optimization method. It combines the surrogate modeling with the optimization algorithm. By this method, not only the accuracy of surrogate model is continuously improved but also the solution of the optimization problem keeps searched in the iterative procedure. An FFD method for ship modification is also used to allow the geometry or a part of the geometry to be bent. The two methods are successfully applied to ship hull form optimization design. Based on the comparison of EGO method and the traditional optimization method, the results demonstrate the usability and efficiency of the EGO methods in ship hull form optimization design. In the future, it will be used to the more ship optimization problem, such as the ship hull form design to improve comprehensive hydrodynamic performance, based on entire CFD.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

- [1] Barr, A. H. (1984) Global and local deformations of solid primitives. *Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, ACM: 21-30.
- [2] Coquillart, S. (1990) Extended free-form deformation: a sculpturing tool for 3D geometric modeling." *SIGGRAPH Comput. Graph.* **24**(4): 187-196.
- [3] Jones, D. R., Schonlau, M., and Welch, W. J. (1998) Efficient global optimization of expensive black-box functions. *Journal of Global Optimization* **13**, 455-492.
- [4] Jeong, S., Murayama, M., and Yamamoto, K. (2015) Efficient optimization design method using kriging model. *Journal of Aircraft* **42**, 413-420.
- [5] Liu, X. Y., Wu, J. W., and Wan, D. C. (2016) Multi-objective optimization for a surface combatant using Neumann-Michell theory and approximation model. *Proceedings of 12th International Conference on Hydrodynamics (ICHHD)*, 18-23.
- [6] Noblesse, F., Huang, F. X. and Yang, C. (2013) The Neumann-Michell theory of ship waves. *Journal of Engineering Mathematics* **79**, 51-71.
- [7] Sederberg, T. W. and S. R. Parry (1986) Free-form deformation of solid geometric models. *SIGGRAPH Comput. Graph.* **20**(4): 151-160.
- [8] Seulgi, Y. I., Kwon, H. I., and Choi, S. (2014). Efficient global optimization using a multi-point and multi-objective infill sampling criteria. *Aerospace Sciences Meeting*.
- [9] Simpson, T. W. (1998) Comparison of response surface and kriging models in the multidisciplinary design of an aerospike nozzle. *Institute for Computer Applications in Science and Engineering (ICASE)*.
- [10] Simpson, T. W., Booker, A. J., Ghosh, D., Giunta, A. A., Koch, P. N. and Yang, R. J. (2004) Approximation methods in multidisciplinary analysis and optimization: a panel discussion. *Structural and multidisciplinary optimization* **27**, 302-313.
- [11] Weihs, C., Herbrandt, S., Bauer, N., Friedrichs, K. and Horn, D. (2016) Efficient global optimization: Motivation, variations and applications.
- [12] Wu J. W., Liu X. Y., and Wan D. C. (2016) Multi-objective hydrodynamic optimization of ship hull based on approximation model. *Proceedings of 26th International Offshore and Polar Engineering Conference (ISOPE)*, Rhodes, Greece, 814-820.
- [13] Wu, J. W., Liu, X. Y., Zhao, M., and Wan, D. C. (2017) Neumann-Michell theory-based multi-objective optimization of hull form for a naval surface combatant. *Applied Ocean Research* **63**, 129-141.
- [14] Yi, S., Lee, D. J., Choi, S., and Im, D. (2015). Efficient design optimization of flexible CROR blade for low-level noise. *AIAA Applied Aerodynamics Conference*.

Fracture of alpha titanium alloys at high strain rates and stress triaxiality

†Vladimir V. Skripnyak¹, Alexander A. Kozulyn¹, Evgeniya G. Skripnyak¹, and
Vladimir A. Skripnyak¹

¹Research Laboratory of Properties of Substances in Extreme States, National Research Tomsk State University, Russia

*Presenting author: skrp2012@yandex.ru

†Corresponding author: skrp2012@yandex.ru

Abstract

The aim of this work was the evaluation of combined effect of stress triaxiality and strain rate on the mechanical behavior of the alpha titanium alloys. Mechanical behaviour of titanium alloy Grade 6 (VT5-1 or Ti-5Al-22,5Sn) and Grade 2 (VT1-0) in a range of strain rates from 0.001 to 1000 1/s and stress triaxiality (0.3–0.6) at room temperature was studied using an Instron VHS 40 / 50-20 servo-hydraulic test machine. Analysis of stress state and strain distribution in smooth and notched samples under tension was carried out by means of computer simulation and analytical relations. The Gurson-Tvergaard-Needleman model, complemented with phenomenological laws for voids nucleation, growth and coalescence, was adopted for describing the fracture of the alpha titanium alloys. It was shown that stress triaxiality is important for prediction of damage evolution and fracture of alpha titanium alloys at high strain rates. The constitutive equation and fracture models have been validated by simulating the tension tests. It was obtained that strain localization phenomena play a major role in the fracture process at lower triaxiality. It is found that the strain to fracture of alpha titanium alloy is strongly depended on the stress triaxiality and strain rate above 100 s⁻¹. The strain to failure of alpha titanium alloys at room temperature decrease by 3.7 times with increasing stress triaxiality from 0.3 to 0.6 in wide range of strain rates.

Keywords: Computer simulation, mechanical behavior, ductility, titanium alloys, high strain rate, stress triaxiality

Introduction

Polycrystalline alpha-titanium alloys belong to the isomechanical group of metal materials with hexagonal close packed (HCP) crystalline lattice. Materials belonging to the same mechanical group possess similarity of mechanical behavior in wide ranges of strain rates and temperature owing to similarity of mechanisms of plastic deformation and fracture. [1].

Generalization of data on the laws of deformation and fracture of alpha titanium alloys in a wide range of strain rates will allow developing a method for prediction of titanium structural elements under dynamic impacts [2]-[5]. It was shown that the ductility and strength of titanium alloys in a wide range of strain rates depends on the grain size and a grain size distribution [5]-[9]. There is evidence that the ductility and fracture of the hexagonal close packed polycrystalline metals and alloys is strongly dependent on the accumulated plastic strain and stress triaxiality [10]-[13]. The dimensionless stress triaxiality η , defined as the ratio between the hydrostatic stress and equivalent stress [10][11][14].

A significant distinction has been noted between the regimes of high and low stress triaxiality. High values of triaxiality (i.e., $\eta > 1.5$) may be achieved in local areas, such as at the ends of the cracks or in the center of a necking or notched specimen under tension. Low stress triaxiality takes place at surfaces and protruding corners, where the equivalent shear stress is high relative to the hydrostatic pressure [15]-[19]. At low stress triaxiality the fracture initiation is strongly inhibited.

Several models were proposed for investigation of the effect of triaxiality on the fracture of polycrystalline metals and alloys [20]-[24]. Neilsen and Tvergaard [20][21] showed that ductile fracture can be described using the criterion, depending on the stress triaxiality and the Lode angle. Valoppi and others [11] used the phenomenological Johnson Cook hardening model and damage initiation criterion with an energy-based law describing the damage evolution in the finite element models for titanium alloy. It was shown, the Gurson–Tvergaard–Needleman plasticity model can be complemented with phenomenological laws for void nucleation, growth and coalescence [10][21]. In this research we study the influence of different values of stress triaxiality ($0.33 < \eta < 0.6$) on ductile fracture in a wide range of strain rates using experimental tests and numerical simulation. Four types of VT 5-1 (Ti-5Al-22,5Sn) sheet samples were used in experiments on static and dynamic tension.

1 Material and samples

Sheets of alpha titanium alloys VT 5-1 (Grade 6) and VT1-0 (Grade 2) were studied under tension in a range of strain rates from 0.01 to 1000 s^{-1} at room temperature. The thickness of the sheet samples VT5-1 and VT1-0 was $1.15 \pm 0.05 \text{ mm}$, and $1.1 \pm 0.05 \text{ mm}$, respectively. Samples of alloys VT5-1 and VT1-0 had an average grain size of $40 \text{ }\mu\text{m}$ and $30 \text{ }\mu\text{m}$, respectively. Samples characterized by different geometries were used for tensile test, as shown in Fig. 1(a).

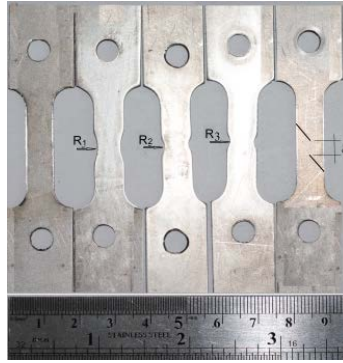


Figure 1. Geometry of the samples used for the tensile and shear tests

The samples were cut using electro erosion cutting method from a sheet of titanium alloy. The initial gauge length l_0 was equal to $24 \pm 0.1 \text{ mm}$. The cross-sectional area of the smooth flat samples was $A_0 = 9.3 \pm 0.05 \text{ mm}^2$, and notch flat specimens had notch radii of $R_1 = 10 \text{ mm}$, $R_2 = 5 \text{ mm}$, $R_3 = 2 \text{ mm}$, respectively. The samples for shear test had distance between cuts $\delta \sim 4.5 \text{ mm}$.

The stress triaxiality η and Lode parameter L defined as [25]:

$$\eta = -p/\sigma_{eq}, L = (2\sigma_{II} - \sigma_I - \sigma_{III})/(\sigma_I - \sigma_{III}), \quad (1)$$

where p is the pressure, σ_I , σ_{II} , σ_{III} are first, second, and third invariants of the Cauchy stress tensor respectively, $\sigma_{eq} = (3\sigma_{II})^{1/2}$ is the equivalent stress.

The initial value of the stress triaxiality η calculated by the analytical formula in the plane stress state [14]:

$$\eta = (1 + 2D) / (3\sqrt{D^2 + D + 1}), D = \ln[1 + a / (4R)], \quad (2)$$

where a is the width of sample, R is the value of curvature radius.

The initial value of the stress triaxiality has varied in the range 0.0 – 0.6 in experiments.

The stress triaxiality η will be slightly different from the initial value when the sample neck is formed.

2 Quasistatic and dynamic tests

The tests were carried out in a range of strain rates ($0.001-1000 \text{ s}^{-1}$) at room temperature using the Instron test machine VHS 40/50-20 with a 50 kN load cell.

The tensile force and displacement of the specimen were recorded at high temporal resolution up to complete fracture of the specimen. The constant tension velocities (20 ± 0.01 , 12 ± 0.01 , 2.4 ± 0.002 , $0.4 \pm 0.001 \text{ m/s}$) were supported during tests.

Tests were divided into three groups. First group is uniaxial tensile tests carried out on smooth specimens, characterized by positive values of both the stress triaxiality and Lode parameter. The second group is pure shear tests. The third group is uniaxial tensile tests carried out on notched specimens. Three values of notch radius, 2 mm, 5 mm and 10 mm, were used in this study.

True strain and true stress at time moments of loading were determined by analytical relations [11]:

$$\varepsilon_1^{true} = \ln(1 + \Delta l / l_0), \quad \sigma_1^{true} = (F / A_0)(1 + \Delta l / l_0), \quad (3)$$

where ε_1^{true} is true strain, σ_1^{true} is true stress, F is tensile force, A_0 is mean initial minimum cross sectional area of sheet sample, Δl , and l_0 are the elongation and the initial length of sample working part.

True stresses and true strains were also determined by computer simulation of the specimen deformation process. The results of numerical simulation were used in which a good agreement was reached between the calculated and experimental values of the sample forces and displacements versus time.

3 Computational model

The computational model uses the theoretical basis of continuum damage mechanics [26]. Mechanical behaviour was described by a system of conservation equations (mass, momentum and energy), kinematic equation and the constitutive equation. Initial and boundary conditions were added to the system of equations. The boundary conditions correspond to the conditions of loading of the 3D body. Dog bone specimens were simulated under axial tension with a constant strain rate. The computer simulations were performed with the use of licensed LS DYNA software, and being a part of the package simulation of the dynamic loads of the package ANSYS WB 15.2. The calculations were carried out using solvers using finite-difference scheme of second order accuracy. The grid model of the samples used in tests is shown in Fig. 2.

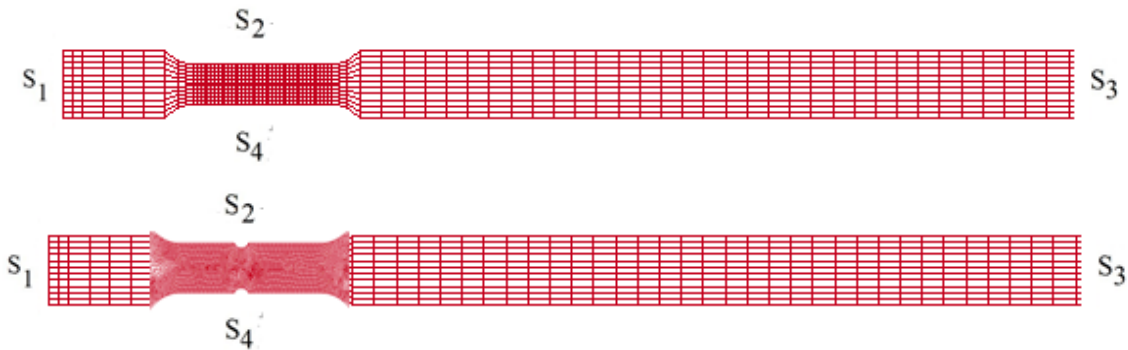


Figure 2. Grid models of flat samples

Boundary conditions corresponding to uniaxial tension of the sample at a constant strain rate have the form:

$$\begin{aligned} u_x|_{S_1} &= 0, & u_x|_{S_3} &= 0, \\ u_y|_{S_1} &= 0, & u_y|_{S_3} &= v_{y_0}, \\ u_z|_{S_1} &= 0, & u_z|_{S_3} &= 0, \\ \sigma_{ij}|_{S_2 \cup S_4 \cup S_5 \cup S_6} &= 0, \end{aligned} \quad (4)$$

where $u_i|_{S_j}$ is the components of the particle velocity vector on the surface S_j , v_{y_0} is the tensile velocity, σ_{ij} is the components of the stress tensor.

The initial conditions correspond to the free stress state of the material in a uniform temperature field.

The flow stress of alloys has been described using a modification of the Johnson-Cook model (5) and the Zerilli–Armstrong model (6) [2] [11][27]

$$\sigma_s = \{A + B(\varepsilon_{eq}^p)^n + k d_g^{-1/2}\} [1 + C \ln(\dot{\varepsilon}_{eq} / \dot{\varepsilon}_{eq0})] \{1 - [\frac{T - T_0}{T_m - T_0}]^m\}, \quad (5)$$

where A, B, C, n, m, k are constants of material, d_g is the grain size, $\dot{\varepsilon}_{eq} = [(2/3)\dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij}]^{1/2}$,

$\dot{\varepsilon}_0 = 1.0 s^{-1}$, $\varepsilon_{eq}^p = \int_0^t [(2/3)\dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij}]^{1/2} dt$ is the plastic strain intensity, T is the temperature, T_0 is the room temperature, and T_m is the melting temperature.

Material constants of the equation (13) of alpha titanium alloys are given in Table 1.

Table 1. Material constants of modified Johnson-Cook model

Coefficient	A, GPa	B, GPa	n	k_g , GPa nm ^{1/2}	C	m at $T \leq T_\beta$	T_m , K
Titanium alloy							
VT1-0	0,244	0,86	0,475	6,2	0,027	0.7	1946
VT5-1	0,760	0,86	0,5	6,2	0,027	0.64	1888

T_β is equal to 1313K and 1183 for VT5-1 and VT1-0, respectively.

Updating of the constitutive equation of the Armstrong - Zerilli model can be written in the following form [27]:

$$\sigma_s = \sigma_{s0} + C_5(\varepsilon_{eq}^p)^{n_1} + k_{hp} d_g^{-1/2} + C_2 \exp\{-C_3 T + C_4 T \ln(\dot{\varepsilon}_{eq} / \dot{\varepsilon}_{eq0})\}. \quad (6)$$

Material parameter $\sigma_{s0}, C_5, n_1, k_{hp}, C_2, C_3, C_4$ of alpha titanium alloys are given in Table 2.

Table 2. Material parameter of the Armstrong-Zerilli model

Coefficient	σ_{s0} , GPa	k_{hp} , GPa nm ^{1/2}	C_2 , GPa	C_3 , K ⁻¹	C_4 , K ⁻¹	C_5 , GPa	n_1
Titanium alloy							
VT1-0 (Grade 2)	0,138	6,2	0,1843	0,000877	0,0004	0.62	0.5
VT5-1 (Grade 6)	0,41	6,2	0,1843	0,000877	0,0004	1.84	0.5

Fig. 3 shows the calculated yield stress and experimental true stress versus true strain curves.

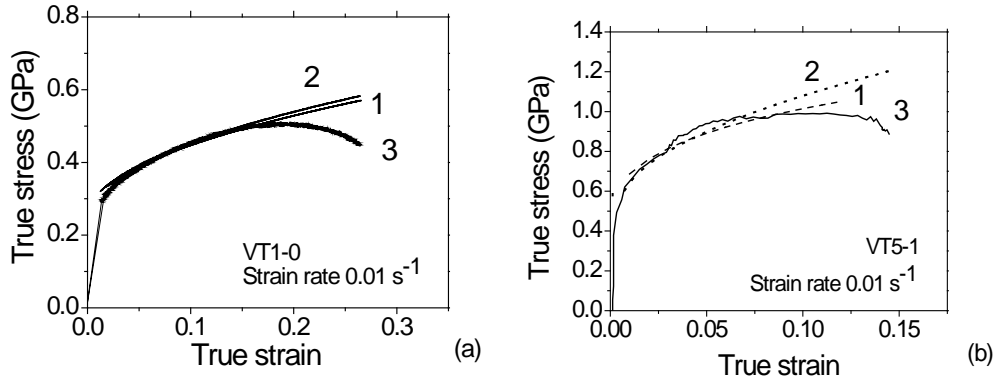


Figure 3. Calculated stress versus strain and experimental true stress vs true strain curves of (a) VT1-0 and (b) VT5-1; Johnson-Cook (curves 1) and modified Zerilli-Armstrong model prediction (curve 2)

Curves 1 were calculated by Johnson-Cook constitutive equation (5) and curves 2 by modified Zerilli-Armstrong equation (6). Curves 3 are experimental data on VT1-0 (Fig3(a)) and VT5-1 titanium alloys (Fig 3(b)). Grain size was assumed $\sim 30 \mu\text{m}$ for VT1-0 and $40 \mu\text{m}$ for VT5-1. Experimental and calculated yield stresses correspond to strain rate of 0.01 s^{-1} .

4 Damage model

Ductile fracture is considered as a result of void nucleation, growth and coalescence under a triaxle stress states or shear localization at low stress triaxiality [10]. One of the most widely used damage models for ductile fracture was proposed by Gurson with subsequent development in the works of Tvergaard and Needleman [15][20][21]. This Gurson-Tvergaard-Needleman (GTN) model [20][21] was used for analysis of stresses and strains in smooth and notched samples of sheet VT 5-1 under tension. The yield criterion has a form:

$$(\sigma_{eq}^2 / \sigma_s^2) + 2q_1 f^* \cosh(-q_2 p / 2\sigma_s) - 1 - q_3 (f^*)^2 = 0, \quad (7)$$

where σ_s is the yield stress, p is the pressure, q_1 , q_2 and q_3 are model parameters, f is the void volume fraction.

The rate of void growth is obtained by assuming mass conservation and depends on the volume change part of the plastic strain.

Consequently, there is no void growth in pure shear deformation. The void nucleation depends on the equivalent plastic strain ϵ_p , here a normal distribution A is used.

A strong coupling between deformation and damage is introduced by a plastic potential function which is dependent on the void volume fraction f^* :

$$\begin{aligned} \dot{f} &= \dot{f}_{nucl} + \dot{f}_{growth}, \\ \dot{f}_{nucl} &= (f_N / s_N) \epsilon^p \exp\{-0.5[\epsilon_{eq}^p - \epsilon_N] / s_N\}^2\}, \\ \dot{f}_{growth} &= (1 - f) \dot{\epsilon}_{kk}^p, \end{aligned} \quad (8)$$

where ϵ_N and s_N are the average nucleation strain and the standard deviation respectively. The amount of nucleating voids is controlled by the parameter f_N .

$$\begin{aligned} f^* &= f \text{ if } f \leq f_c; \\ f^* &= f_c + (\bar{f}_F - f_c) / (f_F - f_c) \text{ if } f > f_c, \end{aligned} \quad (9)$$

where $\bar{f}_F = (q_1 + \sqrt{q_1^2 - q_3}) / q_3$, q_1 , q_2 , and q_3 are constants of the model.

The final stage in ductile fracture comprises in the voids coalescence into the fracture zone. This causes softening of the material and accelerated growth of the void fraction f^* until the fracture void fraction f_F is reached. At this moment the material is fractured.

The model of ductile fracture requires knowledge of 9 parameters: three model parameters (q_1 , q_2 and q_3), the initial void fraction f_0 , three void nucleation parameters (ϵ_N , s_N and f_N), two failure parameters (f_c and f_F). The model parameters for titanium alloy VT 5-1 and VT1-0 were determined by numerical simulation of experiments on the tensile samples in the velocity range from 20 to 0.4 m/s. Numerical values of model parameters were fitted

Numerical values of model parameters are given in Table 3. The model was used for simulation samples under tension at constant velocity from 20 m/s to 0.4 m/s.

Table 3. Dimensionless parameters for the GTN model for alpha titanium alloys

Parameter	q_1	q_2	q_3	f_0	f_N	f_c	f_F	ϵ_N	s_N
VT 5-1 (Grade 6)	1	0.7	1	0.00	0.156	0.117	0.26	0.05	0.1
VT1-0 (Grade 2)	1.5	1	2.25	0.002	0.017	0.26	0.303	0.3	0.1

5 Results and discussion

The character of fracture of the samples at tensile velocities of 20 m / s, 12 m / s, 2.4 m / s, and 0.4 m / s can be seen in Fig. 4. The cracks position indicates that there is a very strong correlation between the localization of the plastic deformation and damage growth.

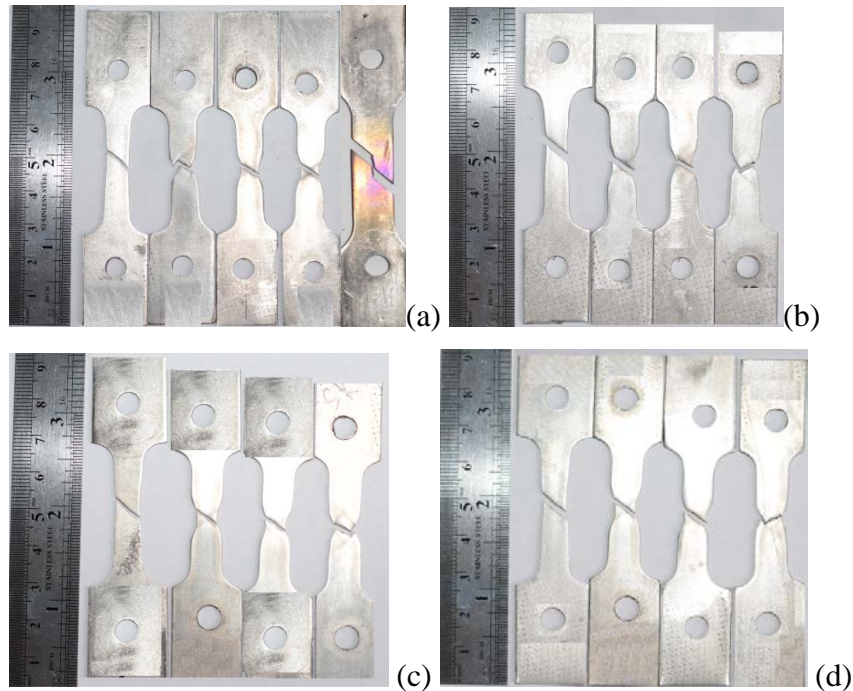


Figure 4. Fractured smooth and notched samples of VT5-1 titanium alloy after tension at the velocity: (a) 20 m/s , (b) 12 m/s, (c) 2.4 m/s, (d) 0.4 m/s

Principal tensile stress, effective stress, effective plastic strain in the notched specimen of an alpha titanium alloy VT5-1 before crack initiation are shown in Fig. 5 (a), (b), (c), respectively. Effective plastic strain near crack is shown in Fig. 5(c, d). Effective plastic strain near crack is shown in Fig. 5(c, d).

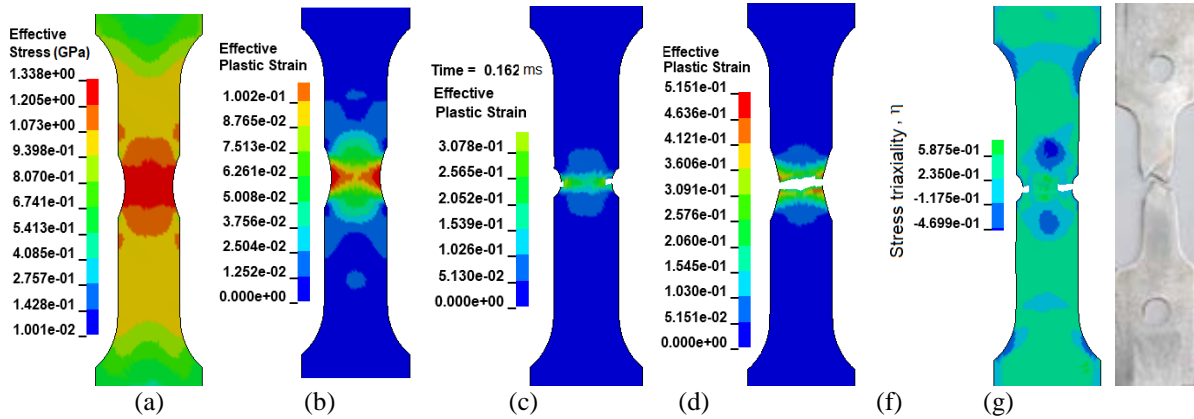


Figure 5. (a) Effective stress, (b), (c), (d) effective plastic strain in the alpha titanium notched specimen under tension at 20 m/s; (f) stress triaxiality; (g) photo of fractured notched specimen

For notched specimens under uniaxial tension damage can initiate at the outer surface of the specimen, and grows from the outer surface to the center of the specimen. The simulations demonstrate the important role of strain localization phenomena in the fracture processes. Figure 6 shows the calculated plastic strain in the necking zone before fracture. Two inclined stationary shear bands were formed in the necking zone. Fig. 5 and 6 showed the calculated configuration of cracks.

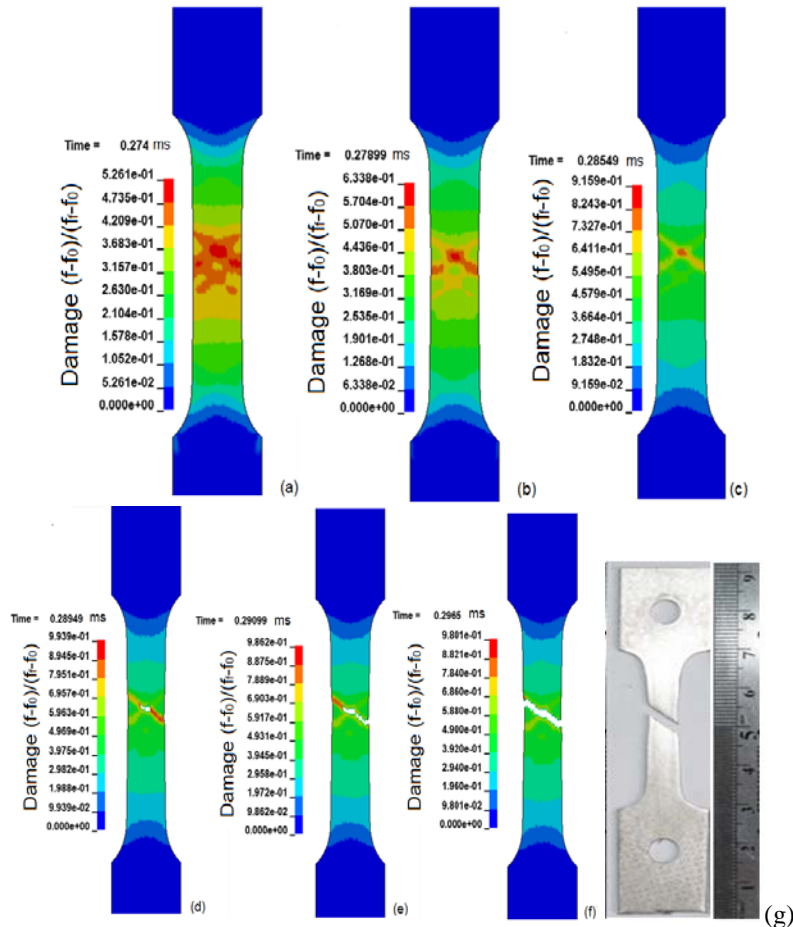


Figure 6. Damage growth and crack initiation in the alpha titanium smooth specimen under tension at 20 m/s; (g) photo of fractured specimen

Calculated configurations have good agreement with experimental data (See Fig.5(a)). Verleysen showed that the phenomena of strain localization play a major role in the fracture process of alpha+beta titanium alloy (Ti6V4Al) at lower stress triaxiality [25]. The experimental and theoretical results obtained in this paper on alpha titanium alloys agree with the results for two-phase alpha + beta titanium alloys. Damage evolution in smooth specimen of alpha titanium alloy (VT5-1) under tension is shown in Fig.6.

The growth of damage and the initiation of cracks in the smooth sample of alpha-titanium under tension at 20 m / s occur in the shear bands. Localized shear bands at the first stages of deformation are non-stationary (See Fig. 6(a,b)) and intersect at angles. Calculated configuration of cracks in the smooth and notched samples of alpha titanium alloys is in good agreement with the experimental data.

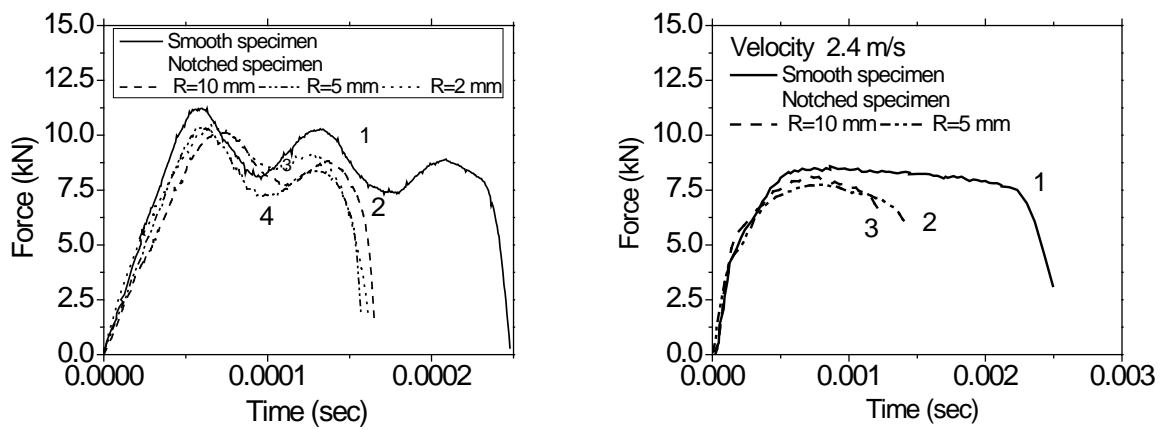


Figure7. Measured force versus time under tension smooth and notch samples (a) at 20.0±0.01 m/s and (b) 2.4 m/s

The force versus time at room temperature and tension velocity of 20 m/s and 2.4 m/s are shown in Fig. 7(a) (b), respectively. The shape of force versus time curves depends on the distance between the grips and the tensile speed.

Using «fast jaw» grips on the Instron test machine VHS 40/50-20 during high-rate tests with different velocities of tension leads to the need to account for changes in the distance between the grips when determining the true stress and true strain curves. Oscillations of the force shown in Fig 7 (a) are caused by the reflection of the force pulses moved through the specimen.

Therefore, the use of the analytical ratio (3) to analyze the results of tests at high tensile velocity leads to the need for averaging the data obtained. The true stress obtained from analytical relations (3) is underestimated relative to the values obtained by numerical simulation. The true stress versus true strain at room temperature and strain rate 834 s^{-1} (at the velocity 20 m/s) are shown in Fig. 8 (a). The introduction of the triaxiality effect in the strain hardening model is important for the prediction of damage evolution under deformation at high strain rates. Damage kinetics in alpha titanium alloys is connected with macroscale plastic instability. Strain to fracture at strain rates 833 ± 5 , 418 ± 2 , 100 ± 1 , $16.7 \pm 1 \text{ s}^{-1}$ of smooth and notched samples are shown in Fig. 8(b). The strain to fracture of the alpha-titanium alloy is reduced simultaneously with the growth of the strain rate from 0.01 to $\sim 1000 \text{ s}^{-1}$ and also reduced with an increase in the stress triaxiality η in the range from 0.33 to 0.55.

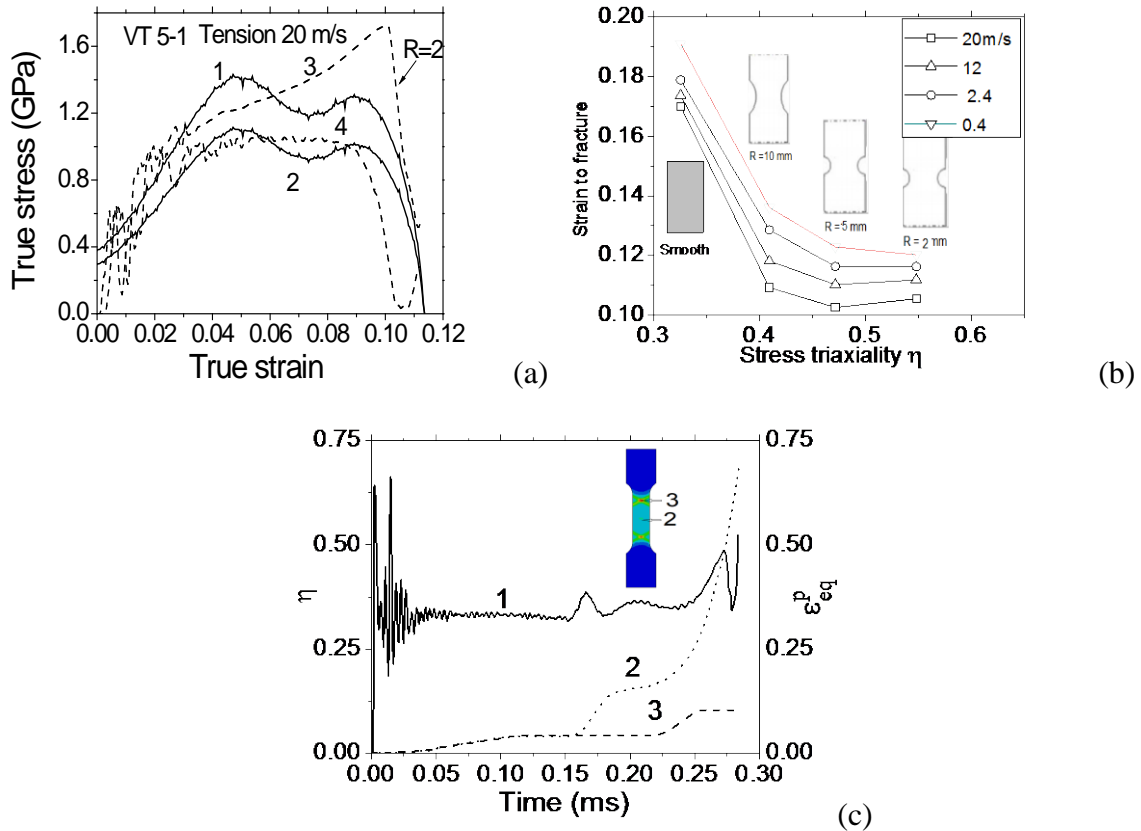


Figure 8. (a) The true stress versus true strain of titanium VT 5-1 at room temperature: analytical curves (1), and (2); calculated curves (3), and (4); (b) strain to fracture samples under uniaxial tension at velocities 20, 12, 2.4, 0.4 m/s and at the initial stress triaxiality η : 0.3333, 0.4087, 0.4681, 0.5491; (c) stress triaxiality in the gauge 2 (curve 1), equivalent plastic strain in gauges 2 and 3 (curves 2,3)

The stress triaxiality at strain rate above 100 s^{-1} is strongly influenced on the plastic flow and fracture of alpha titanium alloy. Strain to failure of alpha titanium alloys at room temperature decrease by 3.7 times with increasing stress triaxiality from 0.3 to 0.6 at strain rates of $0.01\text{--}1000 \text{ s}^{-1}$.

The results on the influence of stress triaxiality on the strain to fracture of the alpha titanium alloys obtained in this paper agree with the data of Bobbili [13].

The simulation results showed that stress triaxiality is changed in the necking zone under tension (see Fig.8(c)). The variation of η in the necking zone is caused by the shear stress relaxation at quasi-stationary shear bands formation.

The use of a model combining the description of the plastic flow of origin and the growth of damage in the calculations made it possible to compare the description of the mechanical behavior of alpha-titanium alloys in the strain rate range from 0.001 to 1000 s^{-1} and different values of stress triaxiality ($0.0 < \eta < 0.6$).

Conclusions

In this paper, mechanical behavior of titanium alloy Grade 6 (this is an analog of VT5-1 or Ti-5Al-22.5Sn) and Grade 2 (VT1-0) was studied under tension in a wide range of strain rates (from 0.001 to 1000 1/s) and stress triaxiality ($0.0\text{--}0.6$) at room temperature. Smooth and notched specimens were tested at tension velocity from 0.04 m/s to 20 m/s using an Instron VHS 40 / 50-20 servo-hydraulic test machine.

Analysis of the experimental results is supported by finite element simulations. Detailed information, complementary to the test results, is obtained on the stress and strain distribution close to the fracture. A ductile fracture at high strain rates occurs in smooth and notched samples.

A coupled elastic-plastic-damage model based on continuum damage theory used to simulate the mechanical behavior of alpha titanium alloys with HCP crystal structure.

The Gurson-Tvergaard-Needleman damage model is adopted for alpha titanium alloys, complemented with phenomenological laws for void nucleation, growth and coalescence.

The model can accurately predict both deformation and damage behaviors of alpha titanium alloys at strain rates from 0.001 to 1000 s⁻¹ and stress triaxiality from 0.3 to 0.6.

The obtained experimental data indicate that there is a very strong correlation between the characteristics of the plastic deformation and the rate of damage growth.

The constitutive and failure model parameters can be determined on the base of tensile test results. The constitutive and fracture models have been validated by simulating the tension tests.

Acknowledgement

This work was supported partially by the Grant from the President of Russian Federation MK-2690.2017.8 and by The Tomsk State University competitiveness improvement programme. The authors are grateful for the support of this research.

References

- [1] Frost H.J., Ashby M.F., *Deformation-mechanism maps*. Oxford: Pergamon Press; 1982. 166 p.
- [2] Gao C.Y., Zhang L.C., and Yan H.X. (2011) A new constitutive model for HCP metals, *Materials Science and Engineering A* **528**, 4445–4452.
- [3] Xu Q., Meyers M.A., and Nesterenko V.F. (2002) Self-organization of shear bands in titanium and Ti–6Al–4V alloy, *Acta Materialia* **50**, 575–596.
- [4] Herzig N., Meyer L.W., Musch D., Halle T., Skripnyak V.A., Skripnyak E.G., Razorenov S.V., Krüger L. (2008) The mechanical behaviour of ultrafine grained titanium alloys at high strain rates, *3rd International Conference on High Speed Forming* – March 11–12, 2008. Dortmund, Germany. -14 p. DOI:10.17877/DE290R-8660
- [5] Skripnyak V. A. and Skripnyak E. G. *Mechanical behaviour of nanostructured and ultrafine-grained metal alloy under intensive dynamic loading*, Chapter 2: *Nanotechnology and Nanomaterials*» "Nanomechanics", Eds.by A. Vakhrushev, ISBN 978-953-51-3182-3, Print ISBN 978-953-51-3181-6, 2017. DOI: 10.5772/Intech open.68291.
- [6] Skripnyak V.A., Skripnyak N.V., Skripnyak E.G., Skripnyak V.V. (2017) Influence of grain size distribution on the mechanical behavior of light alloys in wide range of strain rates, *AIP Conference Proceedings*, 1793, art. no. 110001.
- [7] Skripnyak N.V., Skripnyak E.G., Skripnyak V.A., Skripnyak V.V., Vaganova I.K. (2014) Failure mechanisms of light alloys with a bimodal grain size distribution *11th World Congress on Computational Mechanics, WCCM 2014, 5th European Conference on Computational Mechanics, ECCM 2014 and 6th European Conference on Computational Fluid Dynamics, ECFD*, 3915–392.
- [8] Sharkeev Y.P., Vavilov V.P., Belyavskaya O.A., Skripnyak V.A., Nesteruk D.A., Kozulin A.A., Kim V.M. (2016) Analyzing deformation and damage of VT1-0 titanium in different structural states by using infrared thermography, *Journal of Nondestructive Evaluation*, 35 (3), Art. no. 42.

- [9] Ratochka I.V., Lykova O.N., Geraskina A.Yu. and Skripnyak V.A. (2009) Grain boundary sliding in a submicrocrystalline Ti–6Al–4V titanium alloy under superplastic deformation, *Fiz. Mezomekh* **12**, No. 5, 97-101.
- [10] Y. Bai and T. Wierzbicki, A new model of metal plasticity and fracture with pressure and Lode dependence, *International Journal of Plasticity* **24**, 1071–1096, 2008.
- [11] Valoppi B., Bruschi S., Ghiotti A., Shivpur R. (2017) Johnson-Cook based criterion incorporating stress triaxiality and deviatoric effect for predicting elevated temperature ductility of titanium alloy sheets, *International Journal of Mechanical Sciences* **123**, 94-105.
- [12] Revil-Baudard B., Cazacu O., Flater P., Chandola N., and Alves J.L. (2016) Unusual plastic deformation and damage features in titanium: Experimental tests and constitutive modeling, *Journal of the Mechanics and Physics of Solids* **88**, 100-122.
- [13] Bobbili R., and Madhun V. (2016) Flow and fracture characteristics of near alpha titanium alloy, *Journal of Alloys and Compounds* **684**, 162-170.
- [14] Selini N., Elmequenni M., Benguediab M. (2013) Effect of the Triaxiality in Plane Stress Conditions, *Eng. Technol. & Appl. Sci. Res.* **3**, 373–380.
- [15] Kondori B., Benzerga A. A., Needleman A. (2018) Discrete shear-transformation-zone plasticity modeling of notched bars, *Journal of the Mechanics and Physics of Solids* **111**, 18–42.
- [16] Zhai J., Luo T., Gao X., Graham S. M., Knudsen E. (2016) Modeling the ductile damage process in commercially pure titanium, *International Journal of Solids and Structures* **91**, 26-45.
- [17] Lindner D., Mathieu F., Hild F., Allix O., Minh C.-H., Paulien-Camy O. (2015) On the evaluation of stress triaxiality fields in a notched titanium alloy sample via integrated digital image correlation, *Transactions of the ASME*, **82**, 071014, 1–10.
- [18] Heibel S., Nester W., Clausmeyer T., Tekkayan A. E. (2017) Failure assessment in sheet metal forming using a phenomenological damage model and fracture criterion: experiments, parameter identification and validation, *Procedia Engineering*, **207**, 2066–2071.
- [19] Orozco-Caballero A., Li F., Esqué-de los Ojos D., Atkinson M. D., Quinta da Fonseca J. (2018) On the ductility of alpha titanium: The effect of temperature and deformation mode, *Acta Materialia*, **149**, 1-10.
- [20] Neilsen K.L., and Tvergaard V. (2010) Ductile shear failure or plug failure of spot welds modelled by modified Gurson model, *Eng. Fract. Mech.*, **77**, 1031–1047.
- [21] Tvergaard V. (2015) Study of localization in a void-sheet under stress states near pure shear, *International Journal of Solids and Structures*, **60–61**, 28–34.
- [22] Zhai J., Luo T., Gao X., Graham S. M., Baral M., Korkolis Y. P., Knudsen E. (2016) Modeling the ductile damage process in commercially pure titanium, *International Journal of Solids and Structures*, **91**, 26–45.
- [23] Corigliano A., Mariani S. and Orsatti B. (2000) Identification of Gurson–Tvergaard material model parameters via Kalman filtering technique. I. Theory, *International Journal of Fracture*, **104**, 349–373.
- [24] Mirone G., Barbagallo R., Corallo D. (2016) A new yield criteria including the effect of lode angle and stress triaxiality, *Procedia Structural Integrity*, **2**, 3684–3696.
- [24] Springmann M., Kuna M. (2005) Identification of material parameters of the Gurson–Tvergaard–Needleman model by combined experimental and numerical techniques, *Computational Materials Science*, **32**, 544–552.
- [25] Verleysen P. and Peirs J. (2017) Quasi-static and high strain rate fracture behaviour of Ti6Al4V, *International Journal of Impact Engineering*, **108**, 370-388.
- [26] Zhang W., Cai Y. Continuum Damage Mechanics and Numerical Applications. *Springer Science & Business Media*, 2010 - Technology & Engineering, 1000 p.

- [27] Zerilli F.J., Armstrong R.W. (1992) The effect of dislocation drag on the stress-strain behaviour of F.C.C. metals, *Acta Metall. Mater.*, 40, 1803–1808.

Numerical simulation of mechanical behavior of Zr-Nb alloys in a wide temperature range

†Vladimir A. Skripnyak¹, Evgeniya G. Skripnyak¹, Vladimir V. Skripnyak¹,
Vitas A. Serbenta¹, and Natalia V. Skripnyak¹

¹Research Laboratory of Properties of Substances in Extreme States, National Research Tomsk State University, Russia

*Presenting author: skrp@ftf.tsu.ru

†Corresponding author: skrp@ftf.tsu.ru

Abstract

The aim of this work was the evaluation of temperature and strain rate on the mechanical behavior of the zirconium-niobium alloys. This paper presents results on modelling and theoretical prediction of mechanical and deformation properties of Zr-Nb in a wide range of strain rates and temperature. Mechanical behavior of Zr-1 % Nb was numerically simulated using Johnson-Cook, and Zerilli -Armstrong constitutive equations. Material parameters of these constitutive equations were determined for Zr-1 % Nb alloys (E110). The numerical results on dynamic and quasi-static deformation of Zr-1 % Nb alloy are good agreed with experimental data. Strain rate sensitivity of the yield stress of Zr-Nb alloys at fixed temperature depends on the concentration of Nb, and parameters of grain size distribution. It is shown that the resistance to plastic deformation of Zr-Nb alloys is different under compression and tension at high-strain rates.

The results can be used in engineering analysis of designed technical systems for nuclear reactors.

Keywords: Computer simulation, mechanical behavior, ductility, zirconium-niobium alloys, high strain rate, temperature

Introduction

Improvement of technology of fabrication of fuel claddings and some constructional elements of nuclear reactors is connected with computer simulation of mechanical properties and structural evolution of radiation-resistant alloys Zr-Nb [1]. In this regard, there is an increasing need to develop computational models of the mechanical behaviour of advanced Zr-Nb in loading conditions close to operating ones. The Zr-Nb has a unique complex of physical and mechanical properties and is considered as promising structural alloys for nuclear reactors of IV generation. Coarse grained (CG) and ultrafine grained (UFG) zirconium alloys with a concentration of Nb below 2.5 weight % and additionally doped with Mo, Fe, Cr for the stabilization of precipitations of beta-phase Zr were studied during last decade [2]-[4]. It is known that the formation of ultrafine-grained structures in the alloys of Zr-Nb not only improves the yield and strength of the alloy, but also prevents the formation of cracks at the mesoscopic level [5]-[6]. In this connection, the mechanical behavior of Zr-Nb alloys was studied by numerical simulation method in the practically important temperature range from 297 K to 1243 K.

1 Material and samples

Mechanical behavior of ultrafine -grained (UFG) and coarse grained (CG) zirconium E110 (Zr-1%Nb) alloy was studied under tension at room temperature. Ultrafine grained specimens

of E110 alloy were produced by the severe plastic deformation method [7]. Sizes of specimens are shown in Table 1.

Table 1. Sizes of specimens of E100 (Zr-1%Nb) alloy

Coarse grained E110 alloy ($d_g^{(*)} \approx 55 \mu\text{m}$)			
Designation	Width of gage part of specimens (mm)	Thickness of specimens (mm)	Length of gage part of specimens (mm)
19	8.9±0.01	0.88±0.01	20.3±0.01
20	9.0±0.01	0.90±0.01	20.2±0.01
21	9.0±0.01	0.90±0.01	20.4±0.01
22	8.9±0.01	0.92±0.01	20.8±0.01
23	8.9±0.01	0.87±0.01	20.7±0.01
24	8.9±0.01	0.87±0.01	20.5±0.01
25	9.0±0.01	0.90±0.01	20.5±0.01
Ultrafine grained E110 (grain size $d_g \approx 0.5 \mu\text{m}$)			
26	9.0±0.01	0.90±0.01	20.5±0.01
27	9.0±0.01	0.90±0.01	20.3±0.01
28	9.0±0.01	0.85±0.01	20.3±0.01
29	8.9±0.01	0.90±0.01	20.5±0.01
30	9.0±0.01	0.90±0.01	20.5±0.01
31	9.0±0.01	0.90±0.01	20.5±0.01
32	8.9±0.01	0.90±0.01	20.5±0.01
33	8.9±0.01	0.90±0.01	20.3±0.01
34	9.0±0.01	0.90±0.01	20.7±0.01

(*) d_g is the grain size.

2 Quasistatic and dynamic tests

The tests were carried out at room temperature and strain rate 0.01 s^{-1} using the Instron test machine VHS 40/50-20 with a 50 kN load cell. The tensile force and displacement of the specimen were recorded at high temporal resolution up to complete fracture of the specimen. True strain and true stress at time moments of loading were determined by analytical relations [8]:

$$\varepsilon_1^{true} = \ln(1 + \Delta l / l_0), \quad \sigma_1^{true} = (F / A_0)(1 + \Delta l / l_0), \quad (1)$$

where ε_1^{true} is true strain, σ_1^{true} is true stress, F is tensile force, A_0 is mean initial minimum cross sectional area of sheet sample, Δl , and l_0 are the elongation and the initial length of sample gage part.

True stresses and true strains were also determined by computer simulation.

3 Computational model

The computational model is based on the theoretical basis of continuum damage mechanics [9]. Mechanical behaviour was described by a system of conservation equations (mass, momentum and energy), kinematic equation and the constitutive equation. Initial and boundary conditions were added to the system of equations. The boundary conditions

correspond to the conditions of loading of the 3D body. Dog bone shape specimens were simulated under axial tension with a constant strain rate. The computer simulations were performed with the use of licensed AUTODYN software, and being a part of the package simulation of the dynamic loads of the package ANSYS WB 14.5. The calculations were carried out using solvers using finite-difference scheme of second order accuracy. The grid model of the samples used in tests is shown in Fig. 1.

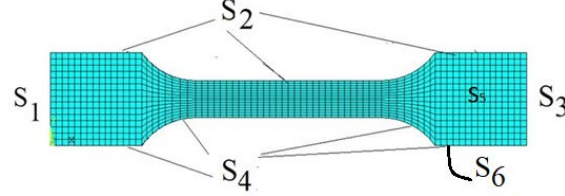


Figure 1. Grid model of flat sample

Boundary conditions corresponding to uniaxial tension of the sample at a constant strain rate have the form:

$$u_1|_{S_1}=0, u_1|_{S_3}=0; u_2|_{S_1}=0, u_2|_{S_3}=V_2^{(0)}; u_3|_{S_1}=0, u_3|_{S_3}=0; \sigma_{ij}|_{S_2 \cup S_4 \cup S_5 \cup S_6}=0, \quad (1)$$

where $u_i|_{S_j}$ are the components of the particle velocity vector on the surface S_j , v_{y_0} is the tensile velocity, σ_{ij} is the components of the stress tensor.

The initial conditions correspond to the free stress state of the material in a uniform temperature field.

$$\rho|_{t=0}=\rho_0, \varepsilon_{ij}|_{t=0}=0, \sigma_{ij}|_{t=0}=0, T|_{t=0}=T_{\text{test}}, \quad (2)$$

where t is a time, ρ is a mass density, T_{test} is the initial temperature of specimen.

The flow stress of zirconium alloys under loading has been described using a modification of the Johnson-Cook model (3) and the Zerilli–Armstrong model (4) [8][10]. The flow stress of alloys has been described using a modification of the Johnson-Cook model (3) and the Zerilli–Armstrong model (4):

$$\sigma_s = \{A + B(\varepsilon_{\text{eq}}^p)^n + k d_g^{-1/2}\} [1 + C \ln(\dot{\varepsilon}_{\text{eq}} / \dot{\varepsilon}_{\text{eq}0})] \{1 - [\frac{T - T_0}{T_m - T_0}]^m\}, \quad (3)$$

where A, B, C, n, m, k are constants of material, d_g is the grain size, $\dot{\varepsilon}_{\text{eq}} = [(2/3)\dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij}]^{1/2}$,

$\dot{\varepsilon}_0 = 1.0 \text{ s}^{-1}$, $\varepsilon_{\text{eq}}^p = \int_0^t [(2/3)\dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij}]^{1/2} dt$ is the plastic strain intensity, T is the temperature, T_0 is the room temperature, and T_m is the melting temperature.

Material constants of alloy the equation (3) of E110 (Zr-1% Nb) are given in Table 1.

Table 1. Material constants of modified Johnson-Cook model for E100 (Zr-1%Nb)

Coefficient	A,	B,	n	k_g ,	C	m	T_m ,
Zr-Nb alloy	GPa	GPa		$\text{GPa nm}^{1/2}$			K
E110	0,290	0,386	0,11	368 for $1.1 \mu\text{m} < d_g < 100 \mu\text{m}$	0,14	0.6 at $T < 1070 \text{ K}$ 0.14 at $T > 1070 \text{ K}$	1946

T_β is equal to 1070 K for Zr-1%Nb.

Updating of the constitutive equation of the Armstrong - Zerilli model can be written in the following form [10]:

$$\sigma_s = \sigma_{s0} + C_5 (\varepsilon_{eq}^p)^{n_1} + k_{hp} d_g^{-1/2} + C_2 \exp\{-C_3 T + C_4 T \ln(\dot{\varepsilon}_{eq} / \dot{\varepsilon}_{eq0})\}. \quad (4)$$

where σ_{s0} , C_5 , n_1 , k_{hp} , C_2 , C_3 , C_4 are material coefficients.

Material coefficients of E110 (Zr-1% Nb) alloy are given in Table 2.

Table 2. Material parameter of the Armstrong-Zerilli model

Coefficient	σ_{s0} ,	k_{hp} ,	C_2 ,	C_3 ,	C_4 ,	C_5 ,	n_1
Zr-Nb alloy	GPa	GPa nm ^{1/2}	GPa	K ⁻¹	K ⁻¹	GPa	
E110	0.110	368 for 1.1 μm < d _g < 100 μm	1.015	8.77	3.95	4.05	0.19
(Zr-1%Nb)	0.445	40 for 0.08 μm < d _g < 1.1 μm	1.015	10 ⁻³	10 ⁻⁴	10 ⁻³	

Results and discussion

True stress versus true strain curves of E110 alloy are shown in Fig.2. The obtained data are in good agreement with the results [11]-[14].

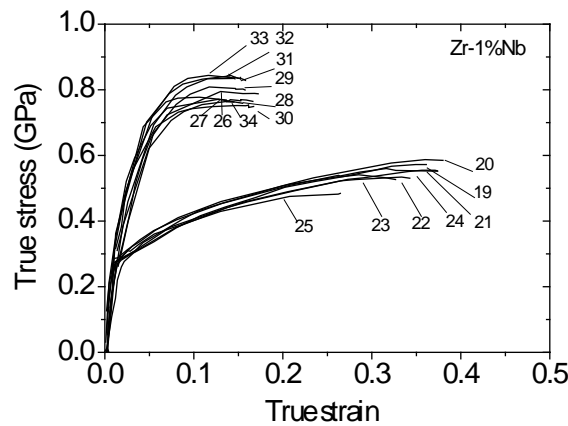


Figure 2. True stress versus true strain curves of ultrafine grained (curves 26-33) and coarse grained (curves 19-25) E100 (Zr-1%Nb) alloy under tension at 0.01 s⁻¹ strain rate

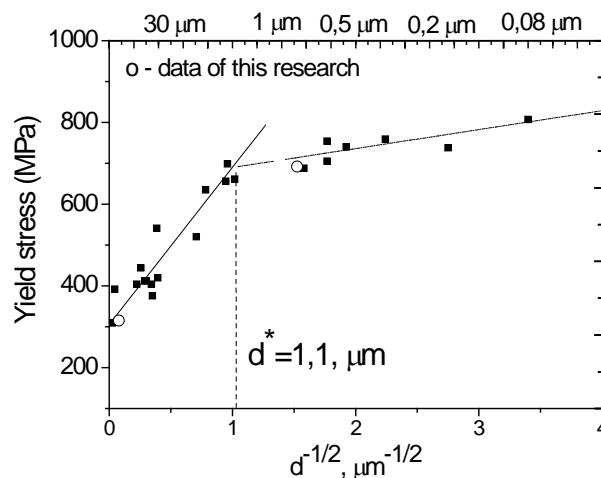


Figure 3. Yield stress versus grain size of zirconium alloys; symbols are experimental data [6][11]-[14]

The dependence of the yield stress on the grain size of coarse grained Zr–1%Nb alloys can be described by the Hall-Petch relation

$$\sigma_s = \sigma_{s0} + k_{hp} d_g^{-1/2}, \quad (5)$$

where σ_s is the yield strength, σ_{s0} and k_{hp} are parameters of material.

Parameters σ_{s0} and k_{hp} of zirconium-niobium alloys is equal to $\sigma_{s0} = 322$ MPa, $k_{hp} = 368$ MPa $\mu\text{m}^{1/2}$, when grain size is in the range $1.1 \mu\text{m} < d_g < 100 \mu\text{m}$. The values of σ_{s0} and k_{hp} change to $\sigma_{s0} = 652$ MPa, $k_{hp} = 40$ MPa $\mu\text{m}^{1/2}$ for alloys with grain size in the range of $0.08 \mu\text{m} < d < 1.1 \mu\text{m}$. These parameters were used for numerical simulation.

Numerical values of σ_{s0} and k_{hp} depend on distribution of grain sizes [15]–[17].

Fig. 4(a) shows the calculated stress versus equivalent plastic strain curves for uniaxial tension of Zr–1%Nb–1%Sn alloy (E110) at strain rate of 10^{-3} s^{-1} .

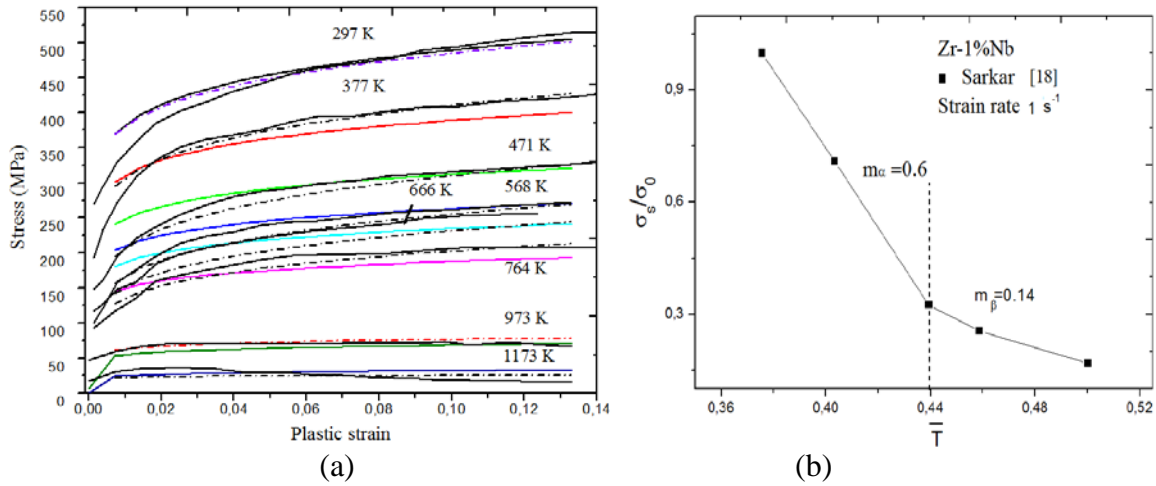


Figure 4. (a) Stress versus plastic strain under uniaxial tension of Zr-1Nb-1Sn alloy at the strain rate of 10^{-3} s^{-1} ; (b) normalized yield strength versus normalized temperature of Zr-1%Nb alloy at strain rate 1 s^{-1} (b). Symbols are experimental data [18]

Solid black curves indicate experimental true stress versus true strain [2]–[3]. Colored and dashed curves were obtained using equation (4) and (3), respectively. Results of simulation good agree with experimental data within temperature range from 297 K to 1173 K [2]–[3][18][19]. The constitutive equation (4) describes the change of the strain hardening in the temperature range more adequately in comparison with the constitutive equation (3).

The dependence of the normalized yield strength of alloy Zr–1% Nb under tension with a strain rate of 1 s^{-1} on the normalized temperature $\bar{T} = (T - T_r)/(T_m - T_r)$ is shown in Fig. 4(b).

The change of the curve slope σ_s/σ_0 (\bar{T}^v) at $\bar{T} = 0.44$ ($T = \sim 1070 \text{ K}$) is the result of a phase transition in Zr–1% Nb alloys. Phase transition from alpha phase (HCP lattice) to the beta phase (BCC lattice) depends on concentration of niobium.

Numerical means of m or C_3 , in constitutive equations (3) and (4) are changed if the temperature exceeds the temperature of beginning of phase transformations, $\alpha \rightarrow \beta$ ($\sim 1070 \text{ K}$).

Fig. 5 shows calculated yield strength versus logarithm of normalized strain rate under tension of the Zr–Nb. The average grain size was $15 \mu\text{m}$. Solid curves calculated at temperatures from 295 K to 1273 K and strain rates from 10^{-3} s^{-1} to 10^2 s^{-1} . The dashed curve calculated at the room temperature, and range of strain rates from 10^{-3} to 10^6 s^{-1} . Experimental data [4][18]–[19] are shown by filled symbols. Thus, it was shown that the dependence of the yield strength of the Zr–1%Nb alloy on the logarithm of normalized strain rate is close to linear in the temperature range from 297 K to 1273 K and strain rates from 10^{-3} to 10^2 s^{-1} .

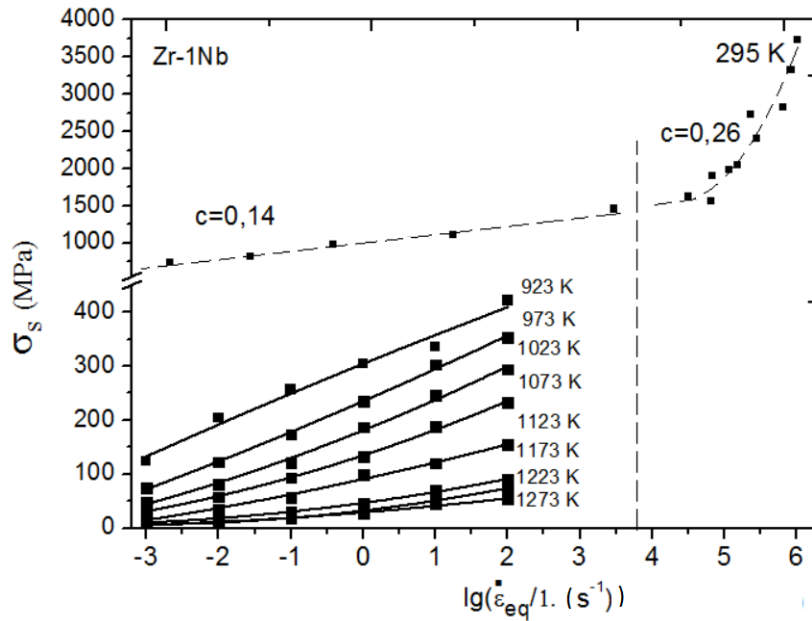


Figure 5. Normalized yield strength versus logarithm of normalized strain rate under tension for the samples of the Zr-1%Nb alloy; symbols are experimental data [4][18]-[19]

Both models (Zerilli–Armstrong and Johnson–Cook) allow obtaining satisfactory predictions of the yield stress under tension in the range of strain rates from 10^{-3} to $\sim 10^3 \text{ s}^{-1}$ and temperature from 297 K to 1273 K. It was found, that it is necessary to change the numerical value of the coefficient c in the model of Johnson–Cook to obtain a satisfactory agreement of calculated yield strength of with experimental data [18]-[19]. Calculated stress – strain curves of Zr–1Nb under tension and compression and calculated stress versus equivalent plastic strain of UFG and CG Zr–1Nb is shown in Fig. 6 (a) and (b), respectively. Results of simulation is showed that the macroscopic flow stress of Zr–Nb alloys under dynamic compression and tension are various (See Fig.6(a)). This effect is caused by instabilities of plastic flow of Zr–1%Nb at the macro– and meso– scale levels under quasi-static and dynamic loads.

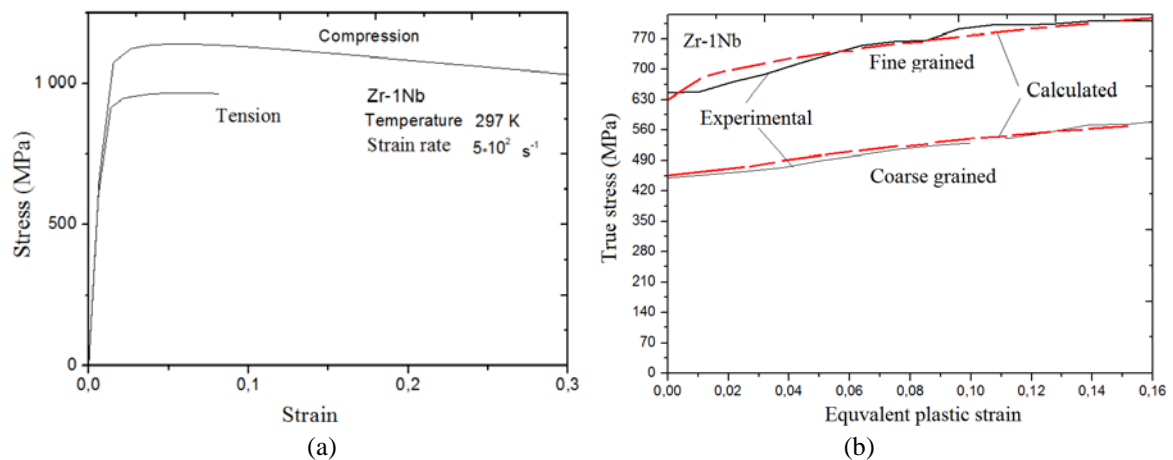


Figure 6. (a) Calculated stress versus strain curves of Zr–1%Nb under tension and compression; (b) calculated stress versus equivalent plastic strain of UFG and CG Zr-1%Nb (b); solid curves are experimental data [5]; dashed curves are calculations by the Zerilli–Armstrong model

Results of numerical simulation showed decreasing of the yield stress under shear bands origin. Calculated stress – strain curves for CG and UFG specimens of the alloy E110 (Zr–1%Nb) at room temperature are shown in Figure 6 (b). The calculated yield stresses of ultrafine-grained Zr-1%Nb alloys are in good agreement with the experimental data due to using bilinear relation (5).

Conclusions

Mechanical behavior of ultrafine-grained and coarse grained Zr–Nb alloys was studied by experimental and computer simulation methods.

Experimental true stress versus true strain curves of E110 (Zr-1%Nb) alloy were obtained at room temperature.

Modified Zerilli–Armstrong and Johnson–Cook constitutive equations were used for numerical simulation of response of specimens under tension and compression.

Modifications of both models allow obtaining satisfactory predictions of the yield stress under tension within the range of strain rates from 10^{-3} to $\sim 10^3$ s⁻¹ and temperature from 297 K to 1273 K.

It was shown that the dependence of normalized yield strength of Zr-1%Nb on normalized temperature can be approximated by a bilinear relation. The bilinear relation is due to change of mechanical properties of alloy caused by phase transition $\alpha \rightarrow \beta$ ($T \sim 1070$ K) in Zr–1%Nb alloys.

Acknowledgement

This work was supported partially by the Grant from the President of Russian Federation MK-2690.2017.8 and by The Tomsk State University competitiveness improvement programme. The authors are grateful to Professor Y. V. P. Sharkeev and researcher O. A. Belyavskaya for useful discussions and cooperation in experimental studies of alloys.

References

- [1] Blokhin D. A., Chernov V. M., Blokhin A. I. (2011) Nuclear and physics properties of zirconium alloys E-110 and E-635 under long time neutron irradiation in the VVER-1000 reactor *Advanced materials*, № 5 23–29.
- [2] Fong R. W. L. (2013) Anisotropic deformation of Zr–2.5Nb pressure tube material at high temperatures *Journal of Nuclear Materials*, **440** 467–476.
- [3] Kapoor R., Wadekar S. L., Chakravartty J. K. (2002) Deformation in Zr–1Nb–1Sn–0.1Fe using stress relaxation technique *Materials Sci. and Eng. A*. **328** 324–333.
- [4] Xiao D., Li Y., Hu S. (2010) High Strain Rate Deformation Behavior of Zirconium at Elevated Temperatures *Journal of Materials Science & Technology* **26** 878–882.
- [5] Stepanova E. N., Grabovetskaya G. P. (2015) Effect of hydrogen on the structural and phase state and the deformation behavior of the ultrafine-grained Zr–1Nb alloy *Journal of Alloys and Compounds* **645** 271–274.
- [6] Cao W.Q., Yu S.H., Chun Y.B. (2005) Strain path effects on the microstructure evolution and mechanical properties of Zr702 *Materials Sciences and Engineering A*. **395** 77–86.
- [7] Sharkeev Y. P., Eroshenko A. Y., Kulyashova K. S. (2013) Microstructure, mechanical and biological properties of zirconium alloyed with niobium after severe plastic deformation *Material wissenschaft und Werkstofftechnik* **44** 198–204.
- [8] Valoppi B., Bruschi S., Ghiotti A., Shivpur R. (2017) Johnson-Cook based criterion incorporating stress triaxiality and deviatoric effect for predicting elevated temperature ductility of titanium alloy sheets, *International Journal of Mechanical Sciences* **123** 94–105.

- [9] Zhang W., Cai Y. Continuum Damage Mechanics and Numerical Applications. *Springer Science & Business Media*, 2010 - *Technology & Engineering*, 1000 p.
- [10] Zerilli F.J., Armstrong R.W. (1992) The effect of dislocation drag on the stress-strain behaviour of F.C.C. metals, *Acta Metall. Mater* **40** 1803–1808.
- [11] Kondo R., Nomura N., Suyalatu, Tsutsumi Y., Doi H., Hanawa T. (2011) Microstructure and mechanical properties of as-cast Zr–Nb alloys *Acta Biomaterialia* **7** 4278–4284.
- [12] Ruestes C.J., Bertolino G., Ruda M. (2014) Grain size effects in the deformation of [0001] textured nanocrystalline Zr, *Scripta Materialia* **71** 9–12
- [13] Guo D., Zhang Z., Zhang G., Li M., Shi Y., Ma T., Zhang X. (2014) An extraordinary enhancement of strain hardening in fine-grained Zirconium *Materials Sciences and Engineering A*. **591** 167–172.
- [14] Jiang L., Perez-Prado M.T., Gruber P.A. (2008) Texture, microstructure and mechanical properties of equiaxed ultrafine-grained Zr fabricated by accumulative roll bonding// *Acta Materialia* **56** 1228–1242.
- [15] Skripnyak V. A. and Skripnyak E. G. (2017) *Mechanical behaviour of nanostructured and ultrafine-grained metal alloy under intensive dynamic loading*, Chapter 2: *Nanotechnology and Nanomaterials» "Nanomechanics"*, Eds.by A. Vakhrushev, ISBN 978-953-51-3182-3, Print ISBN 978-953-51-3181-6, 2017. DOI: 10.5772/Intech open. 68291.
- [16] Skripnyak V.A., Skripnyak N.V., Skripnyak E.G., Skripnyak V.V. (2017) Influence of grain size distribution on the mechanical behavior of light alloys in wide range of strain rates, *AIP Conference Proceedings*, 1793, art. No. 110001.
- [17] Skripnyak N.V., Skripnyak E.G., Skripnyak V.A., Skripnyak V.V., Vaganova I.K. (2014) Failure mechanisms of light alloys with a bimodal grain size distribution *11th World Congress on Computational Mechanics, WCCM 2014, 5th European Conference on Computational Mechanics, ECCM 2014 and 6th European Conference on Computational Fluid Dynamics, ECFD*, 3915-392.
- [18] Sarkar A., Chandanshive S. A., Thota K. M. (2017) High temperature deformation behavior of Zr-1Nb alloy *Journal of Alloys and Compounds* **703** 56-66.
- [19] Kazakov D. N., Kozelkov O. E., Mayorova A. S. (2015) Dynamic behavior of zirconium alloy E110 under sub microsecond shock-wave loading *EPJ Web of Conferences* **94** 1–5.

An optimal reconstruction of Chebyshev-Halley type methods with local convergence analysis

Ali Saleh Alshomrani¹, Ioannis K. Argyros² and †Ramandeep Behl¹

¹Department of Mathematics, King Abdulaziz University, Jeddah 21577, Saudi Arabia.

² Cameron University, Department of Mathematics Sciences Lawton, OK 73505, USA

†Corresponding author: ramanbeh187@yahoo.in

Abstract

In this manuscript, our principle aim is to present a new reconstruction of classical Chebyshev-Halley scheme having optimal fourth and eighth-order convergence for all α unlike the earlier studies. In addition, we analyze the local convergence of them by using supposition requiring the first-order derivative of the involved function f and the Lipschitz conditions. The new approach is not only the extension of earlier studies, but also formulates their theoretical radius of convergence. Several numerical examples originated from real life problems demonstrate that they are applicable to a broad range of scalar equations where previous studies cannot be used. Finally, dynamic study of them also demonstrates that bigger and promising basins of attractions belongs to our iteration functions.

Keywords: Nonlinear equations, Newton's method, Complex dynamics, Chebyshev-Halley method, local convergence analysis.

Introduction

Among the most harder and earlier issues of computational methods and numerical analysis are concerning with the cost-effective and accurate simple zeros of function $f(x)$ in a small number of iterations with specific degree of accuracy (where $f : \mathbb{D} \subset \mathbb{R} \rightarrow \mathbb{R}$ is a univariate sufficiently smooth function in the closed interval \mathbb{D}). It is hard to find analytical methods in the available literature for solving such type of problems. So, there is only one option left for us to find the approximate solutions by using iterative procedures. One of the best and most famous iterative procedure is the classical Newton's method [30, 25]. Several higher-order reconstruction of Newtons procedure have been constructed at the expense of some other values of function/s and/or its derivative/s. We have a good number of cubically convergent iterative procedures, (for the details please see [30]) and one of them is given as below:

$$x_{n+1} = x_n - \left[1 + \frac{1}{2} \frac{L_f(x_n)}{1 - \alpha L_f(x_n)} \right] \frac{f(x_n)}{f'(x_n)}, \quad \alpha \in \mathbb{R}, \quad (1)$$

where $L_f(x_n) = \frac{f''(x_n)f(x_n)}{\{f'(x_n)\}^2}$. This is a well-known family of Chebyshev-Halley iteration functions [14]. We can easily obtain some popular iteration functions from this family. For example, the classical Chebyshev's method [30, 15], Halley's method [30, 15] and super-Halley method [30, 15] if we choose $\alpha = 0$, $\alpha = \frac{1}{2}$ and $\alpha = 1$, respectively. Regardless of cubic convergence, the scheme (1) is consider less practical from a computational point of view because it is not an easy task to find the second-order derivative of every problem.

This fact has motivated many scholars to turn towards the approach of multi-point iteration functions. The principal objective of them is to produce second or higher-order derivative free iteration functions with maximum convergence order by using certain values of function/s and or its first-order derivative/s. In 1964, Traub [30] presented the analysis of multi-point iteration

functions with their properties. Recently, Petković et al. [25] also revised and update the facts about them.

Despite of going in to the detail of them, we have only focused on the higher-order and second derivative free modifications of the family (1). According to our expertise, many researchers from worldwide like, Kou and Li [18, 19], Kou [17], Chun [9], Amat et al. [1], Xiaojian [32] and Arygros et al. [5], proposed higher-order modification of Chebyshev-Halley's iteration functions not using the values of second or higher-order derivative/s.

Recently, Li et al. [21], presented an improvement of Chebyshev-Halley iteration functions, which is defined as follows:

$$\begin{cases} y_n = x_n - \frac{f(x_n)}{f'(x_n)}, \\ z_n = x_n - \left(1 + \frac{f(y_n)}{f(x_n) - 2\alpha f(y_n)}\right) \frac{f(x_n)}{f'(x_n)}, \quad \alpha \in \mathbb{R} \\ x_{n+1} = z_n - \frac{f(z_n)}{f'(x_n) + f''(x_n)(z_n - x_n)}, \end{cases} \quad (2)$$

where $f''(x_n) = \frac{2f(y_n)f'(x_n)^2}{f(x_n)^2}$. The above scheme has minimum fifth-order convergence and further reaches at six for $\alpha = 1$.

Moving ahead in this direction, Sharma [29], also constructed the following new modification of the above scheme (2):

$$\begin{cases} y_n = x_n - \frac{f(x_n)}{f'(x_n)}, \\ z_n = x_n - \left(1 + \frac{f(y_n)}{f(x_n) - 2\alpha f(y_n)}\right) \frac{f(x_n)}{f'(x_n)}, \quad \alpha \in \mathbb{R} \\ x_{n+1} = z_n - \frac{f(z_n)}{f[z_n, y_n] + f[z_n, y_n, x_n](z_n - y_n) + f[z_n, y_n, x_n, x_n](z_n - y_n)(z_n - x_n)}. \end{cases} \quad (3)$$

This family (3) has minimum sixth-order convergence and further attains eighth for $\alpha = 1$. So, it means that this scheme has an optimal eighth-order convergence but only for $\alpha = 1$.

Both of the above mentioned schemes namely, (2) and (3), using three values of the considered function and one derivative of first-order at per step. But, none of them achieved an optimal convergence for each α . No doubts, Sharma got a little success in this path but that one is valid only for $\alpha = 1$, not for other values. According to Kung-Traub conjecture, we can attain maximum eighth-order convergence by using the same functional evaluations.

While keep all these things in our mind, we intend to propose a new powerful and an optimal reconstruction of Chebyshev-Halley iteration functions of order four and eight. In addition, we analyze the local convergence of them by using suppositions requiring first-order derivative of the involved function f and the Lipschitz conditions. Moreover, we also present their theoretical radius of convergence which provides guaranteed convergence of them. Further, we will give a practical exhibition of our iteration functions to many real life situations and conclude that they perform better than the earlier studies. Finally, the dynamical behavior of them also illustrate the above consequences to a great extent.

Construction of higher-order optimal schemes

First of all in this section, we propose a new reconstruction of fourth-order Chebyshev-Halley methods, not requiring the computation of second-order derivative. Then, we extend the same scheme for eighth-order convergence. For this purpose, we consider the well-known second order Newton's method [30], which is given by

$$y_n = x_n - \frac{f(x_n)}{f'(x_n)}. \quad (4)$$

With the help of Taylor series, we will get the following expansions of functions $f(y_n)$ and $f'(y_n)$ about a point $x = x_n$, as follow

$$f(y_n) \approx \frac{(y_n - x_n)^2}{2} f''(x_n), \quad (5)$$

and

$$f'(y_n) \approx f'(x_n) + f''(x_n)(y_n - x_n). \quad (6)$$

By adding the expressions (5) and (6), we get

$$f(y_n) + f'(y_n) \approx \frac{(y_n - x_n)^2}{2} f''(x_n) + f'(x_n) + f''(x_n)(y_n - x_n), \quad (7)$$

which further yields

$$f''(x_n) \approx \frac{2[f'(x_n)]^2(f'(x_n) - f'(y_n) - f(y_n))}{f(x_n)(2f'(x_n) - f(x_n))}. \quad (8)$$

However, this new approximation for $f''(x_n)$ uses four functional evaluations, viz. $f(x_n), f'(x_n), f(y_n), f'(y_n)$. Therefore, in order to reduce the number of functional evaluations, we consider an approximation similar to the King's approximation [16], which is defined as follows

$$f'(y_n) = f'(x_n) \left(\frac{1 + \beta_1 v}{1 + \beta_2 v} \right), \quad (9)$$

where $v = \frac{f(y_n)}{f(x_n)}$ and $\beta_1, \beta_2 \in \mathbb{R}$.

Now, using the above expressions (8) and (9) in expression (1), we obtain a new reconstruction of Chebyshev-Halley family

$$x_{n+1} = x_n - \left[1 + \frac{1}{2} \frac{L_f^*(x_n)}{1 - \alpha L_f^*(x_n)} \right] \frac{f(x_n)}{f'(x_n)}, \quad \alpha \in \mathbb{R}, \quad (10)$$

where $L_f^*(x_n) = \frac{2f(y_n)[(\beta_1 - \beta_2)v^{-1} + \beta_2 v + 1]}{(f(x_n) - 2f'(x_n))(1 + \beta_2 v)}$.

In order to attain eighth-order convergence of the scheme (10), we rewrite in the following way:

$$\begin{aligned} y_n &= x_n - \frac{f(x_n)}{f'(x_n)}, \\ z_n &= x_n - \left[1 + \frac{1}{2} \frac{L_f^*(x_n)}{1 - \alpha L_f^*(x_n)} \right] \frac{f(x_n)}{f'(x_n)}, \quad \alpha \in \mathbb{R}, \\ x_{n+1} &= x_n - \frac{u(f(x_n) - f(z_n))f(y_n) + (x_n - z_n)(v - 1)f(x_n)f(z_n)}{(2v - 1)(x_n - z_n)f'(x_n)f(z_n) + (uf[x_n, z_n] - f(z_n))f(y_n)}, \end{aligned} \quad (11)$$

where $u = \frac{f(x_n)}{f'(x_n)}$ and $v = \frac{f(y_n)}{f'(x_n)}$. The next theorem 1 indicates that under what choices on disposable parameters in (10) and (11), the order of convergence will reach at four and eight, respectively without using any more functional evaluations.

Theorem 1 *Let $f : \mathbb{D} \subset \mathbb{R} \rightarrow \mathbb{R}$ has a simple zero ξ and is a sufficiently differentiable function in the closed interval \mathbb{D} containing ξ . We also assume that initial guess $x = x_0$ is sufficiently close to ξ . Then, the iterative schemes (10) and (11) have fourth and eighth-order convergence, respectively when*

$$\beta_1 = 2(\alpha - 2), \quad \beta_2 = 2(\alpha - 1), \quad (12)$$

where $\alpha \in \mathbb{R}$.

Proof: Let us assume that the error at nth iteration is $e_n = x_n - \xi$. We expand $f(x_n)$ and $f'(x_n)$ around $x = \xi$ with the help of Taylor's series expansion. Then, we have

$$f(x_n) = f'(\xi) \left[\sum_{j=1}^8 c_j e_n^j + O(e_n^9) \right], \quad (13)$$

and

$$f'(x_n) = f'(\xi) \left[\sum_{j=1}^8 j c_j e_n^j + O(e_n^9) \right], \quad (14)$$

where $c_n = \frac{1}{n!} \frac{f^{(n)}(\xi)}{f'(\xi)}$, $n = 2, 3, 4, \dots, 8$.

By using the equations (13) and (14), we get

$$\begin{aligned} u &= e_n - c_2 e_n^2 + 2(c_2^2 - c_3) e_n^3 + (7c_3 c_2 - 4c_2^3 - 3c_4) e_n^4 + (8c_2^4 - 20c_3 c_2^2 + 10c_4 c_2 + 6c_3^2 \\ &\quad - 4c_5) e_n^5 + (52c_3 c_2^3 - 16c_2^5 - 28c_4 c_2^2 + (13c_5 - 33c_3^2) c_2 + 17c_3 c_4 - 5c_6) e_n^6 + O(e_n^7). \end{aligned} \quad (15)$$

By using the expression (15), we obtain

$$\begin{aligned} y_n - \xi &= c_2 e_n^2 - 2(c_2^2 - c_3) e_n^3 + (4c_2^3 - 7c_3 c_2 + 3c_4) e_n^4 + (4c_5 - 8c_2^4 + 20c_3 c_2^2 - 10c_4 c_2 - 6c_3^2) e_n^5 \\ &\quad + (16c_2^5 - 52c_3 c_2^3 + 28c_4 c_2^2 + (33c_3^2 - 13c_5) c_2 - 17c_3 c_4 + 5c_6) e_n^6 - 2(16c_2^6 - 64c_3 c_2^4 \\ &\quad + 36c_4 c_2^3 + 9(7c_3^2 - 2c_5) c_2^2 + (8c_6 - 46c_3 c_4) c_2 - 9c_3^3 + 6c_4^2 + 11c_3 c_5 - 3c_7) e_n^7 + O(e_n^8). \end{aligned} \quad (16)$$

We have the following expansion of $f(y_n)$ about a point $x = \xi$

$$\begin{aligned} f(y_n) = f'(\xi) & \left[c_2 e_n^2 + (2c_3 - 2c_2^2) e_n^3 + (5c_2^3 - 7c_3 c_2 + 3c_4) e^4 - 2(6c_2^4 - 12c_3 c_2^2 + 5c_4 c_2 \right. \\ & + 3c_3^2 - 2c_5) e_n^5 + (16c_2^5 - 52c_3 c_2^3 + 28c_4 c_2^2 + (33c_3^2 - 13c_5) c_2 - 17c_3 c_4 + 5c_6) e_n^6 \\ & - 2(16c_2^6 - 64c_3 c_2^4 + 36c_4 c_2^3 + 9(7c_3^2 - 2c_5) c_2^2 - 9c_3^3 + (8c_6 - 46c_3 c_4) c_2 + 6c_4^2 \\ & \left. + 11c_3 c_5 - 3c_7) e_n^7 + O(e_n^8) \right]. \end{aligned} \quad (17)$$

With the help of expression (13) and (17), we get

$$\begin{aligned} v = c_2 e_n + (2c_3 - 3c_2^2) e_n^2 + (8c_2^3 - 10c_3 c_2 + 3c_4) e_n^3 + (37c_3 c_2^2 - 20c_2^4 - 14c_4 c_2 - 8c_3^2 + 4c_5) e_n^4 \\ + (48c_2^5 - 118c_3 c_2^3 + 51c_4 c_2^2 + (55c_3^2 - 18c_5) c_2 - 22c_3 c_4 + 5c_6) e_n^5 + (344c_3 c_2^4 - 112c_2^6 - 163c_4 c_2^3 \\ + (65c_5 - 252c_3^2) c_2^2 + 2(75c_3 c_4 - 11c_6) c_2 + 26c_3^3 - 15c_4^2 - 28c_3 c_5 + 6c_7) e_n^6 + O(e_n^7). \end{aligned} \quad (18)$$

Using equations (13) – (18), we have

$$\begin{aligned} L_f^*(x_n) = -(\beta_1 - \beta_2) c_2 e_n + \frac{1}{2} (2(\beta_1 - \beta_2)(\beta_2 + 3) c_2^2 + (-\beta_1 + \beta_2 - 2) c_2 + 4(\beta_2 - \beta_1) c_3) e_n^2 \\ + \sum_{r=0}^5 P_r e_n^{r+3} + O(e_n^9), \end{aligned} \quad (19)$$

where $P_r = P_r(\alpha, \beta_1, \beta_2, c_2, c_3, \dots, c_8)$.

By inserting expressions (13) – (19) in the scheme (10), we obtain

$$e_{n+1} = \frac{1}{2} (2 + \beta_1 - \beta_2) c_2 e_n^2 + \sum_{r=0}^5 \bar{P}_r e_n^{r+3}, \quad (20)$$

where $\bar{P}_r = \bar{P}_r(\alpha, \beta_1, \beta_2, c_2, c_3, \dots, c_8)$.

It is clear from the above expression (20) that we obtain at least third-order convergence, when we choose

$$\beta_2 = \beta_1 + 2. \quad (21)$$

Using the expression (21) in $P_0 = 0$, we obtain the following expression

$$(4 + \beta_1 - 2\alpha) c_2^2 = 0, \quad (22)$$

which further yields

$$\beta_1 = 2(\alpha - 2). \quad (23)$$

By substituting the expressions (21) and (23) in (10), we have

$$e_{n+1} = c_2 [(\alpha - 1) c_2 + c_2^2 - c_3] e_n^4 + \sum_{r=2}^5 \bar{P}_r e_n^{r+3} + O(e_n^9). \quad (24)$$

Again by using the expressions (13) – (19) and (21) – (23) in (11), we obtain

$$z_n - \xi = \bar{P}_1 e_n^4 + \bar{P}_2 e_n^5 + \bar{P}_3 e_n^6 + \bar{P}_4 e_n^7 + \bar{P}_5 e_n^8 + O(e_n^9), \quad (25)$$

where $\bar{P}_1 = c_2 [(\alpha - 1) c_2 + c_2^2 - c_3]$, $\bar{P}_2 = [2(3 - 4\alpha + \alpha^2) c_2^3 + \frac{1}{2} c_2^2 (\alpha + 16c_3 - 1) + c_2 \{4(\alpha - 1) c_3 -$

$2c_4\} - 4c_2^4 - 2c_3^2]$, $\bar{P}_3 = \frac{1}{4} \left[-8(7\alpha^2 - 19\alpha + 12)c_2^4 + 2c_2^3(2\alpha^2 - 9\alpha - 60c_3 + 7) + c_2^2 \{ \alpha + 16(3\alpha^2 - 11\alpha + 8)c_3 + 48c_4 - 1 \} + 4c_2 \{ 2(\alpha - 1)c_3 + 6(\alpha - 1)c_4 + 18c_3^2 - 3c_5 \} + 4c_3(4(\alpha - 1)c_3 - 7c_4) + 40c_2^5 \right]$, etc.

We can obtain the following Taylor series expansion from $f(z_n)$ about the point ξ with the help of expression (25)

$$f(z_n) = f'(\xi) \left[\bar{P}_1 e_n^4 + \bar{P}_2 e_n^5 + \bar{P}_3 e_n^6 + \bar{P}_4 e_n^7 + (\bar{P}_1^2 c_2 + \bar{P}_5) e_n^8 + O(e_n^9) \right]. \quad (26)$$

By using the equations (13) – (19) and (21) – (23) and (26), we have

$$\frac{u(f(x_n) - f(z_n))f(y_n) + (x_n - z_n)(v - 1)f(x_n)f(z_n)}{(2v - 1)(x_n - z_n)f'(x_n)f(z_n) + (uf[x_n, z_n] - f(z_n))f(y_n)} = e_n - c_2(c_2^3 - 2c_3c_2 + c_4)P_1e_n^8 + O(e_n^9), \quad (27)$$

Finally by substituting the above expression in scheme (11), we obtain

$$e_{n+1} = c_2^2 \left((\alpha - 1)c_2 + c_2^2 - c_3 \right) (c_2^3 - 2c_2c_3 + c_4)e_n^8 + O(e_n^9). \quad (28)$$

This above expressions (24) and (28) reveal that new constructions of Chebyshev-Halley methods (10) and (11) reach optimal fourth and eighth-order convergence, respectively. This completes the proof. \square

Local Convergence

The local convergence of method (10) was given using hypotheses up to the fourth derivative of function f although only the first derivative appears in this method. The local convergence of method (11) requires the usage of the eighth derivative. These hypotheses limit the applicability of both methods. As a motivational example, define function f on \mathbb{R} , $\mathbb{D} = [-\frac{5}{2}, \frac{1}{2}]$ by

$$f(x) = \begin{cases} x^3 \ln x^2 + x^5 - x^4, & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

Let us assume that our desired root is $\xi = 1$. Then, we have that

$$f'(x) = 3x^2 \ln x^2 + 5x^4 - 4x^3 + 2x^2, \quad f'(1) = 3,$$

$$f''(x) = 12x \ln x^2 + 20x^3 - 12x^2 + 10x$$

and

$$f'''(x) = 12 \ln x^2 + 60x^2 - 12x + 22.$$

Then, obviously, function $f'''(x)$ is unbounded on \mathbb{D} . Hence, the results in [21, 29], cannot apply to show the convergence of method (5) or its special cases requiring hypotheses on the third derivative of function F or higher. Notice that, in-particular there is a plethora of iterative methods for approximating solutions of nonlinear equations [1-32]. These results show that if the initial point x_0 is sufficiently close to the solution ξ , then the sequence $\{x_n\}$ converges to ξ . But how close to the solution ξ the initial guess x_0 should be? These local results give no information on the radius of the convergence ball for the corresponding method. The same technique can be used to other methods.

In this section, we present the local convergence analysis of method (5) using hypotheses only on the first derivative function f and Lipschitz constants. Similarly, we can study method

(14). We expand the applicability of these methods in this way. Moreover, we provide computable radius of convergence, error estimates on the distances $|x_n - \xi|$ and a uniqueness result. Let $L_0 > 0$, $L > 0$, $M \geq 1$ and $\alpha, \beta_1, \beta_2 \in \mathbb{R}$ be given constants. It is convenient for the local convergence analysis that follows to introduce some functions and parameters. Define functions g_1 , p and h_p on the interval $[0, \frac{1}{L_0})$, by

$$\begin{aligned} g_1(t) &= \frac{Lt}{2(1 - L_0 t)}, \\ p(t) &= \begin{cases} (L_0 + 2|\beta_2 - 2\alpha| Mg_1(t) + 2M)t + \frac{(|\alpha\beta_2| + |\alpha + \beta_2|)M^2 g_1(t)}{1 - \frac{L_0}{2}t}, & \text{if } 2|\beta_2 - 2\alpha| > |\beta_2| \\ (L_0 + 2(|\beta_2| + 2|\alpha|) Mg_1(t) + 2M)t + \frac{(|\alpha\beta_2| + |\alpha + \beta_2|)M^2 g_1(t)}{1 - \frac{L_0}{2}t}, & \text{if } 2|\beta_2 - 2\alpha| \leq |\beta_2| \end{cases} \\ h_p(t) &= p(t) - 1, \end{aligned}$$

and parameter r_1 by

$$r_1 = \frac{2}{2L_0 + L}.$$

Then, we have that $r_1 < \frac{1}{L_0}$ and for each $t \in [0, r_1)$, $0 \leq g_1(t) < 1$. Moreover, we have that $h_p(0) = -1$ and $h_p(t) \rightarrow +\infty$ as $t \rightarrow \frac{1}{L_0}^-$. It follows from the intermediate value theorem that function h_p has zeros in the interval $(0, \frac{1}{L_0})$. Denote by r_p the smallest such zero. Furthermore, we also define the following functions g_2 and h_2 on the interval $[0, r_p)$, by

$$g_2(t) = \frac{1}{2(1 - L_0 t)} \left[L + \frac{4M^3 g_1(t) (|\beta_1 - \beta_2| + |\beta_2| g_1(t)t + t)t}{(1 - \frac{L_0}{2}t)(1 - p(t))} \right] t,$$

and

$$h_2(t) = g_2(t) - 1.$$

We get that $h_2(0) = -1 < 0$ and $h_2(t) \rightarrow +\infty$ as $t \rightarrow r_p^-$. Let r_2 be the smallest zero of function h_2 on the interval $(0, r_p)$.

Set

$$r = \min\{r_1, r_2\}. \quad (29)$$

Then, we have that

$$0 < r \leq r_1. \quad (30)$$

and for each $t \in [0, r)$,

$$0 \leq g_1(t) < 1, \quad (31)$$

$$0 \leq p(t) < 1, \quad (32)$$

and

$$0 \leq g_2(t) < 1. \quad (33)$$

Let $U(\gamma, \delta)$, $\bar{U}(\gamma, \delta)$ denote, respectively for the open and closed balls in \mathbb{R} , with center $\gamma \in \mathbb{R}$, and of radius $\delta > 0$. Next, we present the local convergence analysis of method (4), (10) using the preceding notation.

Theorem 2 *Let $f : \mathbb{D} \subset \mathbb{R} \rightarrow \mathbb{R}$ be a differentiable function. Suppose that there exist $\xi \in \mathbb{D}$,*

$L_0 > 0, L > 0, M \geq 1, \alpha, \beta_1, \beta_2 \in \mathbb{R}$ such that for each $x, y \in \mathbb{D}$,

$$f(\xi) = 0, \quad f'(\xi)^{-1} \in L(\mathbb{R}, \mathbb{R}), \quad (34)$$

$$|f'(\xi)^{-1}(f'(x) - f'(\xi))| \leq L_0|x - \xi|. \quad (35)$$

Let us assume that $\Omega_0 = \mathbb{D} \cap U(\xi, \frac{1}{L_0})$.

$$|f'(\xi)^{-1}(f'(x) - f'(y))| \leq L|x - y|, \quad \text{foreach } x, y \in \Omega_0, \quad (36)$$

$$|f'(\xi)^{-1}f'(x)| \leq M, \quad \text{foreach } x \in \Omega_0, \quad (37)$$

and

$$\bar{U}(\xi, r) \subset \mathbb{D}, \quad (38)$$

hold, where the convergence radius r is defined by (29). Then, the sequence $\{x_n\}$ generated for $x_0 \in U(\xi, r) - \{\xi\}$ by method (5) is well defined, remains in $U(\xi, r)$ for each $n = 0, 1, 2, \dots$ and converges to ξ . Moreover, the following estimates hold

$$|y_n - \xi| \leq g_1(|x_n - \xi|)|x_n - \xi| \leq |x_n - \xi| < r, \quad (39)$$

and

$$|z_n - \xi| \leq g_2(|x_n - \xi|)|x_n - \xi| < |x_n - \xi|, \quad (40)$$

where the “ g ” functions are defined previously. Furthermore, for $T \in [r, \frac{2}{L_0})$, the limit point ξ is the only solution of equation $f(x) = 0$ in $\Omega_1 = \bar{U}(\xi, T) \cap \mathbb{D}$.

Proof: We shall show estimates (39)–(40) using mathematical induction. By hypotheses $x_0 \in U(\xi, r) - \{\xi\}$, (29) and (35), we get

$$|f'(\xi)^{-1}(f'(x_0) - f'(\xi))| \leq L_0|x_0 - \xi| < L_0r < 1. \quad (41)$$

It follows from the (41) and the Banach Lemma on invertible functions [3, 4, 26] that $f'(x_0) \neq 0$ and

$$|f'(x_0)^{-1}f'(\xi)| \leq \frac{1}{1 - L_0|x_0 - \xi|}. \quad (42)$$

Hence, y_0 is well defined by the first sub-step of the method (5) for $n = 0$. Then, we have by (29), (31), (36) and (42) that

$$\begin{aligned} |y_0 - \xi| &= |x_0 - \xi - f'(x_0)^{-1}f(x_0)| \\ &\leq |f'(x_0)^{-1}f'(\xi)| \int_0^1 |f'(\xi)^{-1}[f'(\xi + \theta(x_0 - \xi)) - f'(x_0)](x_0 - \xi)d\theta| \\ &\leq \frac{L|x_0 - \xi|^2}{2(1 - L_0|x_0 - \xi|)} \\ &= g_1(|x_0 - \xi|)|x_0 - \xi| < |x_0 - \xi| < r, \end{aligned} \quad (43)$$

which shows (39) for $n = 0$ and $y_0 \in U(\xi, r)$.

The fraction in (10) can be written for $n = 0$ as $\frac{N_0}{D_0}$, where $N_0 = 2f(y_0)[(\beta_1 - \beta_2)f'(x_0) + \beta f(y_0) + f(x_0)]$ and $D_0 = \frac{1}{2}f(x_0)D_1$, $D_1 = 2f(x_0) + 2(\beta_2 - 2\alpha)f(y_0) - \frac{(\alpha + \beta_2)f(y_0)f'(x_0)}{f(x_0)} - \alpha\beta_2 \frac{[f(y_0)]^2}{f(x_0)} - f'(x_0)$.

We need to show that $f(x_0) \neq 0$, $f(x_0) - 2f'(x_0) \neq 0$, $f(x_0) + \beta_2 f(y_0) \neq 0$ and $D_1 \neq 0$ for $x_0 \neq \xi$. Using (29), (34) and (35), we have that

$$|f'(\xi)(x_0 - \xi)|^{-1} (f(x_0) - f(\xi) - f'(\xi)(x_0 - \xi)) \leq |x_0 - \xi|^{-1} \frac{L_0}{2} |x_0 - \xi|^2 = \frac{L_0}{2} |x_0 - \xi| < 1. \quad (44)$$

It follows from (44) that $f(x_0) \neq 0$ and

$$|f'(x_0)^{-1} f'(\xi)| \leq \frac{1}{|x_0 - \xi| \left(1 - \frac{L_0}{2} |x_0 - \xi|\right)}. \quad (45)$$

Then, by (29), (32), (35), (37), (43) and (45), we obtain for $2|\beta_2 - 2\alpha| > |\beta_2|$ and using the first version of the function p in turn that

$$\begin{aligned} & \left| f'(\xi)^{-1} \left[f'(x_0) - f'(\xi) + \frac{\alpha\beta_2[f(y_0)]^2}{f(x_0)} + \frac{(\alpha + \beta_2)f(y_0)f'(x_0)}{f(x_0)} - 2f(x_0) - 2(\beta_2 - 2\alpha)f(y_0) \right] \right| \\ & \leq L_0|x_0 - \xi| + \frac{|\alpha\beta_2|M^2|y_0 - \xi|}{|x_0 - \xi|(1 - \frac{L_0}{2}|x_0 - \xi|)} + \frac{|\alpha + \beta_2|M^2|y_0 - \xi|}{|x_0 - \xi|(1 - \frac{L_0}{2}|x_0 - \xi|)} \\ & \quad + 2M|x_0 - \xi| + 2|\beta_2 - 2\alpha|M|y_0 - \xi| \\ & \leq L_0|x_0 - \xi| + \frac{|\alpha\beta_2|M^2g_1(|x_0 - \xi|)}{1 - \frac{L_0}{2}|x_0 - \xi|} + \frac{|\alpha + \beta_2|M^2g_1(|x_0 - \xi|)}{1 - \frac{L_0}{2}|x_0 - \xi|} \\ & \quad + 2[M + |\beta_2 - 2\alpha|Mg_1(|x_0 - \xi|)]|x_0 - \xi| \\ & = p(|x_0 - \xi|) < p(r) < 1. \end{aligned} \quad (46)$$

Hence, we get that

$$|D_0^{-1} f'(\xi)| \leq \frac{1}{1 - p(|x_0 - \xi|)}. \quad (47)$$

If we use $2|\beta_2 - 2\alpha| \leq |\beta_2|$ then the term $2|\beta_2 - 2\alpha|$ can be replaced by $2(|\beta_2| + 2|\alpha|)$. For this condition, we use the second version of the function p . Further, by using (45) and (47) we have that

$$|D_1^{-1} f'(\xi)| \leq \frac{2}{|x_0 - \xi|(1 - \frac{L_0}{2}|x_0 - \xi|)(1 - p(|x_0 - \xi|))}. \quad (48)$$

Hence, x_1 is well defined by (10) for $n = 0$. We also notice from equation (46) which implies that $f(x_0) - 2f'(x_0) \neq 0$, $f(x_0) + \beta_2 f(y_0) \neq 0$. Further, we have

$$|f'(\xi)^{-1} (f'(x_0) - f'(\xi)) - \frac{f(x_0)}{2}| \leq L_0|x_0 - \xi| + \frac{M|x_0 - \xi|}{2} \leq p(|x_0 - \xi|) < p(r) < 1, \quad (49)$$

and

$$\begin{aligned} & |f'(\xi)(x_0 - \xi)|^{-1} ((f(x_0) - f(\xi) - f'(\xi)(x_0 - \xi)) + \beta_2 f(y_0)) \\ & \leq |x_0 - \xi|^{-1} \left(\frac{L_0}{2} |x_0 - \xi|^2 + |\beta_2|M|y_0 - \xi| \right) \\ & \leq \frac{L_0}{2} |x_0 - \xi| + M|\beta_2|g_1(|x_0 - \xi|) \\ & \leq p(|x_0 - \xi|) < p(r) < 1, \end{aligned} \quad (50)$$

for p given by the first formula, if $2|\beta_2 - 2\alpha| > |\beta_2|$ and from the second formula if $2|\beta_2 - 2\alpha| \leq |\beta_2|$.

Now, we need to estimate the following

$$\begin{aligned} |N_0| &\leq 2|f'(\xi)^{-1}f(y_0)|[|\beta_1 - \beta_2| + |\beta_2||f'(\xi)^{-1}f(y_0)| + |f'(\xi)^{-1}f(x_0)|] \\ &\leq 2M|y_0 - \xi|[|\beta_1 - \beta_2| + |\beta_2|M|y_0 - \xi| + M|x_0 - \xi|] \\ &\leq 2Mg_1(|x_0 - \xi|)|x_0 - \xi|[|\beta_1 - \beta_2| + |\beta_2|Mg_1(|x_0 - \xi|)|x_0 - \xi| + M|x_0 - \xi|]. \end{aligned} \quad (51)$$

Therefore, by (10) (for $n = 0$), (30), (33), (37), (42), (43), (45), (47), (48), and (51) we get in turn that

$$\begin{aligned} |x_1 - \xi| &\leq |x_0 - \xi - f'(x_0)^{-1}f(x_0)| + \frac{2M^3|y_0 - \xi|(|\beta_1 - \beta_2| + |\beta_2||y_0 - \xi| + |x_0 - \xi|)|x_0 - \xi|}{(1 - L_0|x_0 - \xi|)(1 - \frac{L_0}{2}|x_0 - \xi|)(1 - p(|x_0 - \xi|))|x_0 - \xi|} \\ &\leq \frac{2M^3g_1(|x_0 - \xi|)[|\beta_1 - \beta_2| + |\beta_2|g_1(|x_0 - \xi|)|x_0 - \xi| + |x_0 - \xi|]|x_0 - \xi|^3}{(1 - L_0|x_0 - \xi|)(1 - \frac{L_0}{2}|x_0 - \xi|)(1 - p(|x_0 - \xi|))} \\ &\quad + \frac{L|x_0 - \xi|^2}{2(1 - L_0|x_0 - \xi|)} \\ &= g_2(|x_0 - \xi|)|x_0 - \xi| < |x_0 - \xi| < r, \end{aligned} \quad (52)$$

which shows (40) for $n = 0$ and $x_1 \in U(\xi, r)$. By simply replacing x_0, y_0 by x_k, y_k in the preceding estimates we arrive at (39)–(40). Using the estimates $\|x_{k+1} - \xi\| < \|x_k - \xi\| < r$, we deduce that $\lim_{k \rightarrow \infty} x_k = \xi$ and $x_{k+1} \in \Omega_1$. Finally, to show the uniqueness part, let $Q = \int_0^1 f'(y^* + \theta(\xi - y^*))d\theta$ for some $y^* \in \Omega_1$ with $f(y^*) = 0$. Using (35), we get that

$$\begin{aligned} \|f'(\xi)^{-1}(Q - f'(\xi))\| &\leq \|\int_0^1 L_0|y^* + \theta(\xi - y^*) - \xi|d\theta\| \\ &\leq \int_0^1 (1 - t)\|y^* - \xi\|d\theta \leq \frac{L_0}{2}T < 1. \end{aligned} \quad (53)$$

It follows from (53) that Q is invertible. Then, in view of the identity $0 = f(\xi) - f(y^*) = Q(\xi - y^*)$, we conclude that $\xi = y^*$. \square

Remarks

(a) In view of (35) and the estimate

$$\begin{aligned} |f'(\xi)^{-1}f'(x)| &= |f'(\xi)^{-1}(f'(x) - f'(\xi)) + I| \\ &\leq 1 + |f'(\xi)^{-1}(f'(x) - f'(\xi))| \\ &\leq 1 + L_0|x_0 - \xi| \end{aligned}$$

condition (37) can be dropped and M can be replaced by

$$M(t) = 1 + L_0t$$

or by $M(t) = M = 2$, since $t \in [0, \frac{1}{L_0})$.

(b) The results obtained here can be used for operators f satisfying the autonomous differential equation [3, 4] of the form

$$f'(x) = P(f(x)),$$

where P is a known continuous operator. Since $f'(\xi) = P(f(\xi)) = P(0)$, we can apply the results without actually knowing the solution ξ . Let as an example $f(x) = e^x - 1$. Then, we can choose $P(x) = x + 1$.

- (c) The radius r_1 was shown in [3, 4] to be the convergence radius for Newton's method under conditions (35) and (36). It follows from (31) and the definition of r_1 that the convergence radius r of the method (5) cannot be larger than the convergence radius r_1 of the second order Newton's method. As already noted in r_1 is at least as the convergence ball give by Rheinboldt [28]

$$r_R = \frac{2}{3L}.$$

In particular, for $L_0 < L$ we have that

$$r_R < r_1$$

and

$$\frac{r_R}{r_1} \rightarrow \frac{1}{3} \quad \text{as} \quad \frac{L_0}{L} \rightarrow 0.$$

That is our convergence ball r_1 is at most three times larger than Rheinboldt's. The same value for r_R given by Traub [30].

Numerical experiments

In this section, we want to assert that our proposed an optimal family of Chebyshev Halley methods perform better than the families of Chebyshev-Halley methods proposed by Li et al. (2014) and Sharma (2015). Some of the researchers who want to assert that their methods are superior than the other existing methods available in the literature. Generally, they consider either some well-known or standard or self made test problem and then mold the initial approximation to assert that their methods are superior than other methods. Molding the initial guess mean, let A researcher who wants to compare his/her method/methods with B's method/methods by considering a particular test problem with x_0 as initial guess. Now, if A's method/methods does/do not perform better than B's method/methods then A changes the initial guess and continue this process until he/she gets better results than B's method/methods. If A does not get success on that particular test problem on any initial guess then A consider another test problem and continue the same process until A does not get success.

To halt this practice, we consider total six numerical examples out of them first two are chosen from Li et al. [21], third and fourth from Sharma [29] with same initial guesses which are mentioned in their papers. Further, fifth and six test examples are taken from Petković et al. [25]. The details of chosen test problems or functions are available in the Table 1. Further, the initial approximations and zeros of the corresponding test functions are also display in the same table.

To check the effectiveness and validity of the theoretical results, we employ the new optimal family of Chebyshev-Halley methods (11) ($MCHM$), with Chebyshev's method (MCM) ($\alpha = 0$), Halley's method (MHM) ($\alpha = \frac{1}{2}$) and super-Halley method ($MSHM$) ($\alpha = 1$). We shall compare our schemes with a family of Chebyshev-Halley type methods that is very recently proposed by Li et al. [21], out of them we shall pick their best methods (which are claimed by them not by us) namely, Chebyshev's method (LCM) ($\alpha = 0$), Halley's method (LHM) ($\alpha = \frac{1}{2}$) and super-Halley method ($LSHM$) ($\alpha = 1$). Finally, we shall also compare our schemes with the improved Chebyshev-Halley methods which is developed by Sharma [29], between them

we shall choose their best methods namely, ($\alpha = 0$, $\alpha = \frac{1}{2}$, $\alpha = 1$) denoted by *SCM*, *SHM*, and *SSHM*, respectively.

In the Table 2, we display the minimum number of iterations are required to get the desire accuracy to the corresponding zeros of the functions $f_1(x) - f_6(x)$ which are given in Table 1. In addition, we also exhibit the absolute errors $|x_{n+1} - x_n|$ for first three consecutive approximations in this table. Further, the meaning of $(Ae - h)$ is $(A \times 10^{-h})$. Furthermore, we also want to demonstrate the theoretical order of convergence which is proved in section 3. Therefore, to calculate the computational order of convergence, we use the following formula proposed by [31], which is defined as follows

$$\rho \approx \frac{\ln |(x_{n+1} - \xi)/(x_n - \xi)|}{\ln |(x_n - \xi)/(x_{n-1} - \xi)|}.$$

But, this COC requires the exact root ξ and there are many practical situations where the exact root is not known in advance. To overcome this problem, Grau-Sánchez et al. [13], given another definition of COC, which is defined as follows

$$\rho \approx \frac{\ln |\check{e}_{n+1}/\check{e}_n|}{\ln |\check{e}_n/\check{e}_{n-1}|}, \quad (54)$$

where $\check{e}_n = x_n - x_{n-1}$.

All computations have been performed by using the programming package *Mathematica 9* with multiple precision arithmetic. We use $\epsilon = 10^{-300}$ as a tolerance error. The following stopping criteria are chosen for computer programs:

(i) $|x_{n+1} - x_n| < \epsilon$ and (ii) $|f(x_{n+1})| < \epsilon$.

It is noteworthy from the table 2, that our proposed schemes perform better than the Li et al. (2014) and Sharma (2015), when the accuracy is tested in the high precision digits. For better comparison, we give a column by column comparison of different modifications of Chebyshev–Halley methods, so that we can easily see the exact difference between the proposed modifications and existing modifications of Chebyshev–Halley methods. Further, the accuracy in numerical values of approximations to the root by the proposed scheme is higher than the recently improvement of Chebyshev–Halley methods given by Li et al. [21] and Sharma [29]. In general, the our optimal family of Chebyshev–Halley methods (*MCHM*) is superior among all the other proposed methods. This superiority is in accordance because it is an optimal modification of Chebyshev–Halley methods according to Kung–Traub conjecture [20], which is discussed in the previous section. The computational order of convergence (COC) and dynamic study of these methods also confirmed the above conclusions to a great extent.

Now, we also demonstrate the theoretical results which we proposed in section 4, by the applying on some other numerical examples, which are defined as follows:

Example 1 Let f be a function defined on $\mathbb{D} = \bar{U}(0, 1)$, which is given as follows

$$f_7(x) = e^x - 1. \quad (55)$$

Then the first-derivative is $f'_7(x) = e^x$. We get that $L_0 = e - 1 < L = e^{\frac{1}{e-1}}$, $\alpha = 1$ $M = 2$, $\beta_1 = 2(\alpha - 2)$ and $\beta_2 = 2(\alpha - 1)$. By substituting different values of parameters, we get different radius of convergence which are display in the Table 3.

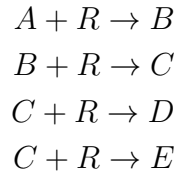
Table 1: Test problems

$f(x)$	Initial guess	Root(ξ)
$f_1(x) = 10xe^{-x^2} - 1$; (see [21])	1.7	1.67963061042845...
$f_2(x) = (x + 2)e^x - 1$; (see [21])	-0.5	-0.442854401002389...
$f_3(x) = e^x + 2^{-x} + 2 \cos x - 6$; (see [29])	3.5	1.82938360193385...
$f_4(x) = (x - 2)(x^{10} + x + 1)e^{-x-x}$; (see [29])	2.5	2
$f_5(x) = e^{x^2+7x-30} - 1$; (see [25])	3.3	3
$f_6(x) = (x - 2)^2 - \log x - 33x$; (see [25])	33	36.9894735829447...

Example 2 Returning back to the motivation example at the introduction on this section, we have $L = L_0 = 146.6629073$, $M = 2$, $\beta_1 = 2(\alpha - 2)$ and $\beta_2 = 2(\alpha - 1)$. By substituting different values of parameters, we get different radius of convergence which are display in the Table 4.

Example 3 Continuous stirred tank reactor (CSTR)

Let us consider the isothermal continuous stirred tank reactor (CSTR). Components A and R are fed to the reactor at rates of Q and $q - Q$, respectively. Then, we obtain the following reaction scheme in the reactor (for the details see [10]):



The problem was analyzed by Douglas [12] in order to design simple feedback control systems. He presented the following expression for the transfer function of the reactor

$$K_C \frac{2.98(x + 2.25)}{(x + 1.45)(x + 2.85)^2(x + 4.35)} = -1,$$

where K_C is the gain of the proportional controller. The control system is stable for values of K_C that yields roots of the transfer function having negative real part. If we choose $K_C = 0$ we get the poles of the open-loop transfer function as roots of the nonlinear equation:

$$f_8(x) = x^4 + 11.50x^3 + 47.49x^2 + 83.06325x + 51.23266875. \quad (56)$$

No doubts, the above function f_1 has four zeros $x^* = -1.45, -2.85, -2.85, -4.35$. However, our required zero is $x^* = -4.35$ for expression (57). Let us also consider $\mathbb{D} = [-4.5, -4]$.

Then, we obtain

$$L_0 = L = 2.760568793, \quad M = 2$$

Now, with the help of different values, we get different radius of convergence displayed in Table 5.

Table 2: (Comparison of different multi-point methods)

$f(x)$		LCM	LHM	$LSHM$	SCM	SHM	$SSHM$	MCM	MHM	$MSHM$
$f_1(x)$	$ x_2 - x_1 $	1.4e-1	7.6e-1	5.9e-1	2.5e-1	1.9e-1	8.1e-1	6.3e-3	3.8e-3	3.2e-2
	$ x_3 - x_2 $	4.2e-5	4.5e-2	4.2e-3	1.3e-4	1.3e-5	1.4e-10	8.1e-20	1.2e-22	2.9e-14
	$ x_4 - x_3 $	5.7e-23	4.2e-8	4.2e-6	8.8e-24	2.3e-30	1.4e-80	5.7e-155	1.6e-178	1.1e-110
	n	6	7	6	6	6	5	5	5	5
	ρ	5.000	5.000	6.000	6.000	6.000	8.000	8.000	8.000	8.000
$f_2(x)$	$ x_2 - x_1 $	1.5e-2	1.1e+0	2.1e-1	1.1e-1	9.0e-2	4.8e-2	2.3e-2	2.3e-2	2.6e-2
	$ x_3 - x_2 $	2.5e-7	2.0e-1	1.4e-2	8.0e-4	1.7e-5	9.0e-8	9.3e-11	1.2e-10	2.9e-10
	$ x_4 - x_3 $	2.0e-31	1.1e+0	1.6e-9	1.1e-15	2.2e-20	2.7e-53	7.0e-78	7.3e-77	1.1e-73
	n	96	6	6	6	6	5	5	5	5
	ρ	5.000	5.000	6.000	6.000	6.000	8.000	8.000	8.000	8.000
$f_3(x)$	$ x_2 - x_1 $	8.9e-10	3.7e-14	5.4e-12	2.7e-10	6.4e-11	6.2e-15	8.8e-15	6.3e-15	3.9e-15
	$ x_3 - x_2 $	7.3e-47	4.3e-49	1.4e-69	1.3e-57	5.6e-62	4.0e-115	9.2e-114	4.9e-115	7.2e-117
	$ x_4 - x_3 $	2.7e-232	9.0e-244	3.9e-415	1.3e-341	2.5e-368	1.1e-916	1.4e-905	1.4e-916	8.8e-931
	n	5	5	4	4	4	4	4	4	4
	ρ	5.000	5.000	6.000	6.000	6.000	8.000	8.000	8.000	8.000
$f_4(x)$	$ x_2 - x_1 $	2.0e-7	8.7e-8	1.2e-9	2.7e-8	6.1e-9	2.4e-12	4.4e-13	1.4e-13	1.5e-13
	$ x_3 - x_2 $	8.3e-35	7.2e-37	1.4e-55	2.7e-46	8.8e-51	2.3e-95	5.1e-102	1.3e-106	2.8e-106
	$ x_4 - x_3 $	1.1e-171	2.7e-182	2.8e-331	2.5e-274	7.4e-302	1.6e-759	1.6e-813	9.0e-851	4.4e-848
	n	5	5	4	5	4	4	4	4	4
	ρ	5.000	5.000	6.000	6.000	6.000	8.000	8.000	8.000	8.000
$f_5(x)$	$ x_2 - x_1 $	D	1.3e-2	1.1e+0	1.2e-1	1.2e-1	1.0e+0	3.6e-2	3.6e-2	3.7e-2
	$ x_3 - x_2 $	D	9.2e-3	2.9e-1	1.6e-2	8.7e-3	5.7e-4	1.9e-9	2.1e-9	3.1e-9
	$ x_4 - x_3 $	D	5.9e-3	2.3e-6	4.8e-7	4.8e-9	1.4e-21	5.6e-68	1.6e-67	5.2e-66
	n	D	D	7	7	7	6	5	5	5
	ρ	NC	5.000	6.000	6.000	6.000	8.000	8.000	8.000	8.000
$f_6(x)$	$ x_2 - x_1 $	7.2e-4	1.2e-4	1.7e-6	7.2e-4	1.2e-4	1.7e-6	2.4e-6	9.3e-6	1.7e-6
	$ x_3 - x_2 $	5.7e-26	2.8e-30	9.4e-47	7.8e-27	3.6e-32	7.7e-58	3.8e-47	1.1e-50	7.7e-58
	$ x_4 - x_3 $	1.5e-136	2.3e-158	2.4e-288	1.3e-164	3.3e-197	1.4e-468	1.8e-381	3.1e-410	1.4e-468
	n	5	5	5	5	5	4	4	4	4
	ρ	5.000	5.000	6.000	6.000	6.000	8.000	8.000	8.000	8.000

(D: stands for divergence. NC means no need to calculate.)

Table 3: Behavior of scheme (10) on example (1)

α	r_1	r_2	$r = \min\{r_1, r_2\}$	x_0	$ x_2 - x_1 $	$ x_3 - x_2 $	$ x_4 - x_3 $	ρ
$\alpha = 0$	0.382692	0.066575	0.066575	0.65	$3.4\text{e}(-6)$	$2.8\text{e}(-23)$	$1.3\text{e}(-91)$	4.000
$\alpha = \frac{1}{2}$	0.382692	0.0877379	0.0877379	0.85	$4.1\text{e}(-6)$	$2.3\text{e}(-23)$	$2.1\text{e}(-92)$	4.000
$\alpha = 1$	0.382692	0.0877468	0.0877468	0.86	$2.2\text{e}(-6)$	$1.0\text{e}(-24)$	$4.2\text{e}(-98)$	4.000

Table 4: Behavior of scheme (10) on example (2)

α	r_1	r_2	$r = \min\{r_1, r_2\}$	x_0	$ x_2 - x_1 $	$ x_3 - x_2 $	$ x_4 - x_3 $	ρ
$\alpha = 0$	0.00454557	0.00109685	0.00109685	1.0009	$5.5\text{e}(-12)$	$7.4\text{e}(-45)$	$2.5\text{e}(-176)$	4.000
$\alpha = \frac{1}{2}$	0.00454557	0.00167861	0.00167861	1.0011	$1.9\text{e}(-11)$	$1.6\text{e}(-42)$	$7.6\text{e}(-167)$	4.000
$\alpha = 1$	0.00454557	0.00167861	0.00167861	1.0011	$2.5\text{e}(-11)$	$7.0\text{e}(-42)$	$4.1\text{e}(-164)$	4.000

Example 4 In the study of the multi-factor effect, the trajectory of an electron in the air gap between two parallel plates is given by

$$x(t) = x_0 + \left(v_0 + e \frac{E_0}{m\omega} \sin(\omega t_0 + \alpha) \right) (t - t_0) + e \frac{E_0}{m\omega^2} \left(\cos(\omega t + \alpha) + \sin(\omega + \alpha) \right), \quad (57)$$

where e and m are the charge and the mass of the electron at rest, x_0 and v_0 are the position and velocity of the electron at time t_0 and $E_0 \sin(\omega t + \alpha)$ is the RF electric field between the plates. We choose the particulars parameters in the expression (57) in order to deal with a simpler expression, which is defined as follows:

$$f_9(x) = x - \frac{1}{2} \cos(x) + \frac{\pi}{4}. \quad (58)$$

The required zero of the above function $\alpha = -0.309093271541794952741986808924$.

Then, we have

$$L_0 = L = M = 1.523542095.$$

So, we obtain the different radius of convergence which are displayed in Table 6 by using the above values.

Attractor basins in the complex plane

In this section, we present the dynamics of the proposed method based on visual display of their basins of attraction when $f(x)$ is a given fixed complex polynomial $q(z)$. We further

Table 5: Behavior of scheme (10) on example (3)

α	r_1	r_2	$r = \min\{r_1, r_2\}$	x_0	$ x_2 - x_1 $	$ x_3 - x_2 $	$ x_4 - x_3 $	ρ
$\alpha = 0$	0.241496	0.0473699	0.0473699	-4.396	$2.1\text{e}(-5)$	$1.1\text{e}(-18)$	$8.4\text{e}(-72)$	4.000
$\alpha = \frac{1}{2}$	0.241496	0.0650116	0.0650116	-4.41	$4.5\text{e}(-5)$	$1.9\text{e}(-17)$	$5.9\text{e}(-67)$	4.000
$\alpha = 1$	0.241496	0.0650156	0.0650156	-4.41	$3.2\text{e}(-5)$	$3.6\text{e}(-18)$	$5.3\text{e}(-70)$	4.000

Table 6: Behavior of scheme (10) on example (4)

α	r_1	r_2	$r = \min\{r_1, r_2\}$	x_0	$ x_2 - x_1 $	$ x_3 - x_2 $	$ x_4 - x_3 $	ρ
$\alpha = 0$	0.437577	0.066575	0.066575	-0.374	1.2e(-6)	1.5e(-25)	3.1e(-101)	4.000
$\alpha = \frac{1}{2}$	0.437577	0.124466	0.124466	-0.419	4.1e(-6)	7.5e(-24)	8.3e(-95)	4.000
$\alpha = 1$	0.382692	0.124493	0.124493	-0.419	1.9e(-6)	1.8e(-25)	1.4e(-101)	4.000

investigate some dynamical properties of the attained simple root finders in the complex plane by analyzing the structure of their basins of attraction. It is known that the corresponding fractal of an iterative root-finding method is a boundary set in the complex plane, which is characterized by the iterative method applied to a fixed polynomial $q(z) \in \mathbb{C}$, see e.g. [27, 24, 7, 6].

The aim herein is to use basin of attraction as another way for comparing the iterative methods. Therefore, we here investigate the dynamics of the listed simple root finders in the complex plane using basins of attraction which gives important information about convergence and stability of the method. To start with, let us recall some basic concepts which are related to basins of attractions. To start with, let us recall some basic concepts which are related to basins of attractions. Let $g : \mathbb{C} \rightarrow \mathbb{C}$ be a rational map on the Riemann sphere. The orbit of a point $z \in \mathbb{C}$ under g is defined

$$\{z, g(z), g^2(z), \dots, g^n(z), \dots\},$$

which consists of successive images of z by the rational map g . The dynamic behavior of the orbit of a point of g would be characterize by its asymptotic behavior. We first introduce some notions of a point in the orbit under g : a point $z_0 \in \mathbb{C}$ is known as a *fixed point* of g , if $g(z_0) = z_0$. In addition, z_0 is known as a *periodic point of period* $m > 1$, if $g^m(z_0) = z_0$, where m is smallest such integer. Further, if z_0 is a periodic point of period m then it is a fixed point for g^m . Moreover, there are mainly four types of fixed points of a map g , which are based on the magnitude of the derivative. A fixed point z_0 is known as:

If ξ is a root of $f(x)$, then the basin of attraction of ξ , is the collection of those initial approximations x_0 which converge to ξ . It is mathematically defined as follows:

$$B(\xi) = \{z_0 \in \mathbb{C} : g^n(z_0) \rightarrow \xi \text{ as } n \rightarrow \infty\}.$$

Arthur Cayley was the first person who considered the concept of the basins of attraction for Newton's method in 1879. Initially, he considered this concept for the quadratic polynomial. After some time, he also considered cubic polynomials, but was unable to find an obvious division for the basins of attraction as he earlier defined for the quadratic equations. In the early of 20th century, the French mathematicians Gaston Julia and Pierre Fatou started to understand the nature of complex cubic polynomials. The Julia set of a nonlinear map $g(z)$, called $J(g)$, is the closure of the set of its repelling fixed points and establishes the borders between the basins of attraction. On the other hand, the complement of $J(g)$ is known as the Fatou set $F(g)$. In simple words, the basins of attraction of any fixed point belongs to the Fatou set $F(g)$ and the boundaries of these basins of attraction belong to the Julia set $J(g)$. For the details of these concepts please see [11, 27, 24]. The aim herein is to use the basins of attraction as another way for characterizing initial approximations converging to the desired root ξ for the listed iteration algorithms. That is to say, the basins of attraction play a role representing a valuable dynamics

of the iteration schemes under consideration.

In order to achieve a vivid description from a dynamical point of view, we consider a rectangle $\mathbb{D} = [-3, 3] \times [-3, 3] \in \mathbb{C}$ with a 400×400 grid, and we assign a color to each point $z_0 \in D$ according to the simple root at which the corresponding iterative method starting from z_0 converges, and we mark the point as black if the method does not converge. In this section, we consider the stopping criterion for convergence to be less than 10^{-4} wherein the maximum number of full cycles for each method is considered to be 200. In this way, we distinguish the attraction basins by their colors for different methods. For concrete examples of dynamics of the listed methods behind the basins of attraction, we present several test problems described below.

Test problem 1. Let $p_1(z) = (z^4+1)$, having simple zeros $\{-0.707107-0.707107i, -0.707107+0.707107i, 0.707107-0.707107i, 0.707107+0.707107i\}$. It is straight forward to see from Fig. 1 – 3 that our methods, namely *MCM*, *MHM* and *MSHM* are the best methods in terms of less chaotic behavior to obtain the solutions. Further, our methods also have the largest basins for the solution and is faster in comparison to all the mentioned methods.

Test problem 2. Let $p_2(z) = (z^3+2z)$, having simple zeros $\{0, -1.41421i, 1.41421i\}$. Based on Fig. 4 – 6 , it is observe that our proposed methods namely, *MCM*, *MHM* and *MSHM* are the best methods because they have larger and brighter basin of attraction in comparison to the methods namely, *LCM*, *LHM*, *LSHM*, *SCM*, *SHM* and *SSHM*, respectively.

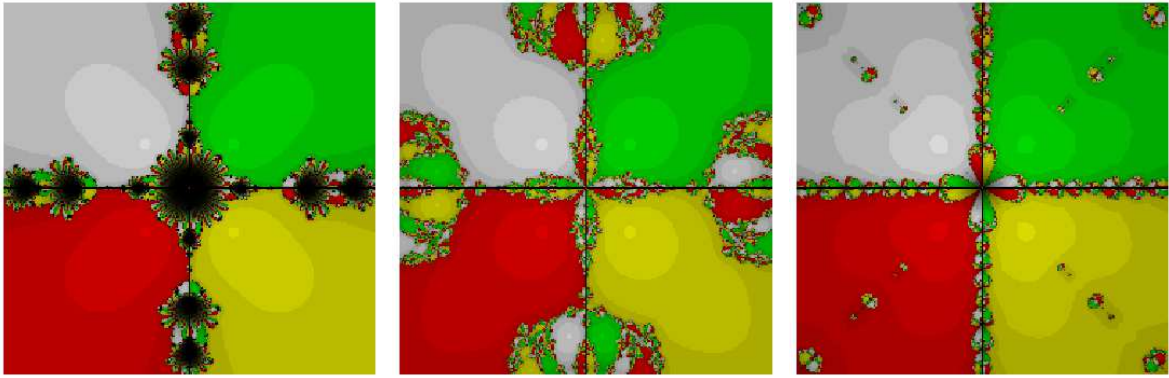


Figure 1: The methods *LCM*, *LHM* and *LSHM*, respectively for test problem 1.

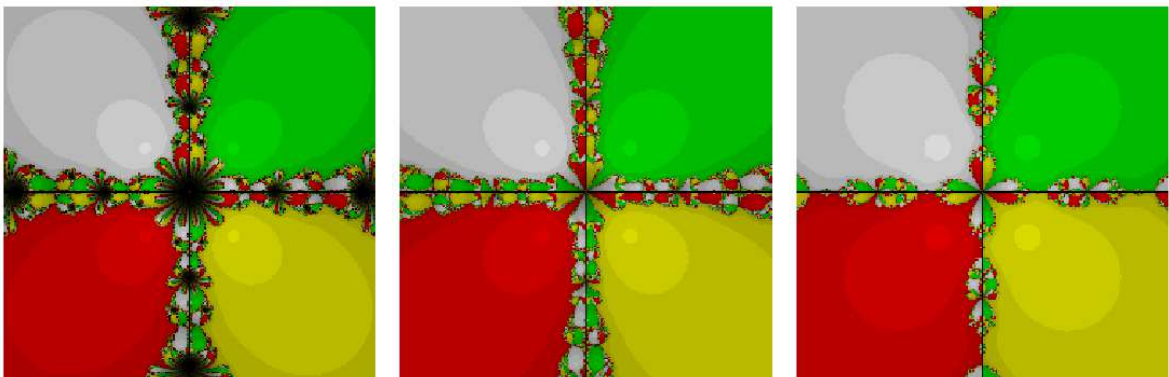


Figure 2: The methods *SCM*, *SHM* and *SSHM*, respectively for test problem 1.

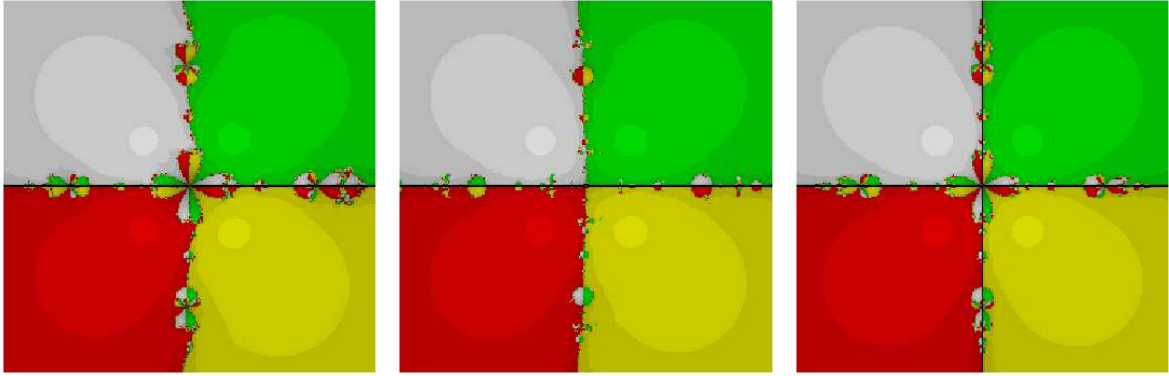


Figure 3: The methods *MCM*, *MHM* and *MSHM*, respectively for test problem 1.

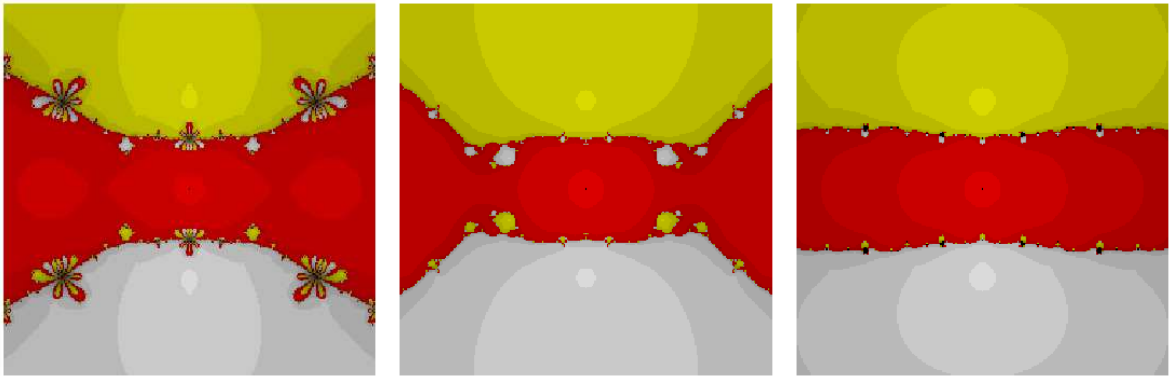


Figure 4: The methods *LCM*, *LHM* and *LSHM*, respectively for test problem 2.

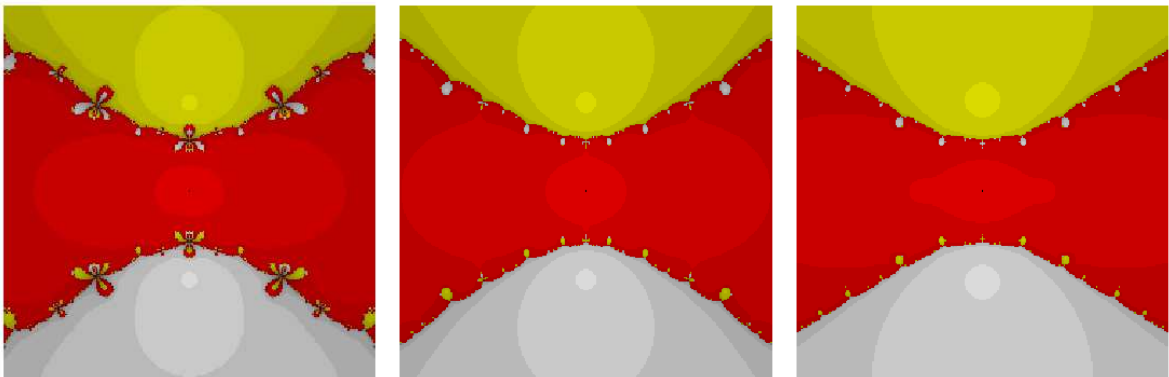


Figure 5: The methods *SCM*, *SHM* and *SSHM*, respectively for test problem 2.

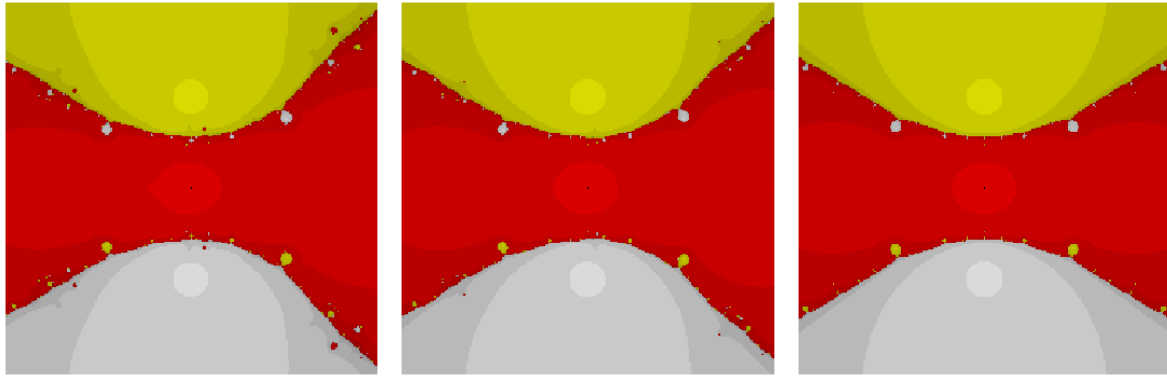


Figure 6: The methods MCM , MHM and $MSHM$, respectively for test problem 2.

Conclusions

The present contribution of this study is not only to increase the order of convergence of classical Chebyshev-Halley method. But, we also provide theoretical radius of convergence which guaranteed for the convergence of iterative methods. In addition, our schemes (10) and (11) can further produce many more new optimal methods of order four and eight, respectively for each α . On the other hand, Sharma [29] and Li et al. [21] didn't get guess in order to obtain optimal methods for each α in their studies. On the accounts of results obtained in the Table 2, it can be concluded that the proposed methods are highly efficient as compared to the existing methods in term of computational efficiency and speed. We are claiming the superiority of our methods because we compare them on the same test problems with same initial approximations which are they taken in their papers (for detail please see Table 1). We also verify in section 4 that these methods converge to the required root even though the third derivative is not bounded. Finally, the dynamical behaviors of our methods also demonstrate the superiority to the other known methods in terms of larger and brighter basin of attraction and less chaotic.

References

- [1] Amat, S., Hernández, M.A., Romero, N. (2008) A modified Chebyshev's iterative method with at least sixth order of convergence. *Appl. Math. Comput.* **206**(1), 164–174.
- [2] Argyros, I.K. (1993) A note on the Halley method in Banach spaces. *Appl. Math. Comput.* **58**, 215–224.
- [3] Argyros, I.K. (2008) Convergence and Application of Newton-type Iterations. *Springer*.
- [4] Argyros, I.K. and Hilout, S. (2013) Numerical methods in Nonlinear Analysis. *World Scientific Publ. Comp. New Jersey*.
- [5] Argyros, I.K., Ezquerro, J.A., Gutierrez, J.M., Hernandez, M.A. and Hilout, S. (2011) On the semilocal convergence of efficient Chebyshev-Secant-type methods. *J. Comput. Appl. Math.* **235** 3195–3206.
- [6] Behl, R., Cordero, A., Motsa, S.S. and Torregrosa, J.R. (2015) On developing fourth-order optimal families of methods for multiple roots and their dynamics. *Appl. Math. Comput.* **265**, 520–532.
- [7] Behl, R. and Motsa, S.S. (2015) Geometric Construction of Eighth-Order Optimal Families of Ostrowski's Method. *Sci. Wor. J. Article ID 614612*.
- [8] Chen, D., Argyros, I.K., Qian, Q.S. (1994) A local convergence theorem for the super-Halley method in a Banach space. *Appl. Math. Lett.* **7**, 49–52.
- [9] Chun, C. (2007) Some second-derivative-free variants of Chebyshev-Halley methods. *Appl. Math. Comput.* **191**, 410–414.
- [10] Constantinides, A., Mostoufi, N. (1999) Numerical Methods for Chemical Engineers with MATLAB Applications. *Prentice Hall PTR, New Jersey*.

- [11] Devaney, R.L. (1992) A First Course in Chaotic Dynamical Systems, *Westview Press*.
- [12] Douglas, J.M. (1972) Process Dynamics and Control, *vol. 2 Prentice Hall, Englewood Cliffs*.
- [13] Grau-Sánchez, M., Noguera, M. and Manuel Gutiérrez, J. (2010) On some computational orders of convergence. *Appl. Math. Lett.*, **23**, 472–478.
- [14] Gutiérrez, J.M. and Hernández, M.A. (1997) A family of Chebyshev-Halley type methods in Banach spaces. *Bull. Aust. Math. Soc.*—, **55**, 113–130.
- [15] Kanwar, V., Singh, S. and Bakshi, S. (2008) Simple geometric constructions of quadratically and cubically convergent iterative functions to solve nonlinear equations. *Numer. Algor.* **47**, 95–107.
- [16] King, R.F. (1973) A family of fourth order methods for nonlinear equations. *SIAM J. Numer. Anal.* **10**, 876–879.
- [17] Kou, J. (2007) On Chebyshev-Halley methods with sixth-order convergence for solving non-linear equations. *Appl. Math. Comput.* **190**(1), 126–131.
- [18] Kou, J. and Li, Y. (2007) Modified Chebyshev's method free from second derivative for non-linear equations. *Appl. Math. Comput.* **187**, 1027–1032.
- [19] Kou, J. and Li, Y. (2007) Modified Chebishev–Halley method with sixth-order convergence. *Appl. Math. Comput.* **188**(1), 681–685.
- [20] Kung, H.T. and Traub, J.F. Optimal order of one-point and multi-point iteration. *J. Assoc. Comput. Mach.* **21** 643–651.
- [21] Li, D., Liu, P. and Kou, J. (2014) An improvement of Chebyshev–Halley methods free from second derivative. *Appl. Math. Comput.* **235**, 221–225.
- [22] Magreñán, Á. A. (2014) Different anomalies in a Jarratt family of iterative root-finding methods. *Appl. Math. Comput.* **233**, 29–38.
- [23] Magreñán, Á. A. (2014) A new tool to study real dynamics: The convergence plane. *Appl. Math. Comput.* **248**, 215–224.
- [24] Neta, B., Scot, M. and Chun, C. (2012) Basins of attraction for several methods to find simple roots of nonlinear equations. *Appl. Math. Comput.* **218**, 10548–10556.
- [25] Petković, M.S., Neta, B., Petković, L.D. and Džunić, J. (2012), Multipoint methods for solving nonlinear equations. *Academic Press*.
- [26] Potra, F.A. and Pták, V. (1984) *Nondiscrete induction and iterative processes. Research Notes in Mathematics, vol. 103, Pitman, Boston*.
- [27] Scott, M., Neta, B. and Chun, C. (2011) Basins attractors for various methods, *Appl. Math. Comput.* **218**, 2584–2599.
- [28] Rheinboldt, W.C. (1978) An adaptive continuation process for solving systems of nonlinear equations. *Polish Academy of Science, Banach Ctr. Publ.* **3**, 129–142.
- [29] Sharma, J.R. (2015) Improved Chebyshev–Halley methods with sixth and eighth order convergence. *Appl. Math. Comput.*, **256**, 119–124.
- [30] Traub, J.F. (1964) Iterative methods for the solution of equations. *Prentice-Hall, Englewood Cliffs*.
- [31] Weerakoon, S., Fernando, T.G.I.(2000) A variant of Newton's method with accelerated third-order convergence. *Appl. Math. Lett.* **13**, 87–93.
- [32] Xiaojian, Z. (2008) Modified Chebyshev-Halley methods free from second derivative. *Appl. Math. Comput.* **203**(2), 824–827.

A Measure of Engineering Students' Pro-industry Behavior Adjust Industry 4.0

Chun-Mei Chou^{1*} Chien-Hua Shen² Hsi-Chi Hsiao³ Tsu- Chuan Shen⁴

^{1*} Institute of Vocational and Technological Education, National Yunlin University of Science & Technology, Yunlin, Taiwan *E-mail: choucm@yuntech.edu.tw*

²Department of Business Administration, Transworld Institute of Technology, Yunlin, Taiwan

³Department of Business Administration, Cheng Shin University, Kaohsiung, Taiwan

⁴Department of Information Engineering, Feng Chia University, Taichung, Taiwan, R.O.C.

Abstract

This study describes the development and validation of the Pro-industry Behavior scale, a quantitative 24-item scale that measures an Asian country university students' behavior toward pro-industry. A total of 814 undergraduate engineering students completed the questionnaire. A pilot study (n= 154) was examined factorial validity and reliability of questionnaire and study objects (n= 658) used confirmatory factor analysis. The Pro-industry Behavior Measure (PBM) has three-factor model (Industry identity, Industry concern, and Industry regulator) was fit using maximum likelihood estimation (MLE). The Pro-Industry Behavior scale could be useful for understanding the ways in which students think about pro-industry issues and could be used to investigate the relationship between pro-industry behavior and other variables. The applications of the PBM were discussed.

Keywords: Pro-industry Behavior; Industry identity; Industry concern; Industry regulator

Introduction

In response to the rapidly growing global industry environment many call for changes in how individuals should deal with the industry 4.0. An important aspect of moving towards an industry sustainable development is to promote pro-industry behavior [1] [2]. Engineering students strengthen the practical skills is important who choose materials and compose ability of the industry practice course [3] [4]. In view of the practical needs, the engineering education curriculum content of the professional subjects is influenced by the industrial development trend. Engineering students' industry professional competence and specialized learning mechanism of pro-industry behavior will be emphasized of industry identity, industry concern, and industry regulator [6] [7].

Engineering students face two challenges of pro-industry behavior, there are: 1. To understand the impact of the pro-industry behaviors on employment abilities and seeking job of the career development, as well as to the industry practice abilities and preparation of industry practice, is very important. 2. The digitalization of work is not just something that lies ahead; it has already changed work more and more over the last few years, e.g. making it more mobile[8] [9]. In view of the industry practical needs of the engineering education contents of Industry 4.0, the content of professional subjects is influenced by the industrial development trend, and the industry practice competence of engineering students. In the face of the industry change, the industry employment ability of engineering students is challenged [6] [10] [11].

Students' pro-industry behavior was conducted in the industry practice specialization process of cognitive process, not only to participate in the common industry-oriented learning behavior, but also to practice quite personal characteristics [12]. The pro-industry behavior process is cognitive adjustment that to use knowledge and main contributions of this study and to set up the industry learning. It is important to understand students' views of pro-

industry behavior and learning in the vocation and technology education of human cultivation and industry connotation. The review of the literature reveals that the research studies conducted in the context of pro-industry learning and applications have measured attitudes and behavior of students regarding industry identity, industry concern, and industry regulator [12] [13]. The purpose of the study was to develop a valid and reliable instrument to be used for measuring engineering students' attitudes toward pro-industry behavior and its applications. With this instrument, it is believed that the gap in the professional literature indicated above will be partially met.

Method

A. Participants

(1) *Pilot study* : A total of 154 students that mean age of the participants was 21.3 (SD = 1.98) and there were 52 (33.8%) females in the sample.

(2) *The present study*: The participants in this study were 658 students from 20 engineering institutes in Asian countries [17]. The sample consisted of 322 women (48.9%), with a mean age of participants of 22.3 years (SD = 2.17). All participants are volunteers and receive no monetary or in-kind rewards. Participants in this study did not participate in the pilot study and collected data from the study within three months of the pilot study.

B. Measure

The purpose of this study was to evaluate the 24-PBM factorial validity. All participants were volunteers and they were briefed on the purpose of this study and informed of their rights not to participate and withdraw from completing the questionnaire at any time during or after the data have been collected [18] [19]. Participants took about 20 min to complete the questionnaire. This study aimed to test and refine the 28 items. These items were presented using a 5-point Likert response scale with 1 = strongly disagree and 5 = strongly agree. The principal component analysis with varimax-rotated see Table 1, and results of confirmatory factor analysis see Table 2.

Table 1 Principal component analysis with varimax-rotated

	II	IC	IR	H2
II 1	.824	.232	.342	.887
II 2	.798	.367	.268	.828
II 3	.837	.355	.289	.845
II 4	.842	.261	.389	.845
II 5	.799	.249	.243	.839
II 6	.812	.324	.317	.841
II 7	.866	.341	.216	.857
II 8	.872	.268	.311	.832
IC 1	.235	.789	.276	.824
IC 2	.341	.823	.312	.798
IC 3	.267	.821	.226	.891
IC 4	.289	.869	.317	.832
IC 5	.311	.732	.238	.819
IC 6	.243	.796	.326	.823
IC 7	.327	.839	.325	.856
IC 8	.345	.823	.210	.844
IC 9	.354	.793	.297	.865
IR 1	.288	.216	.863	.828
IR 2	.419	.329	.782	.867
IR 3	.234	.289	.808	.882
IR 4	.278	.390	.833	.878
IR 5	.342	.306	.842	.797
IR 6	.387	.398	.794	.874
IR 7	.279	.411	.789	.877
Eigenvalue	6.872	7.694	3.498	-
% of variance	32.56	39.34	17.56	-

Table 2 Result of confirmatory factor analysis

Item	Understandardized estimate	Standardized estimate	t value	R ²	α
<i>Industry Identity</i>					.921
II 1	.984	.8783	65.783	.789	
II 2	.955	.874	48.327	.765	
II 3	.992	.891	34.461	.740	
II 4	1.137	.992	24.043	.987	
II 5	1.109	.993	23.093	.935	
II 6	1.056	.992	21.434	.972	
II 7	1.104	.984	22.319	.925	
II 8	1.108	.992	22.378	.943	
<i>Industry Concern</i>					.919
IC 1	1.022	.973	27.389	.589	
IC 2	1.052	.983	26.359	.578	
IC 3	.993	.882	23.598	.542	
IC 4	1.018	.972	39.873	.923	
IC 5	1.361	.992	13.367	.962	
IC 6	1.388	.996	13.024	.938	
IC 7	1.403	.987	13.209	.942	
IC 8	1.387	.978	13.478	.956	
IC 9	1.484	.993	13.459	.971	
<i>Industry Regulator</i>					.934
IR 1	.969	.895	78.256	.965	
IR 2	1.022	.984	89.356	.978	
IR 3	1.001	.985	70.984	.958	
IR 4	.992	.895	79.320	.978	

explained				
-----------	--	--	--	--

Note. II=Industry identity; IC=Industry concern;
IR=Industry regulator. All factor loadings=.74 or
greater are underlined. H2=communality.

IR 5	1.022	.921	78.934	.967	
IR 6	.982	.899	78.544	.958	
IR 7	.965	.917	76.953	.944	

Conclusion

The purpose of this study was to test pro-industry awareness among engineering students using the newly developed Pro-industry Behavior Measure (PBM). The tool provides an alternative to existing measures where engineering students support industrial behavior, with a focus on student industry recognition, industry concerns and industry regulators' views on employment [20] [21]. The consisting of three factors, the PBM measures user perceptions of the industry, industry concerns, and industry regulators' views on using pro-business learning. Pro-business behavior measurement (PBM) is developed and validated through research using separate samples [22] [23].

In general, the validity of this study was found to support PBM as a measure of the utility of industrial students in supporting industry behavior. The results of the CFA show that the data for the third model is the best compared to the two alternative models and that these items have good normalized loading for the hypothetical underlying factors constructs, which are less highly correlated between them (see figure 1). These results provide evidence of the molecular structure of PBM and may be useful to educational researchers. A better understanding of students' understanding of the industry will increase their awareness of industry-related behaviors and will make teaching more meaningful in the field of education. Several researchers have demonstrated a positive relationship between student "pro-industry" awareness and their "pro-industry" behavior and career development. As part of supporting industries as part of teacher education, PBM allows researchers to measure and understand how users respond to instruction [23] [24]. In doing so, the usefulness of PBM can be expanded to further inform researchers about the factors that affect user behavior. Such future research may be based on user demographics, such as the level of industry development, the level of industry experience, and attitudes toward industrial learning.

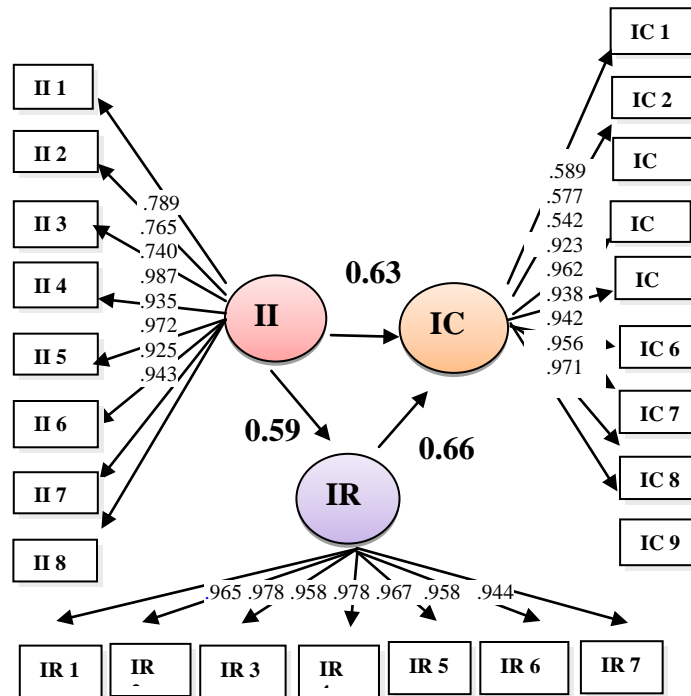


Figure 1 The PBM measures users' perception

ACKNOWLEDGMENT

This paper was written while the authors were supported by a grant from the Ministry of Science and Technology (MOST 105-2511-S-224-001-MY3 & MOST 105-2511-S-224-001-MY3)

REFERENCES

- [1] T. Stock, and G. Seliger, “Opportunities of sustainable manufacturing in industry 4.0,” *Proc. CIRP*, vol. 40, pp. 536–541, 2016.
- [2] D. Gorecky, M. Schmitt, M. Loskyll, and D. Zühlke, “Human-machineinteraction in the industry 4.0 era,” in *Proc. 12th IEEE Int. Conf. Ind. Informat. (INDIN)*, Jul. 2014, pp. 289–294.
- [3] K. Pažur Aničić, B. Divjak, and K. Arbanas, “Preparing ICT graduates for real-world challenges: Results of a meta-analysis,” *IEEE Trans. Educ.*, vol., no. , pp.1-7, 2016.
- [4] S.R. Powers, and K.K. Myers, “Vocational anticipatory socialization: College students’ reports of encouraging/discouraging sources and messages,” *J. Car. Deve.*, pp.1-16, Aug. 2016.
- [5] National Science Council, “Science and Technology of the Republic of China White (Republic of China 100 years to 103 years). Accessed on Aug. 23, 2016. [Online]. Available: <https://www.most.gov.tw/most/attachments/5e33a3c7-589e-4cc5-962f-7f2b4625eb64>, 2015.
- [6] J. Pérez, C. Vizcarro, J. García, A. Bermúdez, and R. Cobos, “Development of procedures to assess problem-solving competence in computing engineering,” *IEEE Trans. Educ.*, vol.60, no.1, pp.22-28, Feb. 2017.
- [7] S. A. Male, M. B. Bush, and E. S. Chapman, “An Australian study of generic competencies required by engineers,” *Eur. J. Eng. Educ.*, vol. 36, no. 2, pp. 151–163, May 2011.
- [8] Gracia, L., “Employability and higher education: contextualising female students’ workplace experiences to enhance understanding of employability development”, *J. Edu. and Work*, Vol. 22, No.4, pp. 301–318, 2009.
- [9] J. M. Foley, S. Daly, C. Lenaway, and J. Phillips, “Investigating student motivation and performance in electrical engineering and its subdisciplines,” *IEEE Trans. Educ.*, vol.59, no.4 , pp.241-249, Nov. 2016.
- [10] W.H. Thomas, and D.C. Feldman, “Personality, social relationships, and vocational indecision among college students: The mediating effects of identity construction”, *Career Development International*, vol. 14, no.4, pp.309-332, 2009.
- [11] T. Niesen, C. Houy, P. Fettke, and P. Loos, “Towards an integrative big data analysis framework for data-driven risk management in industry 4.0,” in *Proc. 49th Hawaii Int. Conf. Syst. Sci. (HICSS)*, Jan. 2016, pp. 5065–5074.
- [12] D. Zhang, J. Wan, C.H.R. Hsu, and A. Rayes, “Industrial technologies and applications for the Internet of Things,” *Comput. Netw.*, vol. 101, pp. 1–4, Jun. 2016.
- [13] C. Faller, and D. Feldmüller, “Industry 4.0 learning factory for regional SMEs,” *Proc. CIRP*, vol. 32, pp. 88–91. 2015.
- [14] R. Dhalla, and C. Oliver, “Industry identity in an oligopolistic market and firms’ responses to institutional pressures,”*Orga. Studies*, vol.34, no.12, pp.1803–1834. 2013.
- [15] J. Nelles, S. Kuz, A. Mertens, and C. M. Schlick, “Human-centered design of assistance systems for production planning and control: The role of the human in industry 4.0,” *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, pp. 2099–2104, Mar. 2016.
- [16] F.T. Cheng, et al., “Industry 4.1 for wheel machining automation,” *IEEE Robot. Autom. Lett.*, vol. 1, no. 1, pp. 332–339, Jan. 2016.
- [17] G.T. Chao, A.M. O’Leary-Kelly, S. Wolf, H.J. Klein and P.D. Gardner “Organizational socialization: Its content and consequences, *Journal of App. Psy.*, vol.79, no.5, pp.730-743, 1994.
- [18] S. Wang, J. Wan, D. Li, and C. Zhang. “Implementing Smart Factory of Industrie 4.0: An Outlook,” *Inter. J. Distri. Sensor Networks*, 2016.
- [19] J.Y. Chen, and Y.C. Liu, “A study on the determinants of regional distribution of new firms in service industry for Taiwan,” *J. of Entre. Res.* vol. 6 , no.2 , pp.1-32, 2011.
- [20] S. Wang, J. Wan, D. Zhang, Li, D., and Zhang, C. “Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination,”*Computer Networks*, vol.101, no.4, pp.158-168, 2016.
- [21] J. Lee, B. Bagheri, and H. A. Kao, “Research letters a cyber-physical systems architecture for Industry 4.0-based manufacturing systems,” *Manu. Letters*, vol. 3,pp.18-23, 2015.
- [22] C. Faller, and D. Feldmüller, “Industry 4.0 learning factory for regional SMEs,”*Procedia CIRP*, vol.32, pp.88-91, 2015.
- [23] G. Schuh, T. Gartzten, T. Rodenhauser, and A. Marks, “Promoting work-based learning through Industry 4.0,” *Pro. CIRP*, vol. 32, p.82-87, 2015.

- [24] R.V. Krejcie, and D.W. Morgan, "Determining sample size for research activities," *Edu. and Psy. Measur.*, vol. 30, no. 3, p.607–107, 1970.

A Novel Method Combining Reverse Modeling and Topological Optimization for Lightweight Design of Automobile Wheel Hubs with Hollow Ribs

P.F.Xu¹, †S.Y.Duan, S. Liu¹

¹ College of Mechanical Engineering, HeBei University of Technology, Tianjin City, 300401 China

†Corresponding author: dsy605106@163.com

Abstract

The lightweight of wheel hubs is very important for reducing the unsprung mass and critical to improve the vehicle dynamic and braking performance of vehicles, and hence the control stability and comfortability. However, many current experience-based lightweight designs of wheel hubs have shortcomings resulting in uneven distribution of materials in terms of mechanics principles. This work develops a new method that combines the reverse modeling with topological optimization techniques to obtain lightweight wheel hubs, based on the principles of mechanics. A reverse modeling technique is firstly used to scan and create the initial 3D geometry of the wheel hub with solid ribs. The finite element (FE) method is then used to perform stress analysis to find the maximum stress and its location of wheel hub under multiple conditions. The FE model is then divided into optimization region and unoptimized region: the former the interior portion, and the latter is the outer surface of the wheel hub. A topology optimization is next the conducted, leading to a wheel hub with hollow ribs: the interior material of the ribs of the wheel hub is automatically removed. Finally, the hollow wheel hub is reconstructed with wall thickness about 5mm, via again a reverse modeling technique. Our results show that the reconstructed model can reduce the mass of 12.7% compared to the pre-optimized model (from 11.26 kg to 9.83 kg). The present method of combined with reverse modeling and topology optimization can guarantee the optimal distribution of wheel hub material based on mechanics principle. It can be performed automatically and hence shorten the time for optimal lightweight designs to improve the performance of the many other existing structures and structure components.

Keywords: Reverse modeling; hollow wheel hub; Finite element; Topology optimization; Lightweight structures.

The 3D geometry model of the wheel hub

Reverse modeling technique is used to create the 3D geometry of the five-spoke aluminum alloy wheel hub of a selling SUV. The weight and size of wheel hub are 10.5 kg and 17×6.5 inches, respectively. Current mainstream 3D handheld scanner was selected to scan the wheel hub due to that the size of the wheel hub is not very large and the appearance accuracy requirement is not high [5]. Creating the wheel hub model includes the following three main steps, shown as in Fig. 1. In the first step, the data acquisition is to scan the wheel hub with a hand-held laser scanner with an accuracy of 0.5 mm, thereby obtaining a reverse mesh as shown in Fig. 1 (a). In the second step, the reverse mesh editing is the optimization of the surface contour of the reverse mesh in the reverse software as shown in Fig. (b). In the last step, the 3D geometry model of the wheel hub available for topology optimization with in the solid modeling step is obtained based on the reverse mesh.

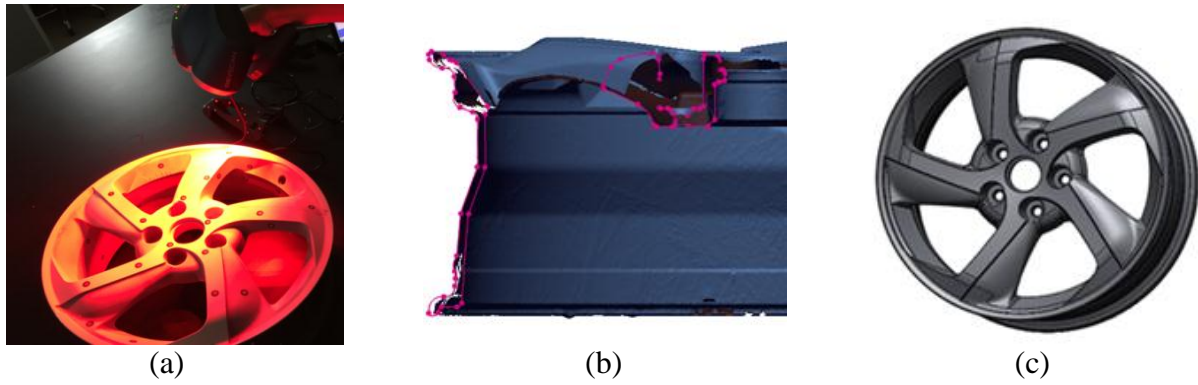


Figure 1. Reverse modeling, (a) data acquisition, (b) reverse mesh editing, (c) solid modeling

Numerical analysis of the wheel hub

Aluminum alloy A356 material is selected for casting wheel hubs. The material parameters of the aluminum alloy A356 material are shown in Table 1. [2]. The finite element analysis model is assembled from a wheel hub model and a torque shaft as shown in Fig. 2 [6]. For the one part, the mass of wheel hub model can be calculated in Abaqus and it is 11.26 kg. Due to the modeling error in reverse modeling, the mass of the FE wheel hub model is not equal to that of the real wheel hub. For the other part, axle was modeled using the material parameters of Q235-B material, and according to the formula (1) the length of axle is 500mm. Load: The static F_i ($i=1,2,3,\dots,24$) was applied every 15° within 360° of the axle end, the static load F_i applied at the axle end is 5265N, the role of F_i points and directions as shown in Fig. 2. In order to simulate the actual operation of the wheel hub in a cycle when the load suffered. The bending moment of a single wheel in this cycle was provided by the automobile manufacturer for 2632Nm. In 错误!未找到引用源。 , the axle length of the loading moment and it is 500mm, and according to the formula (1):

$$M = FL \quad (1)$$

Boundary conditions: The six degrees of freedom of the rim based plane and the rim edge surface were constrained. as shown in the red area of 错误!未找到引用源。 [7].

Table 1. Material properties of A356

Elasticity Modulus	Poisson's ratio	Density	Yield stress	Ultimate stress
72.00GPa	0.33	$2.67 \times 10^3 \text{ kg/m}^3$	218.0MPa	283.0MPa

The 10-node modified quadratic tetrahedron was applied to assembly model. This element type makes it easy to mesh complex contact-containing models. The elements number of the wheel hub FE model was 169173. The node number of the wheel hub was 281034. The maximum stress position within wheel hub rotation could be obtained, and the maximum stress was 128.29MPa as shown in Fig. 4(a). Topology optimization will be based on this.

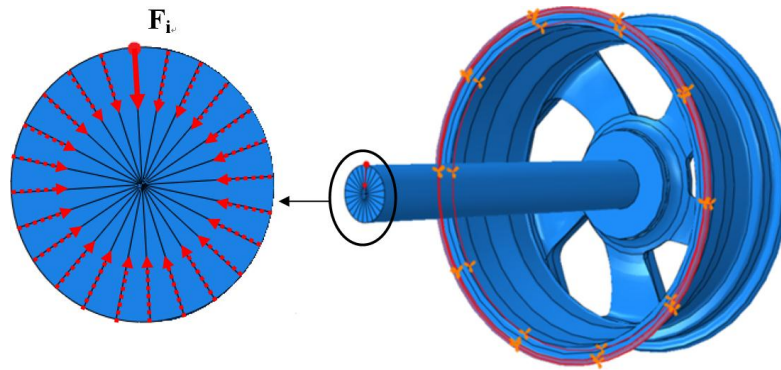


Figure 2. Finite element model of wheel hub

Structure optimization and Model reconstruction

In this section, a model of hollow wheel hub was obtained by topology optimization technology. According to material mechanics, it could be known that the normal stress of a beam was proportional to the distance from the neutral layer as Eq. (2).

$$\sigma = E(y/\rho) \quad (2)$$

The surface material of the wheel hubs withstands most of the load on wheel hubs. The material of wheel hubs with less efficient can be removed. Variable density topology optimization was used to reduce the wheel hub material in Abaqus. The FE model is then divided into optimization region and unoptimized region: the former the interior portion, and the latter is the outer surface of the wheel hub. The design area allows to reduce the material during optimization and it was spokes of wheel as shown in Fig. 2. Cell density is design variable. The optimization objective function is set to minimize strain energy values. The purpose is to maximize the overall stiffness of the hub. The optimization Constraint is set to optimized volume $\leq 75\%$ of the original volume. The method of freezing the surface load area can be chosen to preserve the material of the wheel hub surface. The load is a pressure of 1 Pa applied to the surface of the design area during optimization (This stress generated by this pressure was $1 \times 10^{-6} \%$ of the maximum stress on the FE model), so it could be neglected. The topology optimization would remove the material inside the wheel hub mesh model. The hollow structure could be obtained as shown in Fig. 3.

The optimized mesh model was derived and reconstructed using the same method with the reverse modeling. A new weight of the hollow wheel hub can be obtained as 9.83 kg. The weight of optimized wheel hub is 12.7% (1.43 kg) less the original wheel hub model. The same finite element analysis was applied to the optimized model. The larger element can speed up computing, and the result as shown in Fig. 4(b), the maximum stress position had changed. It located in the area where spoke and rim connecting position, with a maximum stress of 130MPa. From the FE results, it can be seen that the wheel hub mass is significantly reduced while the maximum stress of the original wheel hub and the optimized wheel hub is almost the same. This result is satisfactory [8].

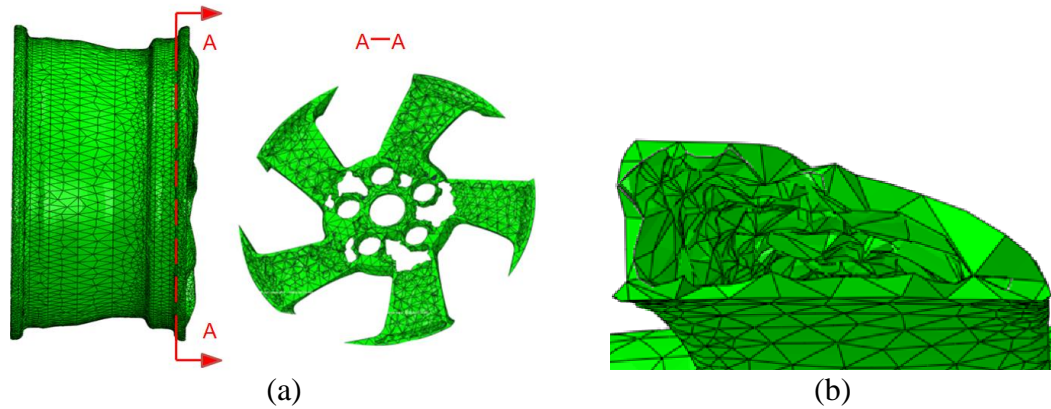


Figure 3. The optimized wheel hub model, (a) the optimized model is cut along the A-A section the, (b) radial sectional view of a hollow single-spoke.

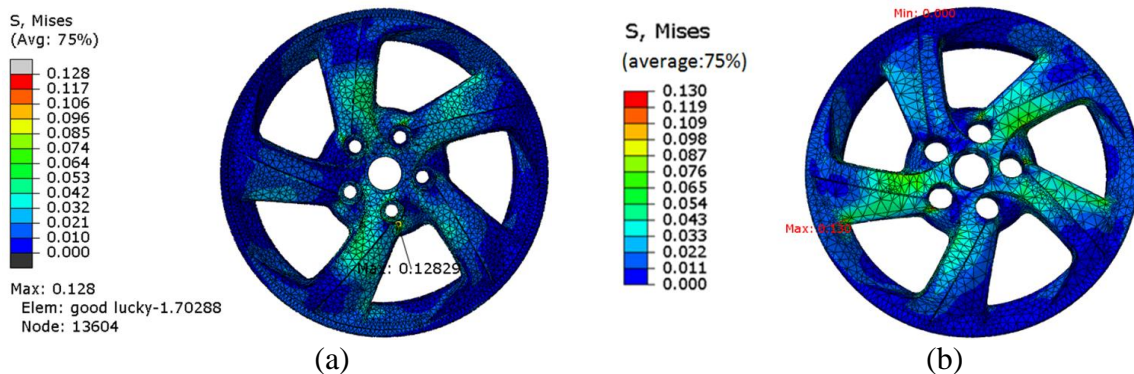


Figure 4. (a) Original wheel hub stress cloud. (b) Optimized wheel hub stress cloud

References

- [1] Shuyong Duan, Yourui Tao, Xu Han, Xujing Yang, Shujuan Hou, Zhangping Hu. (2014) Investigation on structure optimization of crashworthiness of fiber reinforced polymers materials. Part B 60 471-478.
- [2] Ch. P. V. Ravi Kumar, Prof. R. Satya Meher. (2013) Topology Optimization of Aluminium Alloy Wheel. International Journal of Modern Engineering Research. pp-1548-1553.
- [3] G. Chiandussi, I. Gaviglio, A. Ibba. (2004) Topology optimization of an automotive component without final volume constraint specification. Advances in Engineering Software 35 609-617.
- [4] Denghong Xiao, Hai Zhang, Xiandong Liu, Tian He and Yingchun Shan. (2014) Novel steel wheel design based on multi-objective topology optimization. Journal of Mechanical Science and Technology 28 (3) 1007-1016.
- [5] Li Z, Xiang HY, Li ZQ, Han BA, Huang JJ. (2013) The Research Of Reverse Engineering Based On Geomagic Studio. Applied Mechanics and Materials, ISSN: 1662-7482, Vols, 365-366, pp 133-136.
- [6] ISO 3006:2005-Road vehicles--Passenger car wheels for road use--Test methods.
- [7] Liu GR, Quek SS. (2013). *The finite element method: A practical course, 2nd edition*.
- [8] Barbieri SG (Barbieri Saverio Giulio); Giacomini M (Giacomini Matteo); Mangeruga V (Mangeruga Valerio); Mantovani S (Mantovani Sara). (2017) A design strategy based on topology optimization techniques for an additive manufactured high performance engine piston. Procedia Manufacturing, Volume 11, Pages 641-649.

Shape identification of unsteady heat convection fields to control temperature distribution

*†E. Katamine¹ and N. Okada²

¹Dep. Mechanical Engineering, National Institute of Technology, Gifu College, 2236-2 Kamimakuwa, Motosu-shi, Gifu, Japan

²Isuzu Engineering Co., Ltd. 8 Tsuchidana, Fujisawa-shi, Kanagawa, Japan

*Presenting author: katamine@gifu-nct.ac.jp

†Corresponding author: katamine@gifu-nct.ac.jp

Abstract

This paper presents numerical solution to a shape identification problem to control temperature distribution to a target distribution in sub-domains of unsteady heat convection fields. The square error integral between the actual temperature distributions and the target temperature distributions in the sub-domains during the specified period of time is used as the objective functional. Shape gradient of the shape identification problem is derived theoretically using the Lagrange multiplier method, adjoint variable method, and the formulae of the material derivative. Reshaping is carried out by the traction method proposed as an approach to solving shape optimization problems. Numerical analyses program for the shape identification is developed based on FreeFem++, and the validity of proposed method is confirmed by results of 2D numerical analyses.

Keywords: Inverse problem, Shape identification, Optimum design, Flow control, Traction method

Introduction

Shape design problems that improve the characteristics of heat transfer in thermal convection fields are an important subject in engineering. A typical example of such a problem can be seen in the design process used to create a heat exchanger. In this problem, the shape of the heat exchanger is optimized to maximize the heat discharge on the sub-boundaries of the convection field. Moreover, the problem of determining the boundary shape that can achieve the desired state distribution function of temperature or flow velocity on specified sub-boundaries, or in specified sub-domains, in a heat convection field is known as an inverse problem. If we regard the inverse problem as designing the shape needed to minimize the integrated squared error between the state distribution function of the actual temperature distribution and the target distribution function, then it can be treated as one optimization problem. This study discusses the solution of the inverse problem and the shape optimization problem with regard to the shape design of the heat convection field domain.

The theory of shape optimization for incompressible viscous flow fields was initiated by Pironneau [Pironneau(1973; 1974; 1984)], who formulated a shape optimization problem for an isolated body located in a uniform viscous flow field to minimize the drag power on this body. The distributed shape sensitivity, which is called the shape gradient, was derived with respect to the domain variation by means of an adjoint variable method based on optimal control theory. The adjoint variable method introduces adjoint variables into variational forms of the governing equations as variational variables; it also determines the adjoint variables using adjoint equations derived from criteria defining an optimality condition with respect to the domain variation.

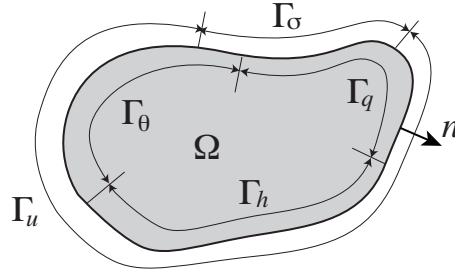


Figure 1: Heat convection field

The present authors have proposed an approach for the shape optimization of such channels or bodies based on a gradient method using the distributed shape sensitivity. In previous studies, the present authors presented a numerical method for the minimization of the dissipation energy of steady-state viscous flow fields [Katamine and Azegami(1995); Katamine et al.(2005)] and extended this method to 3D problems [Katamine et al.(2009)]. Also, the present authors applied this method to the shape optimization solution for the drag minimization and lift maximization of an isolated body located in a uniform viscous flow field [Katamine and Matsui(2012)] and the shape identification problem of flow velocity distribution prescribed problem in sub-domains of steady-state viscous flow fields [Katamine and Kanai(2016)].

The present study describes the extension of this method for solving a shape identification problem of unsteady forced heat convection fields to control temperature distribution to target distribution in sub-domains of the fields. Reshaping is accomplished using the traction method [Azegami et al.(1995; 1997); Azegami(2000)], which was proposed as a means of solving boundary shape optimization problems of domains. In the traction method, domain variations that minimize the objective functional are obtained as solutions of pseudo-linear elastic problems for continua defined in the design domain. These continua are loaded with pseudo-distributed traction in proportion to the shape gradient in the design domain.

In this study, the shape identification problem is formulated in the unsteady heat convection fields. The square error integral between the actual temperature distributions and the target temperature distributions in the sub-domains during the specified period of time is used as the objective functional. Shape gradient of the shape identification problem is derived theoretically using the Lagrange multiplier method, adjoint variable method, and the formulae of the material derivative. Reshaping is carried out by the traction method proposed as an approach to solving shape optimization problems. Numerical analyses program for the shape identification is developed based on FreeFem++, and the validity of proposed method is confirmed by results of 2D numerical analyses.

Governing equations for unsteady heat convection fields

Consider the unsteady heat convection field in the region Ω of \mathbb{R}^d ($d = 2, 3$) in time interval $[0, T]$. Consider determining the flow velocity $u(\vec{x}, t) = (u_i(\vec{x}, t))_{i=1,d}$, pressure $p(\vec{x}, t)$ and temperature $\theta(\vec{x}, t)$ at $\vec{x} \in \Omega$ and time $t \in [0, T]$. The dimensionless forms of the Navier–Stokes equation, continuity equation, and energy equation are the governing equations for unsteady heat convection fields. They can be expressed as follows:

$$\frac{\partial u_i}{\partial t} + u_j u_{i,j} = -p_{,i} + \frac{1}{Re} u_{i,jj}, \quad (\vec{x}, t) \in \Omega \times [0, T], \quad (1)$$

$$u_{i,i} = 0, \quad (\vec{x}, t) \in \Omega \times [0, T], \quad (2)$$

$$\frac{\partial \theta}{\partial t} + u_j \theta_{,j} = \frac{1}{RePr} \theta_{,jj}, \quad (\vec{x}, t) \in \Omega \times [0, T], \quad (3)$$

where the boundary $\Gamma = \partial\Omega = \Gamma_u \cup \Gamma_\sigma = \Gamma_\theta \cup \Gamma_q \cup \Gamma_h$, are Figure 1. Tensors described in this study use the Einstein summation convention and differentiation $(\cdot)_{,i} = \partial(\cdot)/\partial x_i$. The boundary conditions and initial conditions are described below:

$$u_i(\vec{x}, t) = \hat{u}_i(\vec{x}, t), \quad t \in [0, T], \quad \vec{x} \in \Gamma_u, \quad (4)$$

$$\sigma_i(\vec{x}, t) = \hat{\sigma}_i(\vec{x}, t) = (-p\delta_{ij} + \frac{1}{Re}u_{i,j})n_j = 0, \quad t \in [0, T], \quad \vec{x} \in \Gamma_\sigma \quad (5)$$

$$\theta(\vec{x}, t) = \hat{\theta}(\vec{x}, t), \quad t \in [0, T], \quad \vec{x} \in \Gamma_\theta, \quad (6)$$

$$-\frac{1}{RePr}\theta(\vec{x}, t)_{,j}n_j = \hat{q}(\vec{x}, t), \quad t \in [0, T], \quad \vec{x} \in \Gamma_q, \quad (7)$$

$$-\frac{1}{RePr}\theta(\vec{x}, t)_{,j}n_j = \hat{h}(\theta(\vec{x}, t) - \hat{\theta}_f), \quad t \in [0, T], \quad \vec{x} \in \Gamma_h, \quad (8)$$

$$u_i(\vec{x}, 0) = u_{i_{ini}}(\vec{x}), \quad \vec{x} \in \Omega, \quad (9)$$

$$p(\vec{x}, 0) = p_{ini}(\vec{x}), \quad \vec{x} \in \Omega, \quad (10)$$

$$\theta(\vec{x}, 0) = \theta_{ini}(\vec{x}), \quad \vec{x} \in \Omega. \quad (11)$$

Here, \hat{q} represents the heat flux, \hat{h} represents the coefficient of heat transfer, $\hat{\theta}_f$ represents the external temperature, δ_{ij} represents the Kronecker delta, Re is the Reynolds number, and Pr is Prandtl number. $(\hat{\cdot})$ represents the known function on the boundary. \vec{n} is an outward unit normal vector to the boundary. Also, $u_{i_{ini}}$ represents the initial flow velocity, p_{ini} represents the initial pressure, and θ_{ini} represents the initial temperature.

The weak forms of the respective governing equations (1)-(3) can be expressed with adjoint flow velocity $w(\vec{x}, t) = (w_i(\vec{x}, t))_{i=\overline{1,d}}$, adjoint pressure $q(\vec{x}, t)$, and adjoint temperature $\xi(\vec{x}, t)$ as follows:

$$\int_0^T \left\{ t^V(u_{,t}, w) + a^V(u, w) + b^V(u, u, w) + c(w, p) - l(w) \right\} dt = 0, \quad \forall w \in W, \quad (12)$$

$$\int_0^T \left\{ c(u, q) \right\} dt = 0, \quad \forall q \in Q, \quad (13)$$

$$\int_0^T \left\{ t^H(\theta_{,t}, \xi) + a^H(\theta, \xi) + b^H(u, \theta, \xi) + f_q^H(\xi) + f_h^H(\theta, \xi) - f_{hf}^H(\xi) \right\} dt = 0, \quad \forall \xi \in \Xi. \quad (14)$$

Furthermore, $t^V(u_{,t}, w)$, $t^H(\theta_{,t}, \xi)$, $a^V(u, w)$, $b^V(v, u, w)$, $c(w, p)$, $l(w)$, $a^H(\theta, \xi)$, $b^H(u, \theta, \xi)$, $f_q^H(\xi)$, $f_h^H(\theta, \xi)$, and $f_{hf}^H(\xi)$ are defined as follows:

$$\begin{aligned} t^V(u_{,t}, w) &= \int_\Omega w_i \frac{\partial u_i}{\partial t} dx, \quad t^H(\theta_{,t}, \xi) = \int_\Omega \xi \frac{\partial \theta}{\partial t} dx, \\ a^V(u, w) &= \int_\Omega \frac{1}{Re} w_{i,j} u_{i,j} dx, \quad b^V(v, u, w) = \int_\Omega w_i v_j u_{i,j} dx, \\ c(w, p) &= - \int_\Omega w_{i,i} p dx, \quad l(w) = \int_{\Gamma_\sigma} w_i \hat{\sigma}_i d\Gamma, \\ a^H(\theta, \xi) &= \int_\Omega \frac{1}{RePr} \xi_{,i} \theta_{,i} dx, \quad b^H(u, \theta, \xi) = \int_\Omega \xi u_j \theta_{,j} dx, \quad f_q^H(\xi) = \int_{\Gamma_q} \xi \hat{q} d\Gamma, \\ f_h^H(\theta, \xi) &= \int_{\Gamma_h} \hat{h} \xi \theta d\Gamma, \quad f_{hf}^H(\xi) = \int_{\Gamma_h} \hat{h} \xi \hat{\theta}_f d\Gamma. \end{aligned} \quad (15)$$

Here, $(\cdot)_t$ expresses the time derivative of the function. The flow velocity u , its adjoint w , and the other variables are considered to be elements of the following functional spaces:

$$U = \{u(\vec{x}, t) \in H^1(\Omega \times [0, T]) \mid u \text{ satisfies (4) and (9)}\}, \quad (16)$$

$$Q = \{q(\vec{x}, t) \in L^2(\Omega \times [0, T]) \mid \int_{\Omega} q \, dx = 0 \text{ (if } \text{measure}(\Gamma_{\sigma}) = 0)\}, \quad (17)$$

$$\Theta = \{\theta(\vec{x}, t) \in H^1(\Omega \times [0, T]) \mid \theta \text{ satisfies (6) and (11)}\}, \quad (18)$$

$$W = \{w_i(\vec{x}, t) \in H^1(\Omega \times [0, T]) \mid w_i(\vec{x}, t) = 0, \, t \in [0, T], \, \vec{x} \in \Gamma_u, \, w_i(\vec{x}, T) = 0, \, \vec{x} \in \Omega\}, \quad (19)$$

$$\Xi = \{\xi(\vec{x}, t) \in H^1(\Omega \times [0, T]) \mid \xi(\vec{x}, t) = 0, \, t \in [0, T], \, \vec{x} \in \Gamma_{\theta}, \, \xi(\vec{x}, T) = 0, \, \vec{x} \in \Omega\}. \quad (20)$$

Prescribing temperature in sub-domain in unsteady heat convection fields

In this section, the problem of minimizing the square integration errors between the actual temperature $\theta|_{\Omega_D \times [t_1, t_2]}$ from time $t = t_1 \in [0, T]$ to $t = t_2 \in [0, T]$ and the target temperature $\theta_D|_{\Omega_D \times [t_1, t_2]}$ in sub-domain $\Omega_D \subset \Omega$ is formulated. We assume $t_1 < t_2$. The domain transformation of this heat convection field region Ω is denoted by \vec{T}_s , and the domain Ω is assumed to vary to reach $\Omega_s = \vec{T}_s(\Omega)$ [Azegami et al.(1995; 1997)]. For simplicity, we assume that the sub-domains Ω_D and Γ_{σ} are invariable, that is $\vec{T}_s(\Omega_D) = \Omega_D$ and $\vec{T}_s(\Gamma_{\sigma}) = \Gamma_{\sigma}$ or domain variation. The square integration error problem for temperature distribution from time $t = t_1$ to $t = t_2$ is formulated as follows:

$$\begin{aligned} \text{Given } \Omega \quad \text{find } \Omega_s \quad \text{that minimizes } & \int_{t_1}^{t_2} E_{\Omega_D}(\theta) \, dt \quad \text{subject to (12) – (14) and} \\ & \int_{\Omega} dx \leq \beta_V M. \end{aligned} \quad (21)$$

where β_V is a coefficient related to the initial domain measure M , and

$$E_{\Omega_D}(\theta) = \int_{\Omega_D} (\theta - \theta_D)^2 \, dx. \quad (22)$$

The Lagrange function $L(u_i, p, \theta, w_i, q, \xi, \Lambda)$ for this problem is given as follows:

$$\begin{aligned} L = & \int_{t_1}^{t_2} E_{\Omega_D}(\theta) \, dt \\ & - \int_0^T \left\{ t^V(u_{,t}, w) + a^V(u, w) + b^V(u, u, w) + c(w, p) - l(w) \right\} dt - \int_0^T \left\{ c(u, q) \right\} dt \\ & - \int_0^T \left\{ t^H(\theta_{,t}, \xi) + a^H(\theta, \xi) + b^H(u, \theta, \xi) + f_q^H(\xi) + f_h^H(\theta, \xi) - f_{hf}^H(\xi) \right\} dt \\ & + \Lambda \left(\int_{\Omega} dx - \beta_V M \right). \end{aligned} \quad (23)$$

where $w \in W$, $q \in Q$, and $\xi \in \Xi$ were introduced as Lagrange multiplier functions or the adjoint functions with respect to the weak forms. The non-negative real constant number Λ is the Lagrange multiplier with respect to the volume constraint. The derivative of L with respect

to domain variation is derived using the velocity field $\vec{V}(\Omega_s) = \partial \vec{T}_s(\Omega) / \partial s = \partial \vec{T}_s / \partial s(\vec{T}_s^{-1}(\Omega_s))$, as follows [Azegami et al.(1995; 1997)]:

$$\begin{aligned} \dot{L} = & - \int_0^T \left\{ t^V(u, w') + a^V(u, w') + b^V(u, u, w') + c(w', p) - l(w') + c(u, q') \right\} dt \\ & - \int_0^T \left\{ t^H(\theta, \xi') + a^H(\theta, \xi') + b^H(u, \theta, \xi') + f_q^H(\xi') + f_h^H(\theta, \xi') - f_{hf}^H(\xi') \right\} dt \\ & - \int_0^T \left\{ t^V(u', w) + a^V(u', w) + b^V(u', u, w) + b^V(u, u', w) + c(u', q) + b^H(u', \theta, \xi) + c(w, p') \right\} dt \\ & - \int_0^T \left\{ t^H(\theta', \xi) + a^H(\theta', \xi) + b^H(u, \theta', \xi) + f_h^H(\theta', \xi) \right\} dt + \int_{t_1}^{t_2} E_{\Omega_D}(\theta') dt \\ & + \dot{\Lambda} \left(\int_{\Omega} dx - \beta_V M \right) + l_G(\vec{V}). \end{aligned} \quad (24)$$

Here, $(\cdot)'$ represents the derivative with respect to domain variation of the function fixed on the spatial coordinates, and

$$l_G(\vec{V}) = \int_{\Gamma} G \vec{n} \cdot \vec{V} d\Gamma, \quad (25)$$

and assuming that the flow velocity satisfies $u_i = 0$ at the design boundary,

$$\begin{aligned} G &= G_0 + G_1 \Lambda, \\ G_0 &= \int_0^T \left\{ -\frac{1}{Re} w_{i,j} u_{i,j} - \frac{\partial \theta}{\partial t} \xi - \frac{1}{Re Pr} \xi_{,i} \theta_{,i} \right. \\ &\quad \left. - \nabla_n(\xi \hat{q}) - (\xi \hat{q}) \kappa - \nabla_n(\hat{h} \xi \theta) - (\hat{h} \xi \theta) \kappa + \nabla_n(\hat{h} \xi \hat{\theta}_f) + (\hat{h} \xi \hat{\theta}_f) \kappa \right\} dt \\ G_1 &= 1, \end{aligned} \quad (26)$$

where $\nabla_n(\cdot) \equiv \nabla(\cdot) \cdot \vec{n}$, and κ denotes the quantity $(d-1)$ times the mean curvature of boundary, and $u_i, p, \theta, w_i, q, \xi$, and Λ are determined by the following conditions:

$$\begin{aligned} \int_0^T \left\{ t^V(u, w') + a^V(u, w') + b^V(u, u, w') + c(w', p) - l(w') + c(u, q') \right\} dt &= 0 \\ \forall w' \in W, \forall q' \in Q \end{aligned} \quad (27)$$

$$\int_0^T \left\{ t^H(\theta, \xi') + a^H(\theta, \xi') + b^H(u, \theta, \xi') + f_q^H(\xi') + f_h^H(\theta, \xi') - f_{hf}^H(\xi') \right\} dt = 0 \quad \forall \xi' \in \Xi \quad (28)$$

$$\begin{aligned} \int_0^T \left\{ -t^V(u', w) + a^V(u', w) + b^V(u', u, w) + b^V(u, u', w) + c(u', q) \right. \\ \left. + b^H(u', \theta, \xi) + c(w, p') \right\} dt = 0 \quad \forall u' \in U, \forall p' \in Q \end{aligned} \quad (29)$$

$$\int_0^T \left\{ -t^H(\theta', \xi_t) + a^H(\theta', \xi) + b^H(u, \theta', \xi) + f_h^H(\theta', \xi) \right\} dt - \int_{t_1}^{t_2} E_{\Omega_D}(\theta') dt = 0 \quad \forall \theta' \in \Theta \quad (30)$$

$$\Lambda \geq 0, \quad \int_{\Omega} dx \leq \beta_V M, \quad \Lambda \left(\int_{\Omega} dx - \beta_V M \right) = 0. \quad (31)$$

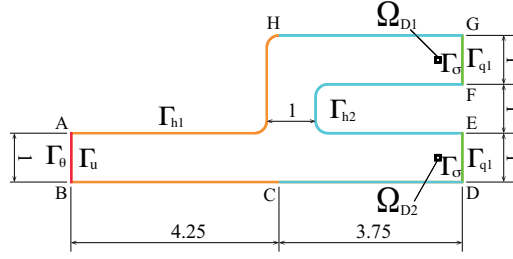
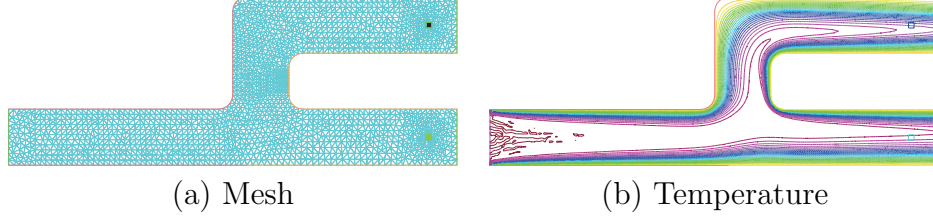
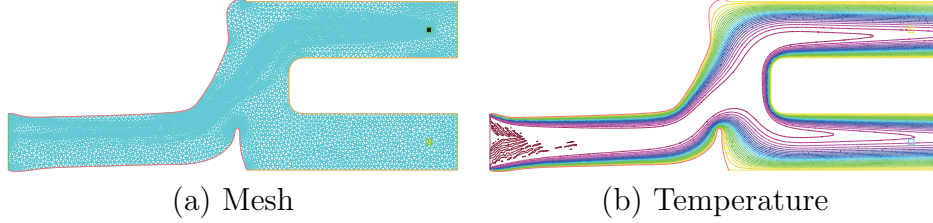


Figure 2: Numerical model: Branch channel, prescribing temperature distribution


 Figure 3: Numerical results: Mesh and temperature distribution at final time ($t=400$) for initial shape

 Figure 4: Numerical results: Mesh and temperature distribution at final time ($t=400$) for the identified shape

The derivative of the Lagrange function agrees with the derivative of the evaluation function, establishing the following relationship:

$$\dot{L}|_{u_i, p, \theta, w_i, q, \xi, \Lambda} = \dot{E}_{\Omega_D}|_{u_i, p, \theta, w_i, q, \xi, \Lambda} = l_G(\vec{V}). \quad (32)$$

Since $G\vec{n}$ in equation (25) is a coefficient function of the velocity field \vec{V} that provides minute variations in the domain, $G\vec{n}$ is referred to as a sensitivity function or shape gradient function. Furthermore, the scalar function G is referred to as the shape gradient density function.

Equation (27) is a weak form of the Navier–Stokes equation and the continuity equation (28) is a weak form of the energy equation in the state equation. Equation (29) is a weak form of the Navier–Stokes equation and continuous state equation for the adjoint problem, equation (30) is a weak form of the energy equation in the state equation for the adjoint problem, and (31) is a constraint equation related to the Lagrange multiplier Λ .

The traction method can be applied if the shape gradient function can be evaluated by analyzing $u_i, p, \theta, w_i, q, \xi$, and Λ based on these equations.

Numerical results

A shape identification problem for prescribing the temperature distribution $\theta|_{\Omega_D \times [t_1, t_2]}$ in an unsteady heat convection field was analyzed for the branch channel model shown in Fig.2.

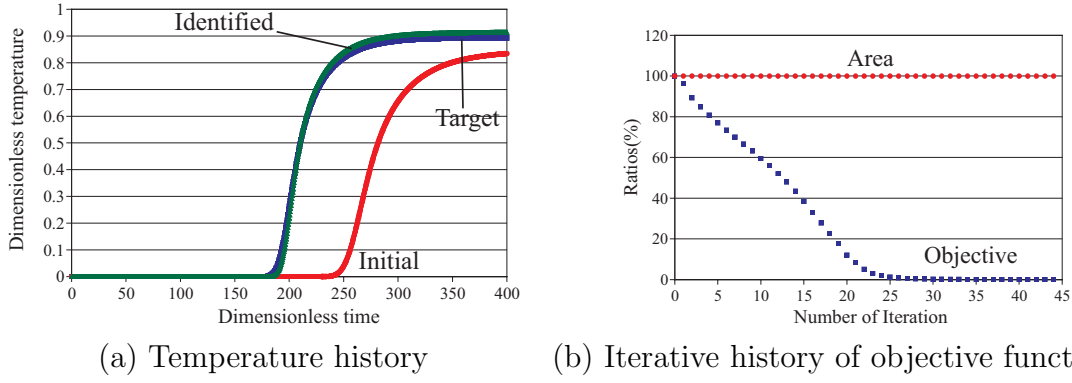


Figure 5: Numerical results: Temperature history in sub-domain Ω_{D1} and iterative history of objective functional

The hot thermal fluid flows in from a boundary Γ_u and flows out from two boundaries Γ_σ . The purpose of shape identification in this analysis is to unify the temperature distribution history near two outlet boundaries Γ_σ during the specified period of time. The temperature distribution history in sub-domain Ω_{D2} was set in the target distribution θ_D in Eq.(21), and the shape identification problem that the temperature distribution history in sub-domain Ω_{D1} agrees with the temperature distribution history in sub-domain Ω_{D2} was analyzed.

The flow boundary conditions included Poiseuille flow on an inlet boundary Γ_u and a natural boundary on two outlet boundaries Γ_σ . The temperature boundary conditions were as follows: $\hat{\theta} = 1$ on the inlet boundary Γ_θ , insulation boundary on the two outlet boundaries Γ_q , and the wall boundaries were heat transfer boundaries Γ_{h1} and Γ_{h2} , with a heat transfer coefficient $\hat{h} = 1$, and an external temperature $\hat{\theta}_f = 0$. The Reynolds number was $Re=100$ and the Prandtl number was $Pr=100$. The initial conditions of the entire domain were set to $\theta_{ini} = 0$ and $u_{i_{ini}} = 0$. The pressure was uniquely set to achieve an average of 0. The time was set to $t_1 = 0$ and $t_2 = T$, and time integration was performed from $t = 0$ to $t = T = 400$ with a $\Delta t = 0.4$ time increment. The two heat transfer boundaries Γ_{h1} of BC and HA were considered to be design boundaries Γ_{design} . Other boundaries were constrained with respect to domain variation. The coefficient β_V , which restrains the size of the domain, was set for $\beta_V = 1$. The heat transfer terms were not considered for the evaluation of the shape gradient density function in Eq.(26).

In this numerical analysis, the flow field velocity \vec{u} , pressure p , temperature θ , adjoint flow velocity \vec{w} , adjoint pressure q , adjoint temperature ξ , and shape updating analysis (velocity field \vec{V}) were all performed using FreeFem++[Ootsuka and Takaishi (2014)], [Hect (2012)]. The mesh and temperature θ at the end time $t = T = 400$ for the initial shape and the identified shape are shown in Fig.3 and Fig.4, respectively. Figure 5(a) shows the temperature history in the initial shape, a target temperature history, and the temperature history for the identified shape in sub-domain Ω_{D1} . Figure 5(b) shows the iterative history for the objective functional. Based on a comparison between Fig.3 and Fig.4, it was observed that the position of branch in the channel moved to upper part in the identified shape so that the temperature history of the two outlet boundaries agreed. In fact, it was confirmed that the temperature history in the sub-domain Ω_{D1} in the identification shape agreed with the target temperature history, and the objective functional approached zero from the result of Fig.5. According to this basic problem, the validity of the proposed method for the shape identification of the unsteady heat convection fields was confirmed.

Conclusions

In the present study, we formulated a shape identification problem in which the square error integral between the actual temperature distributions and the target temperature distributions on the prescribed sub-domains during the specified period of time on unsteady heat convection fields is used as the objective functional. The shape gradient of the shape identification problem was derived theoretically. The validity of the proposed method was confirmed based on the results of a 2D numerical analysis. The present study was supported in part by The OGAWA Science and Technology Foundation in Japan.

References

- [1] Pironneau, O.(1973) On Optimum Profiles in Stokes Flow, *J. Fluid Mechanics* **59**, Part 1, 117-128.
- [2] Pironneau, O.(1974) On Optimum Design in Fluid Mechanics, *J. Fluid Mechanics* **64**, Part 1, 97-110.
- [3] Pironneau, O.(1984) *Optimal Shape Design for Elliptic Systems*, Springer-Verlag.
- [4] Katamine, E. and Azegami, H.(1995) Domain Optimization Analyses of Flow Fields, *Computational Mechanics'95*, S. N. Atluri, G. Yagawa, and T. A. Cruse eds., Springer, **Vol. 1**, 229-234.
- [5] Katamine, E., Azegami, H., Tsubata, T., and Itoh, S.(2005) Solution to Shape Optimization Problems of Viscous Flow Fields, *International Journal of Computational Fluid Dynamics* **19**, 45-51.
- [6] Katamine, E., Nagatomo, Y., and Azegami, H.(2009) Shape optimization of 3D viscous flow fields, *Inverse Problems in Science and Engineering* **17**, No.1, 105-114.
- [7] Katamine, E. and Matsui Y.(2012) Multi-objective shape optimization for drag minimization and lift maximization in low Reynolds number flows, *Theoretical and Applied Mechanics Japan* **61**, 83-92.
- [8] Katamine, E. and Kanai R.(2016) Shape identification of steady-state viscous flow fields to prescribe flow velocity distribution, *Proceedings of the International Conference on Computational Methods*, Berkeley, CA, USA, G.R. Liu and Li, S. Eds, 41-45, ScienTech Publisher.
- [9] Azegami, H., Shimoda, M., Katamine, E., and Wu, Z. C.(1995) A Domain Optimization Technique for Elliptic Boundary Value Problems, *Computer Aided Optimum Design of Structures IV*, Hernandez S. and Brebbia C.A. eds., Computational Mechanics Publications, 51-58.
- [10] Azegami, H., Kaizu, S., Shimoda, M., and Katamine, E.(1997) Irregularity of Shape Optimization Problems and an Improvement Technique, *Computer Aided Optimum Design of Structures V*, Hernandez S. and Brebbia C. A. eds., Computational Mechanics Publications, 309-326.
- [11] Azegami, H.(2000) Solution to Boundary Shape Identification Problems in Elliptic Boundary Value Problems using Shape Derivatives, *Inverse Problems in Engineering Mechanics II*, Tanaka, M. and Dulikravich, G. S. eds., Elsevier, 277-284.
- [12] Ootsuka, K. and Takaishi ,T. (2014) *Finite element analysis using mathematical programming language FreeFem++*, Kyoritsu (in Japanese).
- [13] Hecht, F.(2012) New development in FreeFem++, *Journal of Numerical Mathematics*, **20**, 251-265.

Simulations of Thermal-Hydraulics Two-Phase Flows using Mixture Formulations

†Dia Zeidan¹ and Eric Goncalves²

¹School of Basic Sciences and Humanities, German Jordanian University, Amman, Jordan

²ISAE-ENSMA, Institut Pprime, UPR 3346 CNRS, Poitiers, France

†Corresponding and presenting author: dia.zeidan@gju.edu.jo

ABSTRACT

This short work presents an assessment of the capabilities of a non-equilibrium flow model to solve two-phase flows evolving spatial and temporal discontinuities. The model is based on mixture parameters of state and feature the relative velocity resolution between the two phase systems. It demonstrate an important link between mixture formulations and two-phase flow thermal-hydraulics. Numerical simulations of the two-phase flows are performed using highly accurate and efficient Godunov-type finitevolume solvers. The computational efficiency and ability of both the model equations and the solvers are validated by test cases available in the open literature.

Keywords: Two-phase flow, Non-equilibrium, Mixture model, Godunov methods, Simulation

Introduction and Equations

The various thermal-hydraulics two-phase flow models available in literature are mainly based on either a mixture model, a volume of fluid model or a two-fluid model [3, 5]. The latter is mostly employed in computer codes such as RELAP-5, TRACE and WAHA for the design and safety assessment of nuclear reactors. See for example [7] and references therein. Within individual thermal-hydraulics phenomena where two-phase flow does occur the system of partial differential equations (PDEs) describing such flows is a non-hyperbolic and cannot be written in a conservative form. However, hyperbolicity property can be examined under certain physical restrictions yet the system of equations remains non-conservative [3]. In addition to that, the governing equations cannot handle the relative motion between the two phase system without interphase exchange. Hyperbolic nature of two-phase flow equations is an advantageous property since it allows discontinuities in the solutions related to nuclear thermal-hydraulic systems. These solutions entails the use of different numerical methods of interest within the currently used two-phase flow models context. In this paper, we present recent results of the application of a hyperbolic and conservative system of PDEs to the simulation of thermal-hydraulic mixture of gas and liquid phases [1, 6]. This system enables a fairly straightforward consideration for the relative velocity between the different phases and able to capture shock and expansion waves in two-phase mixture flows. The set of equations is based on three balance equations for mixture mass (ρ), mixture momentum (ρu) and mixture energy (E) as well as a relative

velocity (u_r) balance law written in process without dissipation as [6]:

$$\frac{\partial}{\partial t}(\rho) + \frac{\partial}{\partial x}(\rho u) = 0, \quad (1)$$

$$\frac{\partial}{\partial t}(\rho u) + \frac{\partial}{\partial x}(\rho u^2 + P + \rho c(1 - c)u_r^2) = 0, \quad (2)$$

$$\frac{\partial}{\partial t}(u_r) + \frac{\partial}{\partial x}\left(uu_r + (1 - 2c)\frac{u_r^2}{2} + \psi(P)\right) = 0, \quad (3)$$

$$\frac{\partial}{\partial t}(\rho E) + \frac{\partial}{\partial x}\left(\rho u E + Pu + \rho c(1 - c)u_r\left(uu_r + (1 - 2c)\frac{u_r^2}{2} + \psi(P)\right)\right) = 0, \quad (4)$$

and supplemented by the following gas void fraction (α) and gas mass void fraction (c) balance laws:

$$\frac{\partial}{\partial t}(\rho \alpha) + \frac{\partial}{\partial x}(\rho u \alpha) = 0, \quad (5)$$

$$\frac{\partial}{\partial t}(\rho c) + \frac{\partial}{\partial x}(\rho u c + \rho c(1 - c)u_r) = 0. \quad (6)$$

Here P is the mixture pressure defined as

$$P = \alpha P_2 + (1 - \alpha)P_1,$$

where the relationship between the gas and liquid volumes is shown as follows

$$\alpha + (1 - \alpha) = 1.$$

In addition to that, the ideal compressible equation of state is employed for the gas phase and stiffened gas equation of state is used for the liquid phase. This mixture pressure along with the two different equations of state defines the mixture equation of state. The function $\psi(P)$ is an expression that describes the relationship between the gas and liquid phases through the momentum equations. This is given by

$$\psi(P) = e_2 + \frac{P_2}{\rho_2} - e_1 - \frac{P_1}{\rho_1},$$

where the indices 2 and 1 refer to the gas and liquid phases, respectively.

The properties of system (1)-(4) along with (5)-(6) were studied earlier (see, e.g. [6]). Hyperbolicity of the system of equations were also examined numerically over an entire range of parameters typical of practical applications.

Godunov-Type Resolutions

Finite volume Godunov methods are used to discretise the model equations, (1) to (6), in the form

$$\mathbb{U}_i^{n+1} = \mathbb{U}_i^n - \frac{\Delta t}{\Delta x} \left(\mathbb{F}_{i+\frac{1}{2}}^n - \mathbb{F}_{i-\frac{1}{2}}^n \right), \quad (7)$$

where the interface fluxes, $\mathbb{F}_{i+\frac{1}{2}}^n$ and $\mathbb{F}_{i-\frac{1}{2}}^n$, are calculated using approximate solution of the following Riemann problem

$$\frac{\partial \mathbb{U}}{\partial t} + \frac{\partial \mathbb{F}(\mathbb{U})}{\partial x} = 0, \quad (8)$$

with the initial data defined as

$$\mathbb{U} = \begin{cases} \mathbb{U}_L, & x \leq x_0, \\ \mathbb{U}_R, & x > x_0, \end{cases}$$

where \mathbb{U}_L and \mathbb{U}_R represent the values of gas and liquid properties on a two-phase shock tube at the left and at the right from an interface between the two states at $x = x_0$. Godunov methods of centred-type are considered for the resolution of equation (8). This is due to the large number of unknown variables for the individual phases and the two-phase mixture. Further, the numerical fluxes in (7) are approximated using the Slope-Limited Centered (SLIC) scheme where the solution of the Riemann problem is fully numerical. The SLIC scheme is a second-order in time and space and Total Variation Diminishing (TVD) using any limiter of interest. We refer the reader to [4] for more details of this second-order scheme.

To demonstrate the capability of the equations of the mixture model, the SLIC scheme is employed for the resolution of a well-documented air-water shock-tube problem presented recently in [2]. This problem deals with large physical variations between the two phases where discontinuities appear clearly within thermal-hydraulics, in particular in many types of reactor cores. A tube of 10 m with a diaphragm in the middle which separate the following left and right states are defined as follows at $t = 5$ ms, see [2]:

$$\begin{aligned} (\alpha, \rho_2, u_2, \rho_1, u_1)_L &= (0.25, 57.941, 0.0, 1003.1, 0.0) \quad \text{if } x \leq x_0, \\ (\alpha, \rho_2, u_2, \rho_1, u_1)_R &= (0.1, 25.527, 0.0, 999.85, 0.0) \quad \text{if } x > x_0. \end{aligned}$$

Results are shown in figures 1 at $t = 5$ ms without any source terms consideration. The key reason for employing no source terms effect in the current work is the mixture formulation of the current model equations. Furthermore, unlike other work reported in literature, the results presented in this work are based on compressible liquid where the gas is air. Interestingly, the solution for both phases and mixture flow variables consists of a left rarefaction wave, a contact discontinuity and a right shock wave. The simulation results shown in figure 1 are similar to those presented in [2] except for the middle discontinuities. The relative velocity also is not presented in [2].

Concluding Remarks

A non-equilibrium flow model based on mixture formulation has been proposed for the simulation of two-phase flow thermal-hydraulics. The mixture model is successfully applied to the simulations of a two-phase shock tube with large difference between the gas and liquid flows. It is concluded that this mixture model is capable of solving such two-phase flows without differential closure laws or specific equations of state. The model also provide significant numerical resolutions without any numerical conditions.

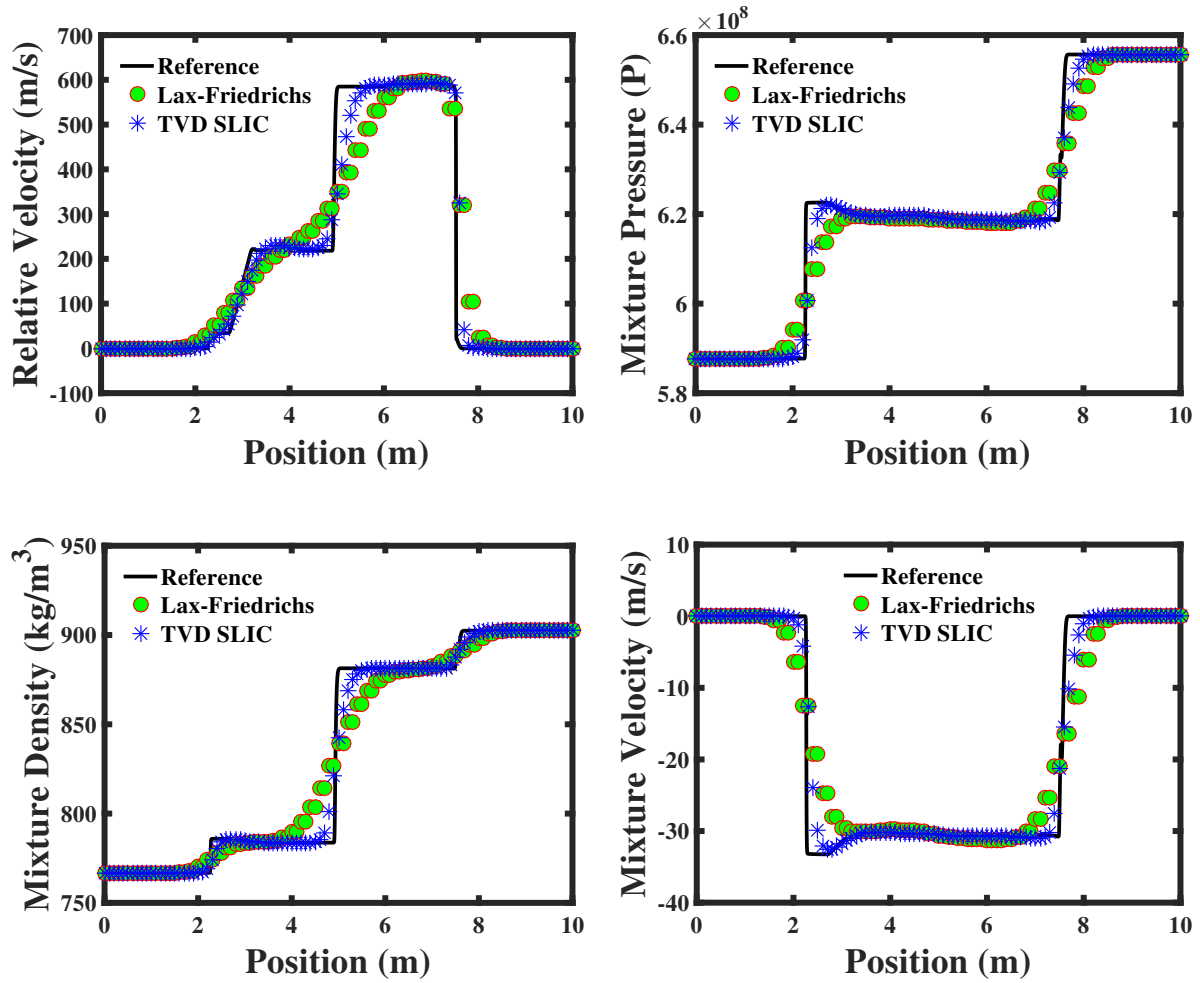


Figure 1: Test 1: Air-water shock-tube problem of [2] at time $t = 5$ ms. The TVD SLIC and first-order Lax-Friedrichs methods are compared with the reference solution results. Coarse meshes, symbols, are provided on 100 cells and very fine meshes of 10000 cells for the solid lines. The waves seen from left to right, repeated left rarefaction and repeated right shock waves separated by a multiple contact discontinuity.

Acknowledgment

This work is supported by Scientific Research Support Fund project No. Bas/1/05/2016, Amman, Jordan, through the German Jordanian University.

References

- [1] Godunov SK and Romenski E, *Elements of Continuum Mechanics and Conservation Laws*, Kluwer Academic/Plenum Publishers, 2003.
- [2] Hu G and Kozłowski T, A Roe-Type Numerical Solver for the Two-Phase Two-Fluid Six-Equation Model with Realistic Equation of State. *Nuclear Engineering and Design* **326**: 354-370, 2018.
- [3] Ishii M and Hibiki T, *Thermo-fluid Dynamics of Two-Phase Flow*, Springer, 2011.
- [4] Toro EF *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*, Springer, 2009.
- [5] Yadigaroglu G and Hewitt GF, *Introduction to Multiphase Flow: Basic Concepts, Applications and*

Modelling, Zurich Lectures on Multiphase Flow, Springer, 2018.

- [6] Zeidan D, Romenski E, Slaouti A and Toro, EF, Numerical Study of Wave Propagation in Compressible Two-Phase Flow, *International Journal for Numerical Methods in Fluids*, **54**: 393-417, 2007.
- [7] Zohuri B and Fathi N, *Thermal-Hydraulic Analysis of Nuclear Reactors*, Springer, 2011.

Numerical investigation of beam-column connections using a new multi-axial-spring model

*Hu Qi¹, †Haishan Guo¹, Kang Liu¹, Lida Tian¹, Geng Jiao¹

¹China State Construction Engineering Corporation, Beijing, China

*Presenting author: 151719077@qq.com

†Corresponding author: 13911507030@139.com

Abstract

Precast concrete moment-resisting frame with hybrid beam-column connections, which is featured by inelastic deformation induced by opening and closing of the interface between precast beam and column, is emphasized in recent years, since it is capable of sustaining design basis earthquake with tiny damage. To explore the opening and closing behavior of the interface of hybrid beam-column connection, a new multi-axial-spring model with only two gap elements whose position and capacity is determined by simple advance section analysis method is proposed. The new multi-axial-spring model, which is obviously with high computational efficiency, is able to tracking accurately the change of compressive zone height of the interface between precast beam and column and count in “beam elongation effects”. The proposed numerical model analysis results are in good agreement with the experimental results.

Keywords: Hybrid beam-column connection, Multi-axial-spring model, Compressive zone height, Precast concrete frame.

Introduction

Compared with conventional monolithic concrete structures, precast concrete systems are advantageous in product quality, cost efficiency, and speed of construction. Precast concrete frame system with hybrid beam-column connections is widely used and commonly accepted in main design codes worldwide, for its capability of sustaining a design level earthquake with limited or negligible damage [1].

In precast frames with hybrid connections, the inelastic deformation demand is concentrated at and provided by opening and closing of the beam-column interface. A variety of numerical models, including macro-models, section analysis model, fiber model, lumped plasticity model, multi-axial-spring model and FEM model(Solid model), have been proposed to predict the behavior of hybrid connections with different level of complexity[1-4]. Fiber and lumped plasticity models are used widely with good accuracy and low computational cost, but for both models, it is difficult to simulation “beam elongation effect”.

Multi-axial-spring model can be used to analyze the rocking behavior and capture the beam elongation effects, such an approach has been widely used to model the hybrid connections. In this model the joint itself was assumed to remain rigid, while beams and columns element are assumed to remain elastic. Inelastic action was supposed to concentrate in the grout(at the beam-column interface) and mild steel bars. Truss elements were used to model the reinforcing steel, while a prestressing element was used to prestress the joint together and 9 gap elements, which are evenly spaced along the height of the interface grout and each with the same area of grout, were used over each side to represent the grout behavior. Further

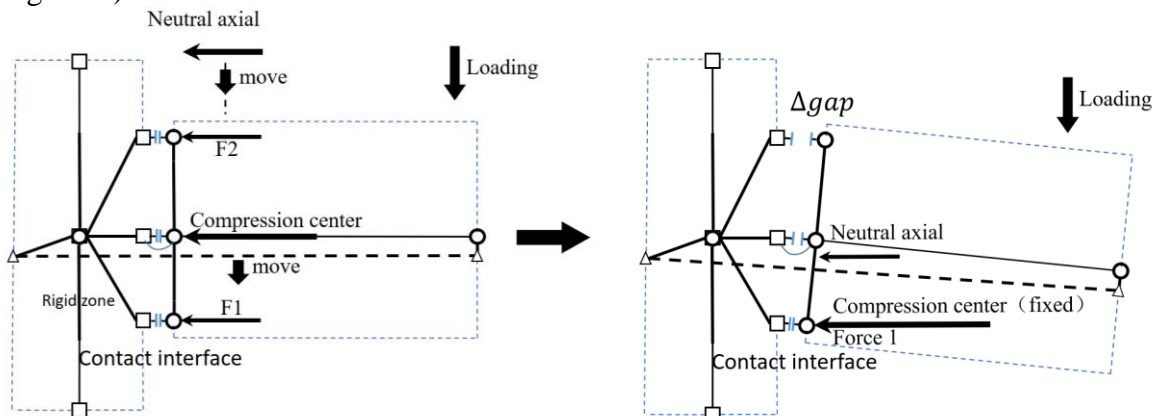
details on the model as well as complete results on beam elongation effects from analytical investigations (using the computer program DRAIN-2DX) on multi-story frame systems can be found in Kim (2002) [5]. Carr, 2004 [6,7] has improved capabilities simulating the contact section interface with an increased number of gap elements. The gap elements is set up for 2 to 10 contact points, which are not evenly spaced along the height of the interface grout and each with the different area of grout. Two different integration schemes, namely, Gauss quadrature and Lobatto integration, can be used to optimize the position of gap elements and calculate their weighting.

In this paper, a new multi-axial-spring model with only two gap elements on half side of interface, which is obviously with better computational efficiency, is proposed and validated. Section analysis method is used to determine the distribution of the two gap elements and a representative area method is used to determine the mechanization property of the gap elements. In this investigation, the proposed model is used to simulate several tests including both hybrid connections and PPEFF beam-column connections [8]. Satisfactory agreement in aspects of overall mechanical property, mild reinforcement strain, prestressed tendon stress and compression center between the analytical and experimental results confirms the validity of the proposed model.

1 Methodology

1.1 Principle

In case of the inelastic deformation of the connections induced by gap opening/closing of the contact interface the model should simulate accurately the local stress and strains in the contact area. The length of inner lever arm (distance from reinforcement to compression center) at the rocking section should be accurately simulated which is proportionate to the capacity of the rocking connection. With increase loading the neutral axis moves from infinity into the contact section and decreases in size with increasing gap opening, finally with the crushing of the edge concrete the compress zone increases in size. Figure 1 shows that two gap elements in compress area (model 2) can simulate the shift of the neutral axis more accurately than one gap element(model 1). In model 2, with suitable location and mechanical property gap elements, the simulate neutral axis firstly move from outside section into the section and the compress center move downward into the scope between F1 and F2, with the loading increasing the F1 begin to decline, the compress center begin to move upward (show in figure1c).



a) Model 1(one spring for compress zone)

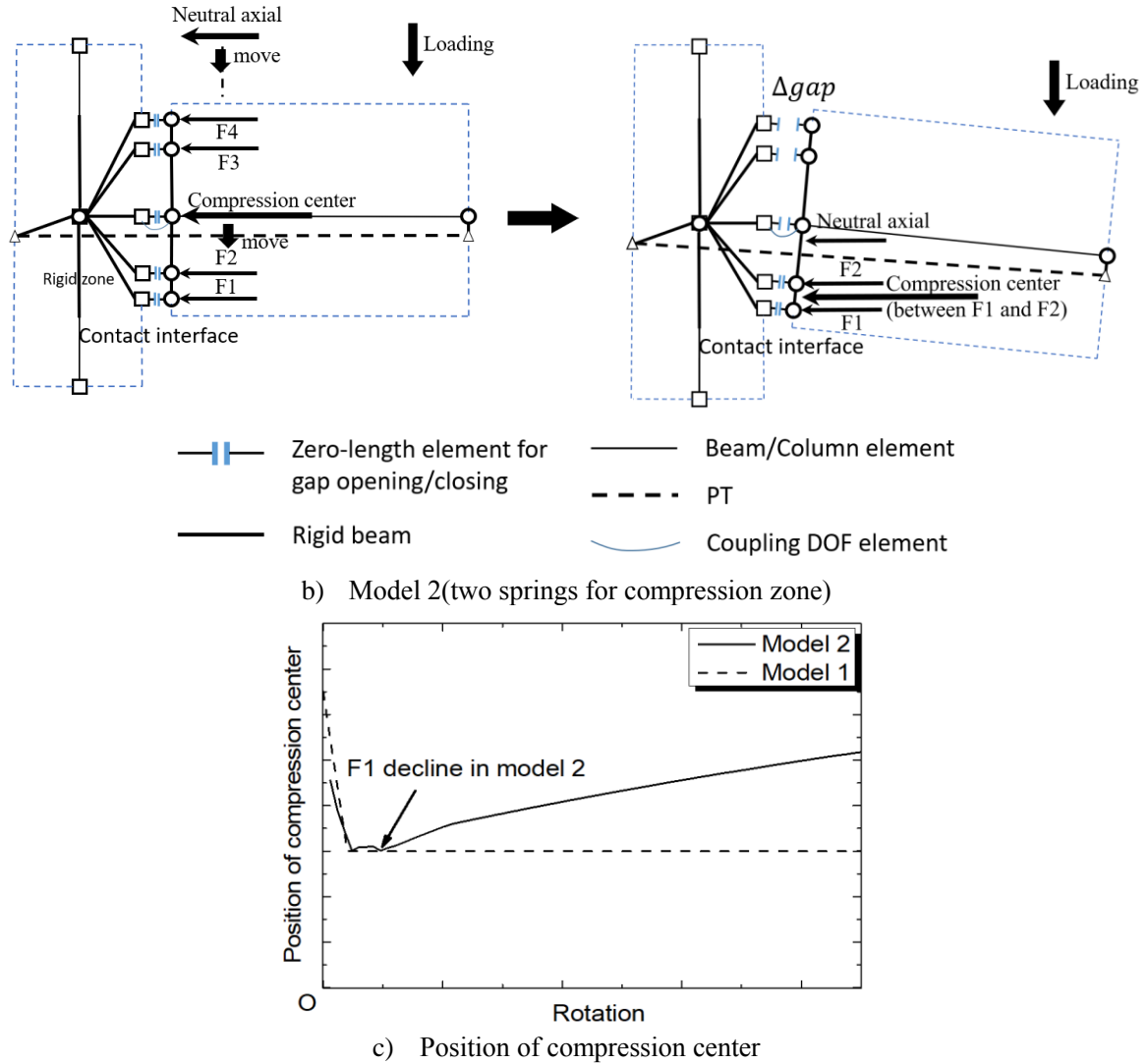


Figure 1 Shift of the compression center(numerical simulated)

1.2 Distribution and mechanical property of the gap elements

This paper proposes a new method to simulate the concrete in compressive zone with only two gap elements, and introduces the technique to determine the position and weighting of the two gap elements. To accurately simulated the shift of the neutral axis, the distribution of the gap elements is calculated by section analysis method [1]. We using the section analysis procedure proposed by reference [1] to simulated the contact interface of the beam column connection, and the position of the neutral axis and compression center can also be calculated (show in figure 2).

If the calculated position of the lowest point of compression center is h_1 (from beam section edge), and the position of the neutral axis is h_2 at maximum rotation. We set the position of gap element 1 at the position of h_1 and the position of gap element 2 at the position of $h_1 + \frac{h_2 - h_1}{2}$ (show in figure 2).

The mechanical property for each gap element is derived from the mechanical property of the scope of concrete it represent. Details of the model will be introduced in section 2.2.

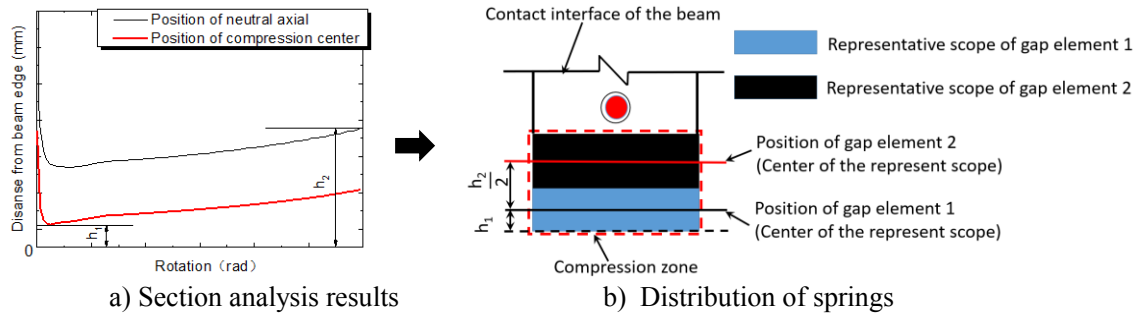


Figure 2 Determine of the position of gap elements

1.3 Preliminary verification

Numerical simulation was conducted using the model presented in section 1.1 and 1.2. The proposed numerical model was developed using the Open System for Earthquake Engineering Simulation (OpenSees) [9]. Figure [3] shows the comparison between the numerical simulation and test result(O-P-Z4) of NISTIR[10], where good agreement is observed both of skeleton curve and hysteretic curve which confirm the validity of the proposed model.

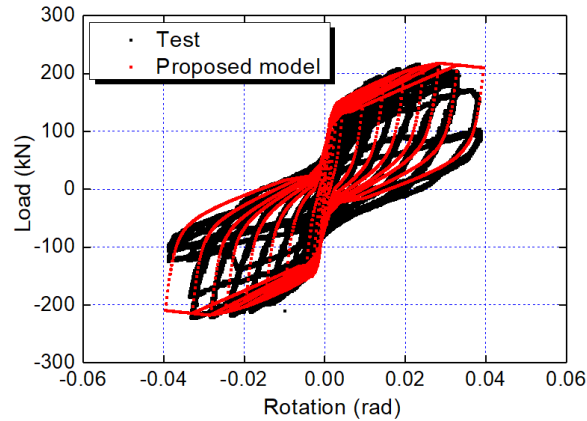


Figure 3 Comparisons of hysteretic behavior between numerical simulation and test

2 Experiment and numerical model

2.1 Experiment

Test were conducted on four specimens of PPEFF beam-column connections [8]: two exterior precast connections and two interior precast connections. The model is shown in Figure 4. Reinforcement information is in Table 1. Detailed information of the specimens are in reference [8]. Quasi-static tests were carried out to investigate the hysteretic behavior, stiffness, bearing capacity and deformation capacity of the beam-to-column dry connection assembled by post-tensioned tendons under slow reversed cycle loading. Experiment result shows PPEFF and hybrid beam-column connections is similarity in working and damage mechanism under low frequency cyclic loading, both of them is featured by inelastic deformation induced by opening and closing of the interface between precast beam and column, the PPEFF joint has slightly better performance [8].

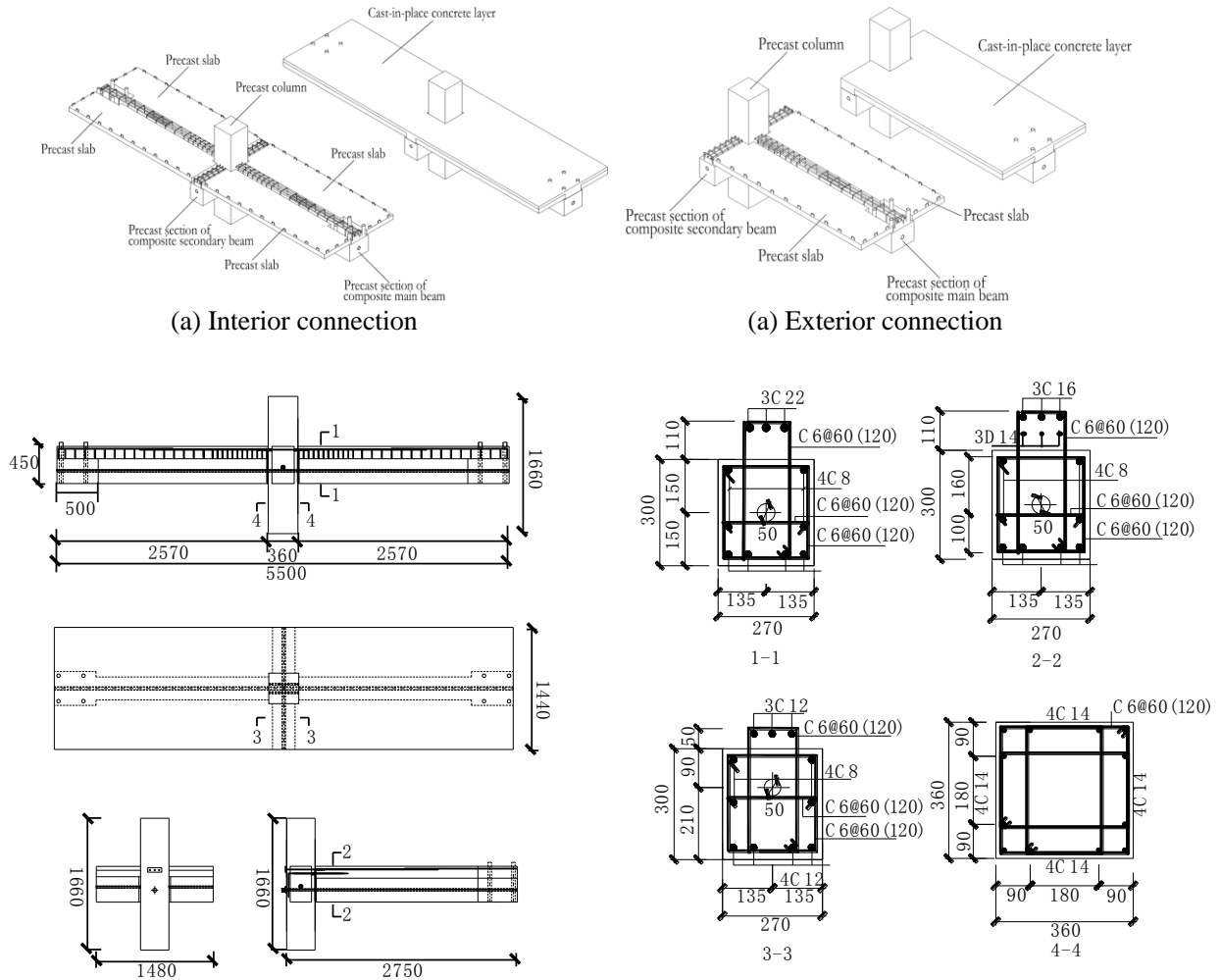


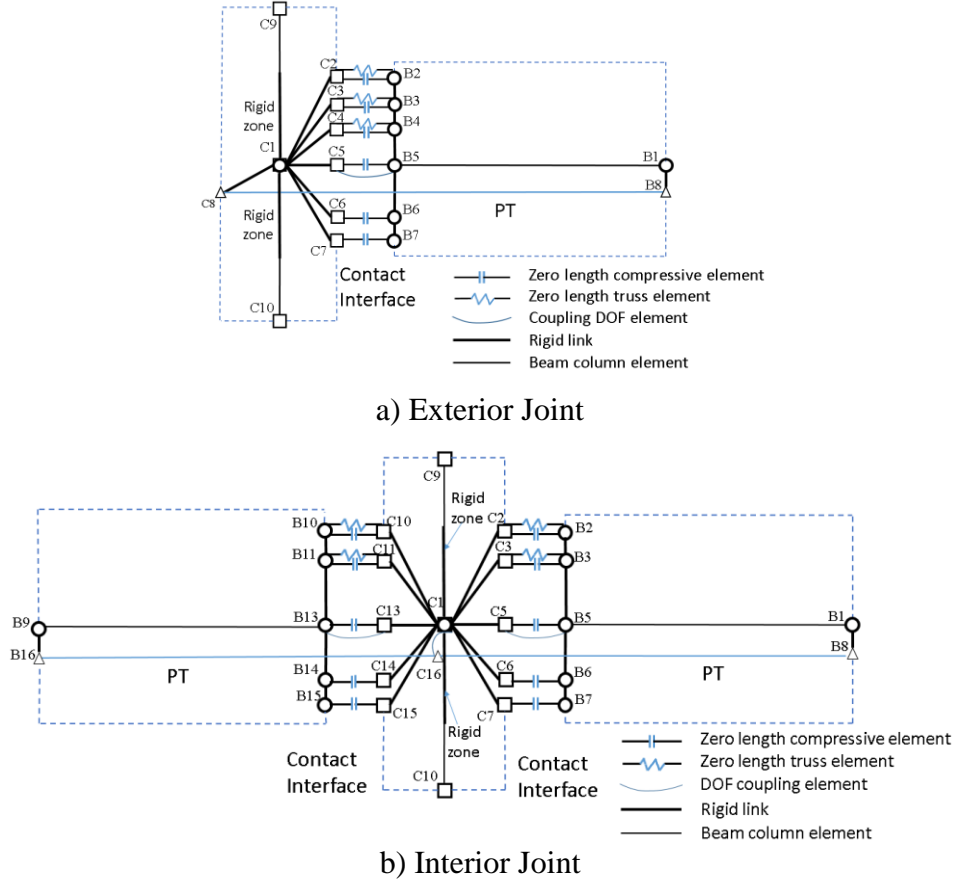
Figure 4 Connection specimens

Table 1 Test specimens

Connection type	Specimen Name	Slab steel	Bending steel	Shear steel	Length of unbonded mild steel(mm)	Unbonded PT
Interior	A2	12C6	3C22	--	360 (inside column)	4Φs15.2
	A3	12C6	3C22	--	120 (outside column) (10% weakened)	4Φs15.2
Exterior	B2	12C6	3C16	3C14	360 (inside column)	4Φs15.2
	B3	12C6	3C16	3C14	120 (outside column) (10% weakened)	4Φs15.2

2.2 Numerical model

The proposed numerical model was developed using OpenSees as shown in Figure 5.


Figure 5 Numerical model

- (1) Zero length elements with compression-only material properties (gap elements) are adopted to simulate the opening/closing behavior of the contact interface.
- (2) Truss element with STEEL02[11] material (assigned initial strain), which takes isotropic hardening and Bauschinger effect into consideration, is used to simulate PT.
- (3) Distribution reinforcement in slab, energy dissipation mild-steel are all simulated by zero length elements using STEEL02 material. CONCRETE02[12] material is used to simulate the concrete, transverse confinement effect is accounted for using the Mander model [13]. The length of truss element in this model (L_{model}) is different from the actual length of unbonded reinforcement in test (L_s), so the material properties and area of the elements require modification [5] using Formula (1).

$$A_1 = A_0 L_{model} / L_s, f_1 = f_0 L_s / L_{model} \quad (1)$$

Where A_1 、 f_1 are the area and yielding strength of the mild-steel in numerical model; A_0 , f_0 are in test.

- (4) Vertical coupling restraint is applied to limit the shear slip between C5 and B5, neglecting shear slip between precast beam and column.
- (5) In the links of C2-B2 and C10-B10, two elements are used: one zero length element (assigned Concrete01 material with compression-only material properties) to simulate contact interface and one zero length element to simulate the behavior of rebar in slab.
- (6) Two zero length elements are used to model the compressive behavior of the contact interface and mild energy dissipation steel for C3-B3 and C11-B11 links.
- (7) Shear reinforcement in exterior connections is simulated by zero length element (assigned Hysteretic material) between C4 and B4. Noting that the bond length of shear reinforcement is $15d$ (d is the diameter of rebar), which is shorter than the anchorage length, the force-

displacement relation of zero length element is derived from bond-slip relation between concrete and rebar.

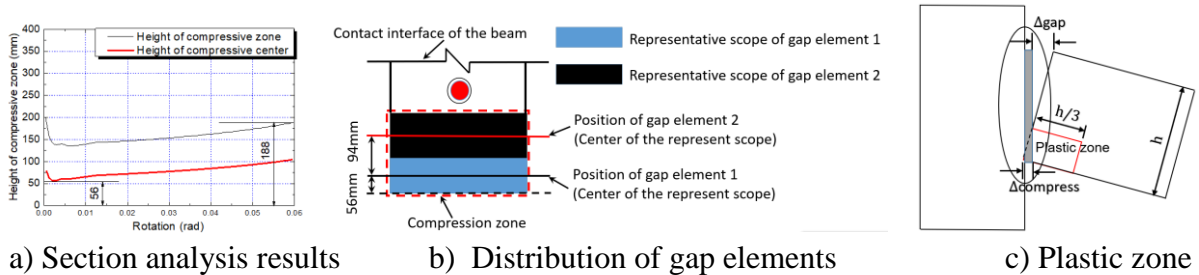
$$F = \pi d * \tau_0 * 15d = 15 \pi d^2 * f(s) \quad (2)$$

Where s stands for slip displacement, $f(s)$ is the bond-slip relation between concrete and rebar[14].

(8) In the links of C6-B6, C14-B14, C7-B7 and C15-B15, zero length elements with compression-only material properties (gap elements) are used to simulated the contact interface. According to section analysis method, the minimum position of compression center is 56mm(take A3 for example), set gap element 1(Figure 6) at this position, which simulates the compressive zone with 112mm high at the bottom of the contact interface. (C6-B6, C14-B14) is designed to simulate the rest concrete when scope of the compressive zone exceeds the representative scope of gap element 1. For A3 the position of gap element 2 is $(188-112)/2+112=150$. Where 188mm is the calculated high of compression zone using section analysis method (show in figure 6). The mechanical property for C6-B6, C14-B14, C7-B7 and C15-B15 is derived from the mechanical property of the scope of representative concrete element with the length of $h/3$ (the length of the plastic zone is $h/3$ [15], where h is the beam section height). The length of truss element in this model (L_{cmodel}) is different from the actual length of plastic zone, so the material properties and area of gap elements require modification [5] using Formula (3).

$$A_c = 3A_{c0}L_{cmodel}/h, f_c(\varepsilon) = f_{c0}(\varepsilon)h/(3L_{cmodel}) \quad (3)$$

Where A_c 、 $f_c(\varepsilon)$ are the area and constitutive relationship of the concrete material in numerical model; A_{c0} 、 $f_{c0}(\varepsilon)$ are in test.



a) Section analysis results b) Distribution of gap elements c) Plastic zone

Figure 6 Position of the gap elements (C6-B6, C14-B14, C7-B7 and C15-B15)

(9) Figure 7 shows that the position of gap elements linking C2-B2, C10-B10, C3-B3 and C11-B11. The mechanical property of the gap elements is derived from the mechanical property of the representative concrete. The rest scope of the interface is simulated by the gap elements linking C5-B5 and C13-B13.

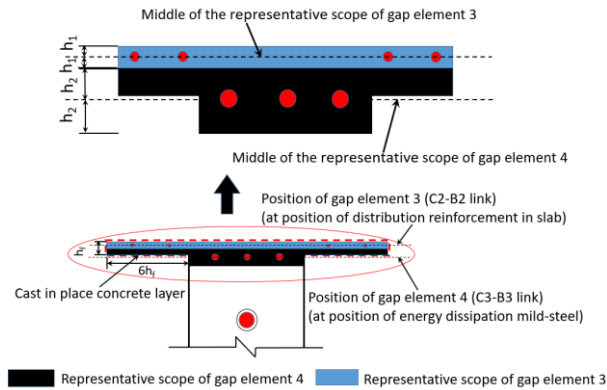


Figure 7 Position of the gap elements (C2-B2, C10-B10, C3-B3, C11-B11)

3 Experimental validation

3.1 Validation of skeleton curves

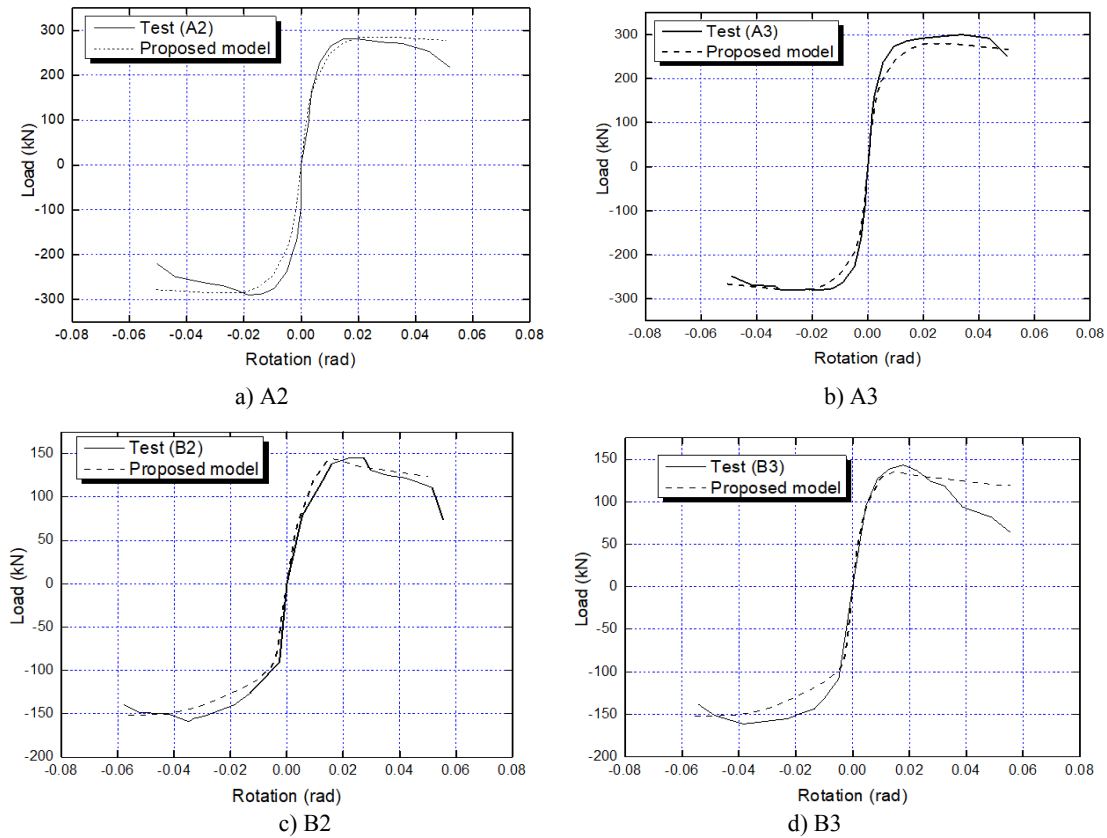


Figure 8 Validation of the proposed model: skeleton curves

Dissymmetry in the skeleton curves of exterior connections, due to asymmetrical reinforcement through the beam-column contact interface, can be recognized in figure 8. For interior connection, the summation of bending capacity contributed by contact interface in either side of the column is symmetry, resulting in symmetrical skeleton and hysteretic curves. The experimental and simulation results of B2 and B3 show that, the contact interface remains elastic before cracking and the section stiffness declines after cracking. Under positive loading (mild steel in tension), the cracking moment is small because the prestressing tendons are at the mid lower portion of beam section, and due to mild steel and distribution steel in the slab, gradual instead of sharp stiffness decrease is observed before yielding of mild steel. Under negative loading (mild steel in compression), however, the cracking moment at the interface is larger and section stiffness decreases greatly as the neutral axis goes upward after cracking.

Under positive loading, the mild steel go gradually from elastic state into yielding and hardening stage as the moment rising, until the exterior concrete in compressive zone reaches its ultimate strength. Afterwards, the concrete at the edge of the contact interface crushed and the compression center moved upward, leading to decent of the bearing capacity. Under negative loading, the bearing capacity of the connection kept increasing under large deformation as the prestress tendon remains elastic and the slight damage of the compressive concrete due to slab. The connection bearing capacity decreased slightly after crushing failure of concrete.

3.2 Validation of hysteretic behavior

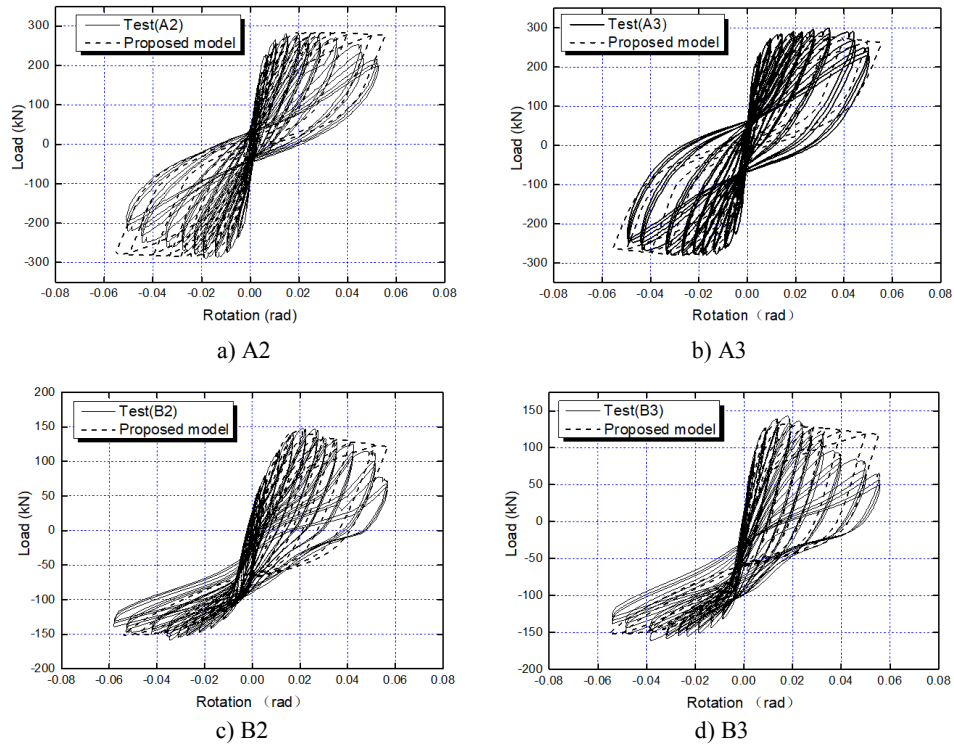
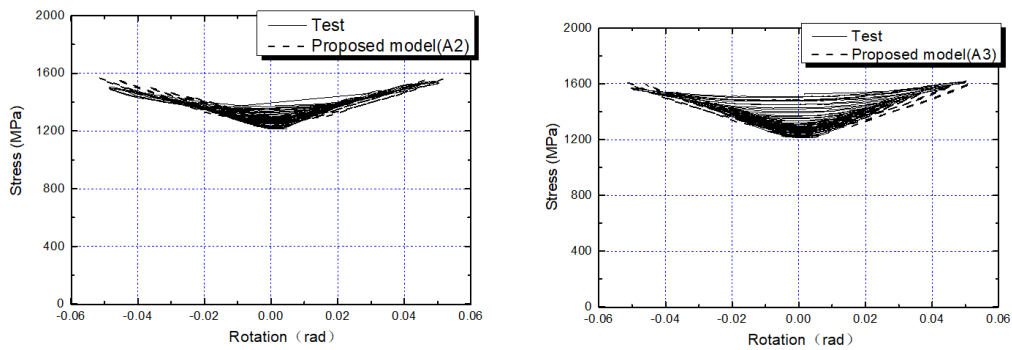


Figure 9 Validation of the proposed model: hysteretic behavior

Figure 9 shows the comparisons between the numerical simulation and test results, where the self-centering behavior and energy dissipation behavior are accurately replicated, implying the effectiveness of the proposed model for the behavior and mechanism of contact interface of the connections.

3.3 Validation of the stress of PT

Figure 10 shows the comparison of simulation and experimental results on stress of PT, where good agreement is observed.



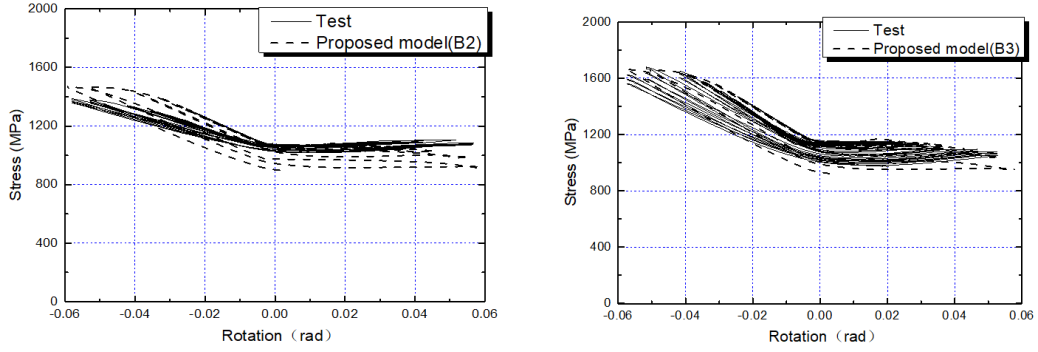


Figure 10 Validation of the proposed model: stress of prestressed tendon

3.4 Validation of the shift of compression center at contact interface

To further testify the accuracy of this model, the positions of compression center in simulation and test are compared for A3 and B3.

Using test data, including moment of the interface, stress of prestressed tendon and rotation of the interface, the position of compression center and stress of mild steel are calculated according to force equilibrium and moment equilibrium on the contact surface. The whole procedure can be summarized as follows.

① Impose a rotation θ , and get the moment M_0 based on test results

② Guess an initial neutral axis position c ,

Calculate $\varepsilon_c = f(c)$, Evaluate corresponding compression force C , and calculate the position of compression center.

③ Calculate tensile force of prestressed tendon F_p according to section rotation and measured

④ Calculate tensile force of mild steel based on Section Equilibrium.

$$F_y = F_0 - F_p$$

⑤ Evaluate moment capacity.

$$M = F_y(C_0 + C_y) + F_p(C_0 + C_p)$$

C_y , C_p are the distances from the neutral axis to the mild steel and prestressed tendons, respectively.

⑥ If $M = M_0$

Yes, Go to Step 7

No, Revise neutral axis position and go to Step 2

⑦ End

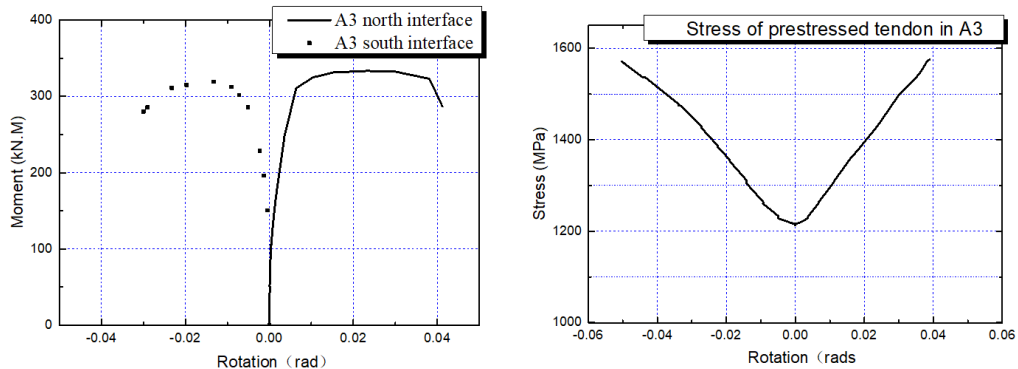


Figure 11 Rotation of contact interface and stress of PT (A3 test)

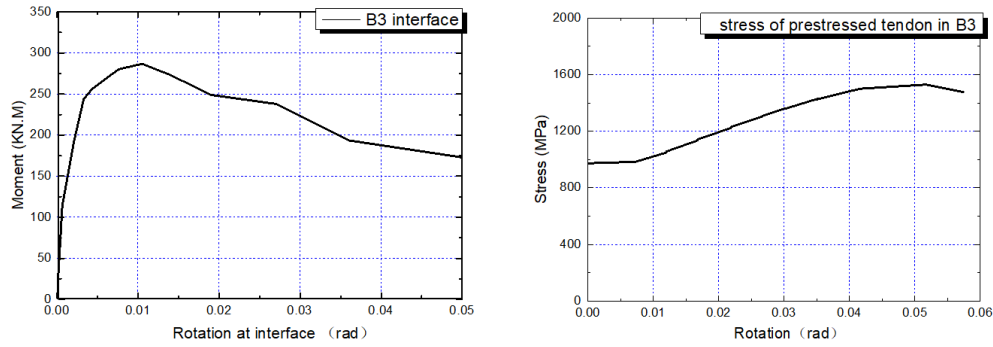


Figure 12 Rotation of contact interface and stress of PT (B3 test)

The shift of compression center calculated from the above procedure is shown in figure 13, using the test data in figure 11 and 12. The simulated compression center using proposed model is also given in figure 13, which shows good agreement of the simulation and test results.

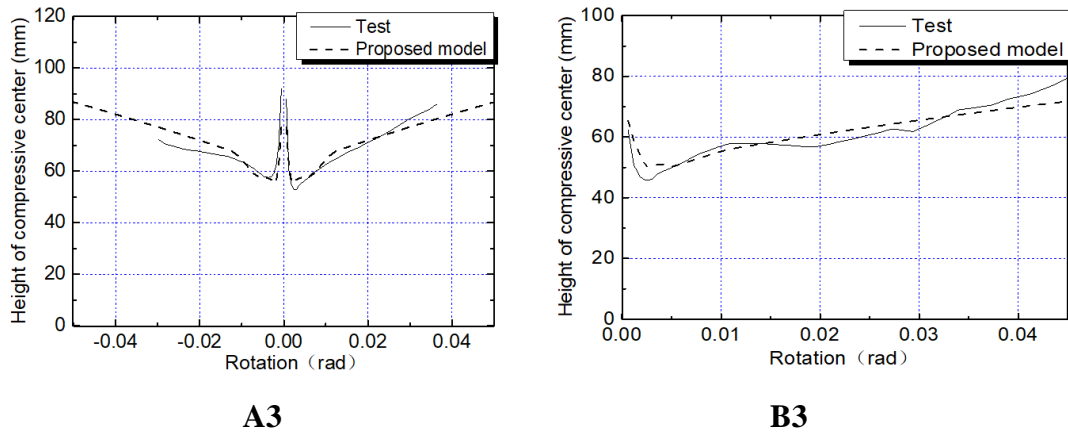


Figure 13 Comparison of compression center position between simulation and test

4 Study of gap elements distribution

Spieth 2004 proposed a multi-spring contact element using Lobatto Integration and Gauss integration to calculate the position of the springs and their weighting. In this paper the effect of different distribution of the springs is investigated. Figure 14 and 15 shows the simulation results by the models using Lobatto integration, Gauss integration and the proposed method in this paper to calculate the position of the gap elements (gap element 1 and 2 in figure 6) and their weighting (shown in Table 2). Figure 16-18 show comparison of PT stress, steel stress and compression center position between proposed model and the model using Gauss integration.

Table 2 Position and weighting of the compression gap elements using numerical integration

Spring	Proposed model (A3)		Proposed model (B3)		Lobatto integration		Gauss integration	
	Position	weighting	Position	weighting	Position	weighting	Position	weighting
1	0.33	0.34	0.44	0.23	0.447	0.83	0.34	0.65
2	0.75	0.50	0.78	0.44	1	0.17	0.86	0.35

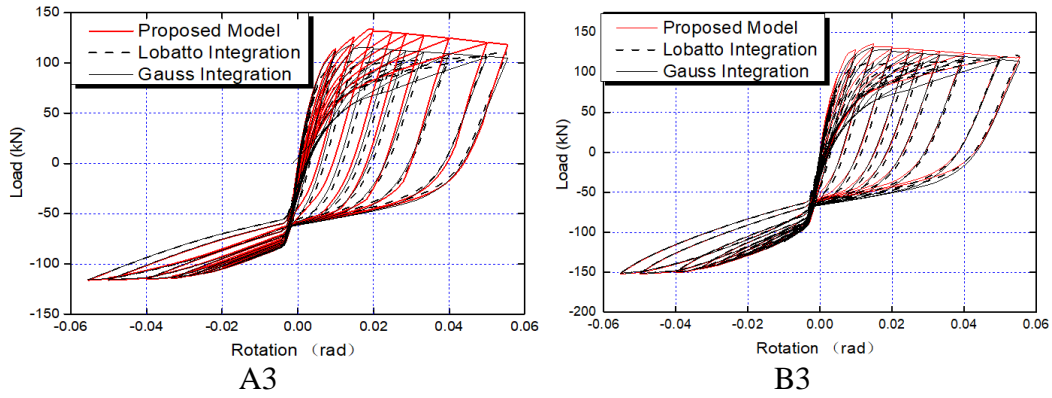


Figure 14 Comparison of the hysteretic behavior of A3(Beam end loading) and B3

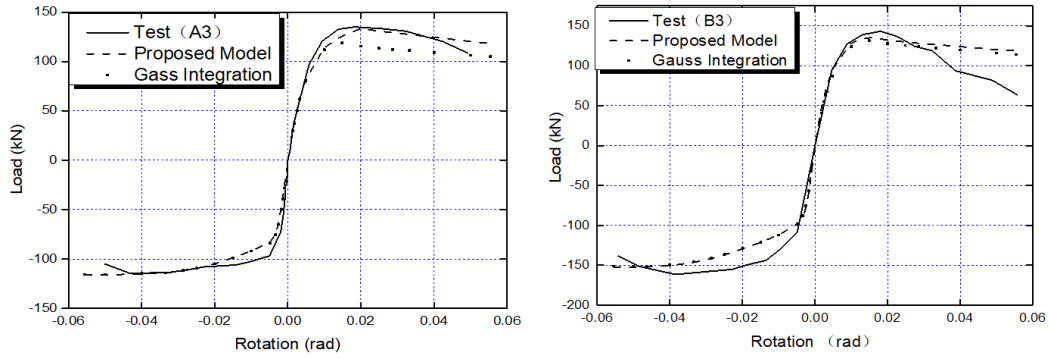


Figure 15 Comparison of the skeleton curves

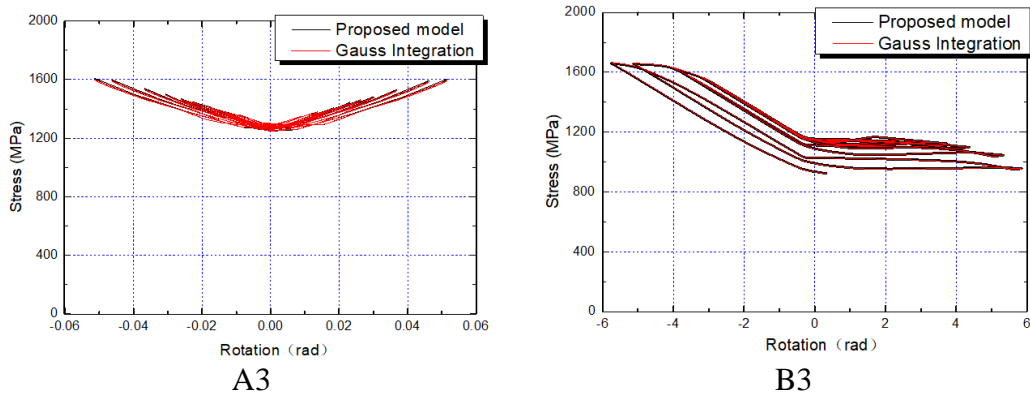


Figure 16 Comparison of stress of prestressed tendon of A3 and B3

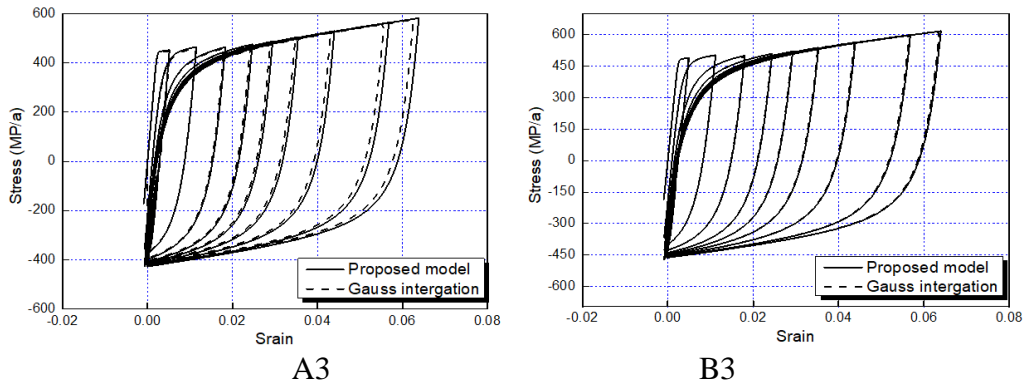


Figure 17 Comparison of steel stress of A3 and B3

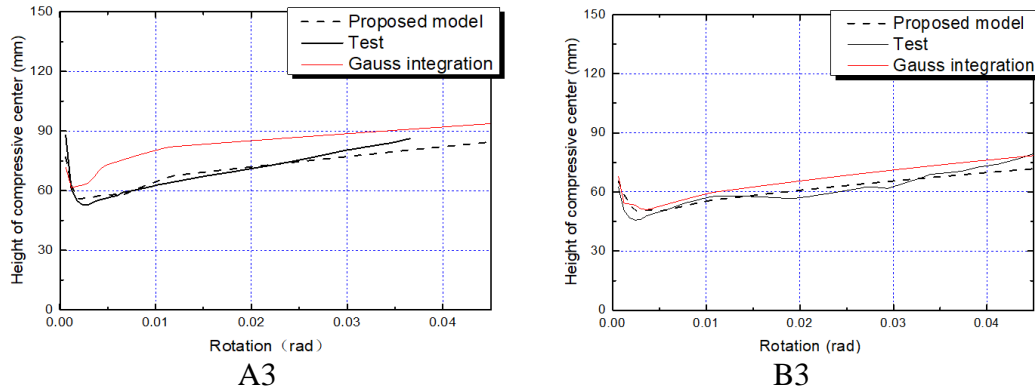


Figure 18 Comparison of compression center of A3 and B3

The method proposed in this paper to determine the position and weight of gap elements takes into consideration the variation of compressive zone, calculated from section analysis procedure, caused by prestressed tendon's area and stress state, area of mild steel, geometry of the contact section, etc. Gauss integration and Lobatto integration, which merely calculates the position and weight of integration point mathematically but without clear physical significance, however, diverges from the actual situation and the simulation results are inferior to that of the proposed method.

5 Conclusions

Based on the comparison of the presented analysis and tests results, conclusions are drawn as follows:

- (1) The new multi-axial-spring model, which is obviously with higher computational efficiency, is able to tracking accurately the change of compressive zone height of the interface between precast beam and column and count in "beam elongation effects".
- (2) The new multi-axial-spring model is validated by several low-cycle loading tests including both hybrid connections and PPEFF beam-column connections and satisfactory agreements in aspects of skeleton curve, hysteretic curve, prestressed tendon stress and compression center are obtained between the analytical and experimental results.
- (3) The new multi-axial-spring model is obviously with better computational efficiency than previous multi-spring models using Lobatto Integration and Gauss Integration since the new one is more accurate with the same number gap elements.

References

- [1] STEFANOPAMPANIN, Priestley, M. J. N., & Sritharan, S. (2001). Analytical modelling of the seismic behaviour of precast concrete frames designed with ductile connections. *Journal of Earthquake Engineering*, 5(3), 329-367.
- [2] El-Sheikh, M., Sause, R., Pessiki, S. and Lu, L.-W. [1999]. seismic behavior of unbonded post-tensioned precast concrete frames", *PCI J.* 44(3).
- [3] Palermo, A., Pampanin, S. and Carr, A. J. [2005b] "Efficiency of simplified alternative modelling approaches to predict the seismic response of precast concrete hybrid systems," fib symposium, Budapest.
- [4] Fenwick, R. C., & Megget, L. M. (1993). Elongation and load deflection characteristics of reinforced concrete members containing plastic hinges. *Bulletin of the New Zealand National Society for Earthquake Engineering*, 26(1), 28-41.
- [5] Kim, J. (2002). Behavior of hybrid frames under seismic loading. *Dissertation Abstracts International, Volume: 63-08, Section: B, page: 3830.;Chairs: John F. Stanto.*

- [6] Carr, A. J. [2004]. RUAUMOKO (Inelastic Dynamic Analysis Software), Department of Civil Engineering, University of Canterbury, Christchurch, New Zealand.
- [7] Spieth, H. A., Carr, A. J., Murahidy, A. G., Arnolds, D., Davies, M., & Mander, J. B. (2004). Modelling of post-tensioned pre-cast reinforced concrete frame structures with rocking column connections. *Journal of Andrology*, 15(5), 468–478.
- [8] Guo Haishan, Liu Kang, Qi Hu, etc. Study on A Novel Precast Unbonded Post-tensioned Concrete Beam-Column Joint[C]. Proceedings of 264th China engineering science and technology forum and the tenth national symposium on earthquake prevention and mitigation engineering. Chengdu, CHINA.2018: 266-276.(in Chinese)).
- [9] McKenna, F., & Fenves, G. L. (2013). Open system for earthquake engineering simulation (opensees). pacific earthquake engineering research center. *University of California*.
- [10] US Department of Commerce, NIST. Performance of 1/3-Scale Model Precast Concrete Beam-Column Connections Subjected to Cyclic Inelastic Loads. Report No. 4[J]. NIST Interagency/Internal Report (NISTIR) - 5436.
- [11] Filippou, F. C., Popov, E. P., Bertero, V. V. (1983). "Effects of Bond Deterioration on Hysteretic Behavior of Reinforced Concrete Joints". Report EERC 83-19, Earthquake Engineering Research Center, University of California, Berkeley.
- [12] Mohd Hisham Mohd Yassin, "Nonlinear Analysis of Prestressed Concrete Structures under Monotonic and Cyclic Loads" PhD dissertation | University of California Berkeley 1994
- [13] Mander, J. B., Priestley, M. J. N., & Park, R. (1988). Theoretical stress-strain model for confined concrete. *Journal of Structural Engineering*, 114(8), 1804-1826.
- [14] 2010. (GB50010-2010) Code for design of concrete structures[S].Beijing: China Architecture & Building Press, 1986.(in Chinese)).
- [15] Chong Xun, Meng Shaoping, Pan Qijian. (2012). Experimental study on seismic performance of Post-Tensioned Prestressed Precast Concrete beam- column assemblages. *CHINA CIVIL ENGINEERING JOURNAL* (12), 38-44. (in Chinese)).

Predicting Plaque Area Increase and Plaque Burden Increase Using Patient-Specific Fluid-Structure-Interaction Models Based on IVUS and OCT Images with Follow-Up

†Xiaoya Guo,¹ Dalin Tang,^{1*,2} David Molony,³ Chun Yang,² Habib Samady³, Jie Zheng⁴, Gary S. Mintz⁵, Akiko Maehara⁵, Jian Zhu⁶, Genshan Ma⁶, Don P. Giddens^{3,7}

¹Department of Mathematics, Southeast University, Nanjing, 210096, China

²Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609 USA

³Department of Medicine, Emory University School of Medicine, Atlanta, GA, 30307, USA

⁴Mallinckrodt Institute of Radiology, Washington University, St. Louis, MO, 63110, USA

⁵The Cardiovascular Research Foundation, Columbia University, New York, NY 10022, USA

⁶Department of Cardiology, Zhongda Hospital, Southeast University, Nanjing, 210009, China

⁷The Wallace H. Coulter Department of Biomedical Engineering,

Georgia Institute of Technology, Atlanta, GA, 30332 USA

† Presenting author: Xiaoya Guo, Southeast University, Nanjing China;

* Corresponding authors: Dalin Tang, Southeast University, Nanjing China, email: dtang@wpi.edu, fax number: 508 831-5824;

Abstract

Atherosclerotic plaque progression may be associated with morphological and mechanical factors. Plaque morphological information on IVUS and OCT images could complement each other and provide for more accurate plaque morphology. Fluid-structure interaction (FSI) models combining intravascular ultrasound (IVUS) and optical coherence tomography (OCT) were constructed to obtain mechanical risk factors. Accuracy and completeness of imaging and advanced modeling lead to accurate plaque progression predictions.

In vivo IVUS and OCT coronary plaque data at baseline and follow-up were acquired from left circumflex coronary and right coronary artery of one patient. Co-registration and segmentation of baseline and follow-up IVUS and OCT images were performed by experts. Baseline and follow-up 3D FSI models based on IVUS and OCT were constructed to obtain plaque stress, strain and flow shear stress for plaque progression prediction. Nine factors including 6 morphological factors and 3 mechanical factors were selected for each slice. Plaque area increase (PAI) and plaque burden increase (PBI) were chosen to measure plaque progression. All possible combinations of nine factors were fed to a generalized linear mixed model for PAI and PBI prediction and quantification of their prediction accuracies.

Prediction accuracy is defined as the sum of sensitivity and specificity. The optimized predictor combining 9 factors gave the best prediction for PAI with accuracy=1.7087 (sensitivity: 0.8679; specificity: 0.8408). Plaque wall strain (PWSn) was the best single-factor predictor for PAI with accuracy=1.5918 (sensitivity: 0.7143; specificity 0.8776). A combination of average cap thickness, calcification area, plaque area, plaque wall stress and plaque wall strain gave the best prediction for PBI with accuracy=1.8698 (sensitivity: 0.8892; specificity: 0.9806). PWSn was the best single-factor predictor with accuracy=1.8461 (sensitivity: 0.8784; specificity 0.9677).

Combining morphological and mechanical risk factors may lead to more accurate progression prediction, compared to the predictions using single factors. IVUS+OCT formed basis for accurate data for morphological and mechanical factors.

Keywords: Vulnerable plaque; OCT; IVUS; plaque progression; patient-specific FSI model.

1. Introduction

Atherosclerotic plaque progression and rupture may be associated with complicated factors including plaque morphology, material properties, mechanical factors, cell and genomic activities, etc. [1-2]. In original study for plaque progression, research groups performed large-scale studies based on histologic sections from autopsy to investigate plaque remodeling and vulnerability [3-5]. For in vivo studies, Mintz et al., Nakamura et al. and among others used medical imaging such as intravascular ultrasound (IVUS) and angiography and indicated that plaque area and lumen cross-sectional area were closely related to plaque progression [6-8]. The limitation of these earlier research is that it only gave one-time plaque data and did not reflect plaque progression. Follow-up studies with advanced medical images can better track the plaque progression. Several research groups used plaque area and plaque burden as the measurement of plaque progression respectively, and investigated the correlation between plaque progression and wall shear stress (WSS) from follow-up data [9-11]. Plaque progression is influenced by the interaction of various morphological factors and mechanical factors including structural and flow conditions, and its mechanism has not been fully understood [2,12]. Wang et al. used fluid-structure interaction (FSI) models with follow-up VH-IVUS data and showed that the combination of morphological and biomechanical factors could improve prediction accuracy, compared to predictions using only morphological features [13].

In recent years, Optical Coherence Tomography (OCT) with high resolution (15-20 μm) gradually became a powerful tool in identifying thin fibrous cap (cap thickness < 65 μm), inflammation and calcification [14-16]. Uemura et al. used 7-month follow-up OCT data from 53 patients to study the relation between morphological characteristics and plaque progression, and found a high correlation between thin-cap fibroatheroma and subsequent luminal progression [17]. One limitation of OCT is its limited penetration: OCT cannot “see” through the whole vessel wall. Plaque morphological information on IVUS and OCT images could complement each other and provide more complete and accurate plaque morphology, especially more accurate fibrous cap thickness measurements [18]. Since accurate cap thickness and stress/strain quantifications are of fundamental importance for vulnerable plaque research, Guo et al. proposed a modeling method to combine IVUS and OCT for more accurate patient-specific coronary morphology and stress/strain calculations [19]. This IVUS+OCT-based modeling approach may provide the basis leading to better plaque stress/strain calculations and progression and vulnerability predictions.

In this paper, patient follow-up IVUS and OCT data were acquired and FSI models were constructed to better quantify human coronary atherosclerotic plaque morphology (especially cap thickness) and plaque stress/strain conditions. Nine selected plaque morphological and mechanical factors and all possible combination were used into generalized linear mixed models (GLMM) to predict plaque progression measured by plaque area increase (PAI) and plaque burden increase (PBI).

2. Data, Models and Methods

2.1 IVUS and OCT data acquisition and image processing

Baseline and one-time follow-up in vivo IVUS/OCT/Angiography data were acquired from two arteries (left circumflex coronary artery and right coronary artery) of one participant (female, 80 age) at Emory University with informed content obtained. IVUS catheterization (Boston Scientific/SCIMED Corp.) was performed with an automatic pullback speed of 0.5mm/s. Following IVUS image acquisition, OCT catheterization (St. Jude, Minnesota, MN) was also performed with an automatic pullback speed of 20mm/s. The IVUS/OCT/Angiography data at baseline (Time 1, T1) and follow-up (Time 2, T2) were acquired uniformly according to the above descriptions. As IVUS and OCT images at T1 and T2 were recorded using different

catheter in four pullbacks, they must be registered both longitudinally and circumferentially in order to be used for modeling. Co-registration and segmentation of paired IVUS and OCT were performed by experts. Paired IVUS and OCT were merged to obtained IVUS+OCT slices, with IVUS providing whole vessel (lumen and out-boundary) contours, and OCT provide more accurate cap thickness and plaque component contours. All image slices were segmented into 3 plaque tissue types: Fibrotic plus Fibro-fatty, Necrotic core, and Dense Calcium.

2.2 The 3D FSI model and Mooney-Rivlin model for material properties

Aortic pressure (136/88 mmHg) obtained by catheter were used as inlet pressure conditions. The modeling procedures, assumptions, governing equations and boundary conditions for the 3D FSI model can be found in our previous publication [20]. Atherosclerotic vessels were stiffer than healthy vessels, axial shrinkage was set at 5% in our models. The anisotropic Mooney-Rivlin model was used for the vessel tissue. Its strain energy density function is:

$$W = c_1(I_1 - 3) + c_2(I_2 - 3) + D_1 [\exp(D_2(I_1 - 3)) - 1] + (K_1/K_2) \{ \exp[K_2(I_4 - 1)^2] - 1 \}, \quad (1)$$

$$I_1 = \sum C_{ii}, \quad I_2 = \frac{1}{2} [I_1^2 - C_{ij}C_{ij}], \quad (2)$$

where I_1 and I_2 are the first and second invariants of right Cauchy-Green deformation tensor \mathbf{C} defined as $\mathbf{C} = [C_{ij}] = \mathbf{X}^T \mathbf{X}$, $\mathbf{X} = [X_{ij}] = [\partial x_i / \partial a_j]$, (x_i) is current position, (a_i) is original position, $I_4 = C_{ij}(\mathbf{n}_c)_i(\mathbf{n}_c)_j$, \mathbf{n}_c is the unit vector in the circumferential direction of the vessel, c_1 , c_2 , D_1 , D_2 , K_1 and K_2 are material parameters [20,21] whose values were determined using in vivo IVUS data [22]: $c_1 = -262.6$ kPa, $c_2 = 22.9$, $D_1 = 125.9$ kPa, $D_2 = 2$, $K_1 = 7.19$ kPa, $K_2 = 23.5$.

Plaque components were assumed isotropic and the isotropic Mooney-Rivlin material model was used to describe their material properties.

$$W_{iso} = c_1(I_1 - 3) + c_2(I_2 - 3) + D_1 [\exp(D_2(I_1 - 3)) - 1], \quad (3)$$

The material parameters: Lipid: $c_1 = 0.5$ kPa, $c_2 = 0$, $D_1 = 0.5$ kPa, $D_2 = 1.5$. Calcification: $c_1 = 92$ kPa, $c_2 = 0$, $D_1 = 36$ kPa and $D_2 = 2$ [22]. The models were solved by a commercial finite element software ADINA (Adina R & D, Watertown, MA, USA) following established procedures [20].

2.3 Data Extraction and Plaque Measurements

The contours segmented from IVUS+OCT slices were used to make FSI models and obtain morphological and mechanical measurements for analysis. Each slice contained 100 evenly-spaced nodal points taken on the lumen, each lumen nodal point was connected to a corresponding point on vessel out-boundary. If the connecting line pass through lipid region, the distance between lumen nodal point and first time the line meets the lipid is defined cap thickness. The average of cap thickness from one slice was defined as average cap thickness (Ave. CT). The area of lipid or calcification (denoted as Ca) in slice was recorded as lipid or Ca area. The area in lumen contour was denoted lumen area (LA). The area between lumen and out-boundary was defined as plaque area (PA). The plaque burden (PB) was defined as the ratio of PA to the sum of PA and LA. Plaque wall stress (PWS) and plaque wall strain (PWSn), WSS were extracted from 3D FSI model solution at 100 lumen nodal points of all slices. Therefore, morphological and mechanical factors used in this study included Ave. CT, lipid/Ca area, LA, PA and PB. WSS, PWS and PWSn.

2.4 Plaque Progression Classification and Prediction

For all paired slices, plaque area increase (PAI) and plaque burden increase (PBI) from T1 to T2 were selected to measure plaque progression:

$$\text{Plaque Area Increase (PAI)} = (\text{PA at T2}) - (\text{PA at T1}). \quad (4)$$

$$\text{Plaque Burden Increase (PBI)} = (\text{PB at T2}) - (\text{PB at T1}). \quad (5)$$

In this work, plaque progression was classified into two types in this work. For a given slice, if $\text{PAI} > 0$, then this slice was labeled 1. If $\text{PAI} \leq 0$, this slice was labeled -1. Slice labeling for PBI was done in the same way as PAI. Generalized linear mixed models (GLMM) were used

to calculate the predictive sensitivity and specificity of all possible combinations of the 9 risk factors (predictors) and find the best combination for plaque progression prediction. Details about GLMM can be found in [23]. A five-fold cross-validation procedure was employed in all 105 slices from two arteries for training and testing sets. For the reliability of results, 100 times repeated experiment were performed. Prediction accuracy is defined as the sum of sensitivity and specificity (Sen+Spe). The receiver operating characteristic curve (ROC) and the area under of the ROC curve were also given to compare the prediction accuracy.

3. Results

3.1 Plaque progression prediction using one single risk factor

For the nine morphological and mechanical factors, each factor was tested to find the best single risk factor for plaque progression prediction. Prediction results from different single factor and plaque progression measurement were compared. According to the sum of sensitivity and specificity, the best five single risk factors for PAI and PBI are showed in Table 1. PWSn was the best predictor for both PAI and PBI. The sum of sensitivity and specificity are 1.5918 and 1.8461 respectively. The ROC curves of PWSn using PAI and PBI were shown in Figure 1. The AUC values were 0.8126 and 0.9529, respectively.

Table 1. Prediction sensitivity and specificity using one single factor using PAI and PBI.

	Predictors	ProbCutoffs	Sensitivity	Specificity	Sen+Spe	AUC
PAI	PWSn	0.5110	0.7143	0.8776	1.5918	0.8126
	PWS	0.5042	0.6679	0.7592	1.4270	0.7477
	Ca Area	0.4988	0.4964	0.9184	1.4148	0.6874
	Ave. CT	0.4606	0.5786	0.7306	1.3092	0.6379
	LA	0.4974	0.5679	0.6939	1.2617	0.6336
PBI	PWSn	0.8304	0.8784	0.9677	1.8461	0.9529
	LA	0.8049	0.6432	0.9935	1.6368	0.8022
	Lipid Area	0.6345	0.7108	0.9032	1.6140	0.8168
	PB	0.7487	0.6838	0.9032	1.5870	0.7606
	FSS	0.7656	0.6811	0.9032	1.5843	0.7840

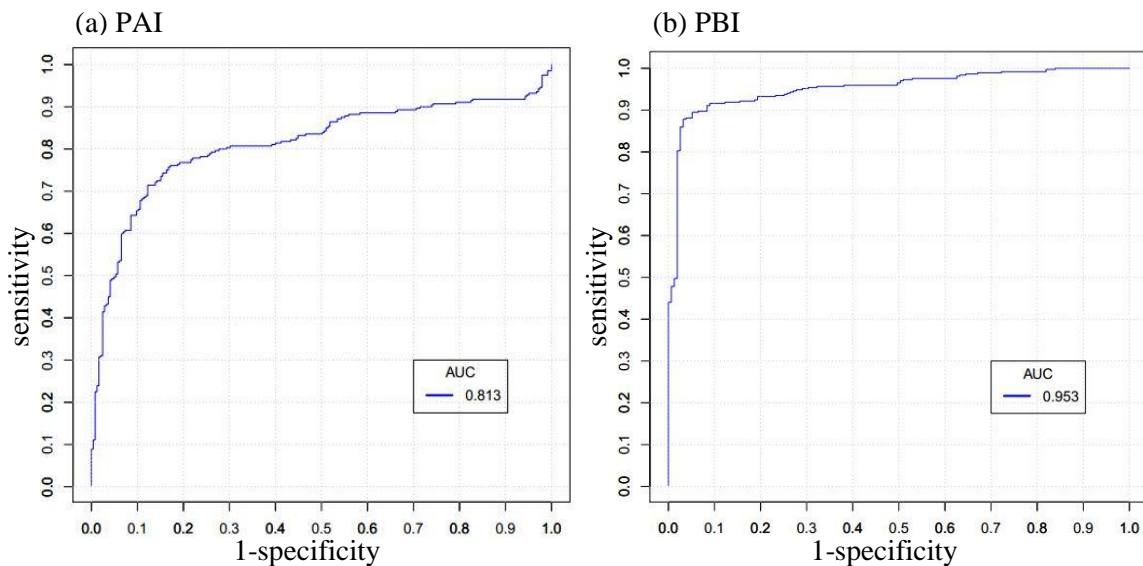


Figure 1. ROC curve and AUC value using PWSn to predict PAI and PBI.

3.2 Plaque progression prediction using combination of risk factors

Table 2 gives two best combinations of nine eight risk factors with PAI and PBI. Using PAI as the measure of plaque progression, the combination of Lipid area, Ave. CT, Ca area, LA, PA, PB, PWS, PWSn, and FSS showed the best prediction accuracy (Sen+Spe: 1.7087). Using PBI as the measure of plaque progression, the combination of Ave. CT, Ca area, PA, PWS, and PWSn gave the best prediction accuracy (Sen+Spe: 1.8698). The ROC curves of best combination using PAI and PBI were shown in Figure 2. The AUC values of best combination were 0.8632 and 0.9584, respectively.

Table 2. Prediction sensitivity and specificity using one single factor using PAI and PBI.

	Predictors	ProbCutoffs	Sensitivity	Specificity	Sen+Spe	AUC
PAI	Lipid Area+Ave. CT +Ca Area+LA+PA+PB +PWS+PWSn+FSS	0.3051	0.8679	0.8408	1.7087	0.8632
	Lipid Area+Ca Area +PA+PB+PWSn	0.3941	0.8857	0.8082	1.6939	0.9215
PBI	Ave. CT+Ca Area+PA +PWS+PWSn	0.8629	0.8892	0.9806	1.8698	0.9584
	Ave. CT+CaArea +PWSn+FSS	0.8373	0.8784	0.9871	1.8655	0.9522

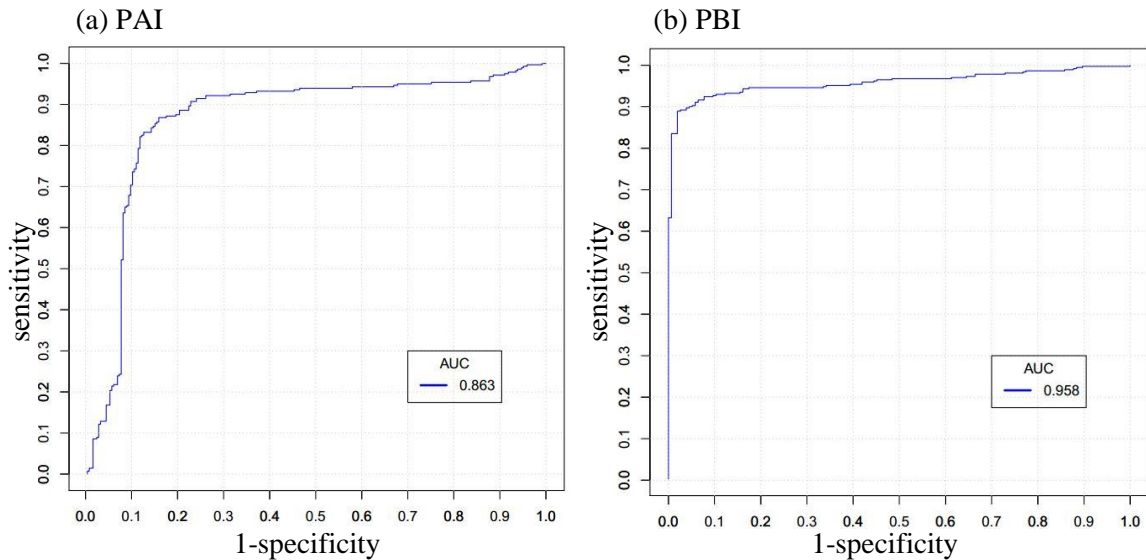


Figure 2. ROC curve and AUC value using best combination of factors to predict PAI and PBI.

4. Discussion

4.1 Significance of combining OCT and IVUS.

The accurate plaque progression prediction depends on accurate simulation, while accurate model depends on high resolution of medical imaging. Imaging resolution has been a major limitation for vulnerable plaque research since the introduction of medical imaging. The resolution for IVUS (100-150 μm) or MRI (300 μm) which is not enough to identify vulnerable

plaques with thin cap thickness < 65 micron. The combination of OCT and IVUS could possess the capabilities of detecting thin fibrous cap and penetrating vessel thickness. OCT+IVUS is able to provide more accurate cap thickness information to promote both the morphological and mechanical analyses in vulnerable plaque research.

4.2 Significance of combining mechanical and morphological risk factors for plaque progression prediction.

Most plaque progression research group paid attention to morphological factors and flow shear stress and seldom considered structural plaque stress and strain [9,10]. While it is well accepted that mechanical forces play an important role in plaque progression, research work based on patient follow-up data demonstrating that is rare. In fact, plaque mechanical state is affected by both fluid and structure forces. Tang's group used FSI models and patient follow-up data to investigate the influence of structural stress/strain for plaque vulnerability and progression [2]. By using OCT and IVUS data with follow-up, we constructed coronary plaque FSI models with cycle bending and perform progression prediction using nine morphological and mechanical risk factors. Our pilot study indicated that combining morphological and mechanical factors could give better predictions.

4.3 Limitations

One major limitation of this study is lack of histology data as the golden standard. Manual segmentation results based on IVUS and OCT images were considered as the alternative to the golden standard. Another limitation is the small sample size of OCT image studied. Large-scale studies with more OCT image are needed to validate and improve the significance of prediction method.

5. Conclusion.

Combining morphological and mechanical risk factors may lead to more accurate progression prediction, compared to the predictions using single factors. IVUS+OCT formed basis for accurate data for morphological and mechanical factors.

Funding:

This research was supported in part by NIH grant R01 EB004759 and a Jiangsu Province Science and Technology Agency grant BE2016785.

There are no conflicts of interest in this study.

References

1. Friedman MH, Krams R, Chandran KB. Flow interactions with cells and tissues: cardiovascular flows and fluid-structure interactions. *Annals Biomedical Engineering* 2010;38(3):1178-1187.
2. Tang D, Kamm RD, Yang C, Zheng J, Canton G, Bach R, Huang X, Hatsukami TS, Zhu J, Ma G, Maehara A. Image-based modeling for better understanding and assessment of atherosclerotic plaque progression and vulnerability: Data, modeling, validation, uncertainty and predictions. *Journal of Biomechanics* 2014; 47(4):834-846.
3. Clarkson TB, Prichard RW, Morgan TM, Petrick GS, Klein KP. Remodeling of coronary arteries in human and nonhuman primates. *JAMA* 1994;271: 289-294.
4. Glagov S, Weisenberg E, Zarins CK, Stankunavicius R, Kolettis GJ. Compensatory enlargement of human atherosclerotic coronary arteries. *N Engl J Med* 1987;316:1371-1375.
5. Zarins CK, Weisenberg E, Kolettis G, Stankunavicius R, Glagov S. Differential enlargement of artery segments in response to enlarging atherosclerotic plaques. *J Vasc Surg.* 1988;7:386-394.
6. Gerber TC, Erbel R, Gorge G, Ge J, Rupprecht H-J, Meyer J. Extent of atherosclerosis and remodeling of the left main coronary artery determined by intravascular ultrasound. *Am J Cardiol* 1994;73:666-671.

7. Mintz GS, Kent KM, Pichard AD, Satler LF, Popma JJ, Leon MB. Contribution of inadequate arterial remodeling to the development of focal coronary artery stenosis. An intravascular ultrasound study. *Circulation* 1997;95:1791–1798.
8. Nakamura Y, Takemori H, Shiraishi K, et al. Compensatory enlargement of angiographically normal coronary segments in patients with coronary artery disease: in vivo documentation using intravascular ultrasound. *Angiology* 1996;47:775–781.
9. Stone GW, Maehara A, Lansky AJ, de Bruyne B, Cristea E, Mintz GS, Mehran R, McPherson J, Farhat N, Marso SP, Parise H, Templin B, White R, Zhang Z, Serruys PW. The PROSPECT Investigators: A prospective natural-history study of coronary atherosclerosis. *N. Engl. J. Med.* 2011;364 (3):226–235.
10. Samady H, Eshtehardi P, McDaniel MC, Suo J, Dhawan SS, Maynard C, Timmins LH, Quyyumi AA, Giddens DP. Coronary artery wall shear stress is associated with progression and transformation of atherosclerotic plaque and arterial remodeling in patients with coronary artery disease. *Circulation* 2011;124, 779–788.
11. Corban MT, Eshtehardi P, Suo J, McDaniel MC, Timmins LH, RassoulArzrumly E, Maynard C, Mekonnen G, King 3rd S, Quyyumi AA, Giddens DP, Samady H. Cobination of plaque burden, wall shear stress, and plaque phenotype has incremental value for prediction of coronary atherosclerotic plaque progression and vulnerability. *Atherosclerosis* 2014;232, 271–276.
12. Maurice RL, Ohayon J, Finet G, Cloutier G. Adapting the Lagrangian speckle model estimator for endovascular elastography: theory and validation with simulated radio-frequency data. *J. Acoust. Soc. Am.* 2004;116 (2), 1276–1286.
13. Wang L, Tang D, Maehara A, Wu Z, Yang C, Muccigrosso D, Zheng J, Bach R, Billiar KL, Mintz GS. Fluid-structure interaction models based on patient-specific IVUS at baseline and follow-up for prediction of coronary plaque progression by morphological and biomechanical factors: A preliminary study. *Journal of biomechanics*. 2018;68:43-50.
14. Tearney GJ, Regar E, Akasaka T, Adriaenssens T, Barlis P, Bezerra HG, Bouma B, Bruining N, Cho JM, Chowdhary S, Costa MA. Consensus standards for acquisition, measurement, and reporting of intravascular optical coherence tomography studies: a report from the International Working Group for Intravascular Optical Coherence Tomography Standardization and Validation. *Journal of the American College of Cardiology*. 2012;59(12):1058-72.
15. Burke AP, Kolodgie FD, Farb A, Weber D, Virmani R. Morphological predictors of arterial remodeling in coronary atherosclerosis. *Circulation*. 2002;105(3):297-303.
16. Brown AJ, Obaid DR, Costopoulos C, Parker RA, Calvert PA, Teng Z, Hoole SP, West NE, Goddard M, Bennett MR. Direct comparison of virtual-histology intravascular ultrasound and optical coherence tomography imaging for identification of thin-cap fibroatheroma. *Circ. Cardiovasc Imaging*. 2015;8(10): e003487.
17. Uemura S, Ishigami KI, Soeda T, Okayama S, Sung JH, Nakagawa H, Somekawa S, Takeda Y, Kawata H, Horii M, Saito Y. Thin-cap fibroatheroma and microchannel findings in optical coherence tomography correlate with subsequent progression of coronary atheromatous plaques. *European heart journal*. 2011;33(1):78-85
18. Ræber L, Heo JH, Radu MD, Garcia-Garcia HM, Stefanini GG, Moschovitis A, Dijkstra J, Kelbaek H, Windecker S, Serruys PW. Offline fusion of co-registered intravascular ultrasound and frequency domain optical coherence tomography images for the analysis of human atherosclerotic plaques. *EuroIntervention*. 2012;8(1):98-108.
19. Guo X, Giddens DP, Molony D, Yang C, Samady H, Zheng J, Mintz GS, Maehara A, Wang L, Pei X, Li ZY. Combining IVUS and Optical Coherence Tomography for More Accurate Coronary Cap Thickness Quantification and Stress/Strain Calculations: A Patient-Specific Three-Dimensional Fluid-Structure Interaction Modeling Approach. *Journal of biomechanical engineering*. 2018;140(4):041005
20. Yang C, Bach R, Zheng J, Naqa IE, Woodard PK, Teng Z, Billiar KL, Tang D. “In vivo IVUS-Based 3D Fluid Structure Interaction Models with Cyclic Bending and Anisotropic Vessel Properties for Human Atherosclerotic Coronary Plaque Mechanical Analysis,” *IEEE Trans. Biomed. Engineering*, 2009;56(10): 2420-2428
21. Holzapfel GA, *Nonlinear Solid Mechanics: A Continuum Approach for Engineering*, Wiley, Chichester, 2000; New York
22. Guo X, Zhu J, Maehara A, Monoly D, Samady H, Wang L, Billiar KL, Zheng J, Yang C, Mintz GS, Giddens DP. Quantify patient-specific coronary material property and its impact on stress/strain calculations using in vivo IVUS data and 3D FSI models: a pilot study. *Biomechanics and modeling in mechanobiology*. 2017;16(1):333-44
23. Wu Z, Yang C, Tang D. In vivo serial MRI-based models and statistical methods to quantify sensitivity and specificity of mechanical predictors for carotid plaque rupture: location and beyond. *Journal of biomechanical engineering*. 2011;133(6):064503.

Failure modes analysis of a multi-story subway station under stochastic earthquake based on probability density evolution method

*Zhiqian Liu ¹ and † Zhiyi Chen ^{1,2}

¹ Department of Geotechnical Engineering, Tongji University, Shanghai 20092, China

² Key Laboratory of Geotechnical and Underground Engineering of Ministry of Education, Shanghai 200092, China

*Presenting author: 1510165@tongji.edu.cn

†Corresponding author: zhiyichen@tongji.edu.cn

Abstract

With the deep development and utilization of underground space, complex underground structures represented by multi-story subway station structure are constantly emerging. The damage of DAIKAI subway station during the 1995 Kobe earthquake indicates that underground structures may be seriously damaged under strong earthquake. Based on the probability density evolution method (PDEM), this paper studies the failure mode of a multi-story subway station structure under rare earthquake from the perspective of stochastic analysis, and gives the vulnerable spot of the multi-story subway station structure. It is found that the story drift of the four story of the structure does not exceed the elastic-plastic limit, but the failure probability of most structure components is more than 0.5. The vulnerable spots of the multi-story subway stations structure are mainly including the central columns, plates in middle stories, top plates, joints of bottom plate and sidewall. Failure of the internal components, such as central columns and plates in middle stories, are mainly bending failure, while failure of the outer frame components, such as the top and bottom plates and side walls, are mainly shear failure. In addition, the failure modes of multi-story subway stations can be roughly divided into four types. In the failure mode type I and II, most of failure appears at both vertical and horizontal components, while in the failure mode type III and IV most of failure occurs at horizontal components.

Keywords: Multi-story subway station, Probability density evolution method, Stochastic earthquake motion, Vulnerable spot, Failure mode

1 Introduction

With the deep development and utilization of underground space, complex underground structures represented by multi-story subway station structure are constantly emerging. During the 1995 Hanshin earthquake, the DAIKAI subway station was badly damaged [1][2], which led to the intensive research of seismic performance of subway station structures [3][4].

Compared with the typical single-story or two-story subway station, the structure of multi-story subway station structure is more complex, which causes the vulnerable spots and failure mode of the structure not easy to be undetected. However, up to now, most scholars have adopted deterministic methods, such as numerical dynamic time history analysis method [5][6] and model test method [7] under one or several given ground motions as well as Pushover analysis method [8][9], to explore the seismic responses and vulnerable spots of the structure without reasonable consideration of the randomness of the ground motion.

In order to fully understand the weak link of subway station structure under earthquake and based on the probability density evolution method (PDEM) and the thought of equivalent extreme event, the seismic reliability of a multi-story subway station structure under random earthquake is studied in this paper by using deformation and component strength as evaluation indexes. The main failure mode types of the multi-story subway stations structure under stochastic earthquake motion are summarized.

2 Project background

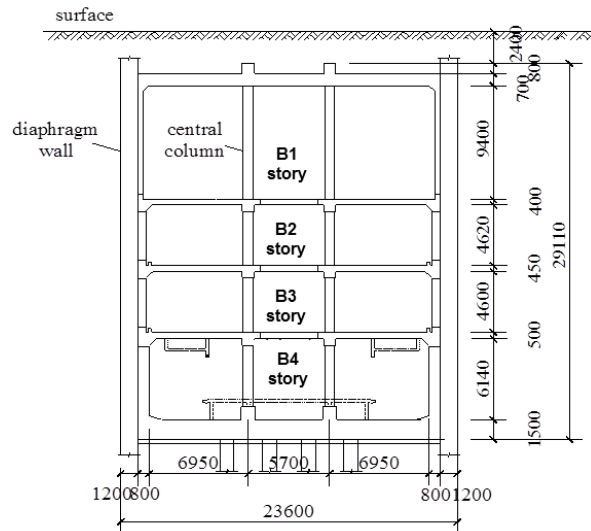


Figure 1. Standard cross section of the subway station, dimensions in mm

The multi-story subway station analyzed in this paper is a four-story three-span island station which located in Shanghai, China. Fig. 1 gives the standard cross section size of the station structure. The width and height of the standard section of the station is 23.6m and 29.1m, respectively, and the buried depth of the roof is 3.2m. The station is constructed by cut and cover method, and the diaphragm wall is used in the enclosure structure. The inner lining wall is connected to the diaphragm wall through the embedded parts, so that both of them bear forces together and form the side wall. The longitudinal distance between central columns is 8m. The cross section size of the column in B1 and B2 story is 1.1m long and 0.7m wide, and the size of the column in the B3 and B4 story is 1.4m and long 0.7m wide. According to “Chinese Code for Seismic Design of Urban Rail Transit Structures” [10], the shear wave velocity of the soil is 150m/s.

3 Numerical model

The dynamic analysis model of soil structure interaction (SSI) is established in the finite element code ABAQUS [11]. Mohr-Coulomb constitutive model with Rayleigh damping is applied to the soil. The Poisson's ratio of soil is 0.3, and the internal friction angle and cohesion are 15° and 20kPa, respectively.

Beam element (B21) is chosen for the structure. Central columns are made of C45 concrete [12], whose elastic modulus, tensile and compression strength are 33.5GPa, 2.51 and 29.6 MPa, respectively. Other parts of the structure were made of C35 concrete, whose parameters are 31.5GPa, 2.20 and 23.4 MPa, respectively. Poisson's ratio of both C35 and C45 concrete

is 0.2. The concrete damaged plasticity model is adopted to better simulate the dynamic response of the structure. The calculation of damage parameters has been done in previous studies [8][9]. Idealized elastic-plastic model was selected for rebar. Rebar used in central columns and other parts were HRB400 and HRB335, respectively. And the yield strength of HRB400 and HRB335 is 400MPa and 335MPa, respectively. The soil-structure interaction is defined by the Coulomb friction law. The coefficient of friction is assumed to be 0.4, which is equivalent to the friction angle of 22° .

The size of the structure is determined by the axis of the standard cross section shown in Fig. 1. In this paper, the single side width of the soil mass is 250m, which is more than 3 times the structural width specified in “Chinese Code for Seismic Design of Buildings” [13]. The depth of soil is 85m. The infinite element boundary is adopted as the lateral soil boundary to reduce the boundary effect [11]. The established numerical model is shown in Fig. 2.

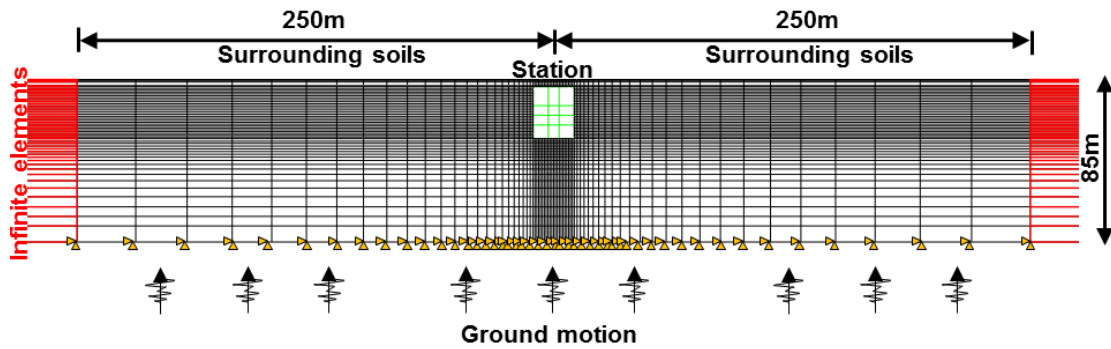


Figure 2. Numerical model of the underground subway station with surrounding soils

4 Process of PDEM

4.1 Numerical analysis process of PDEM

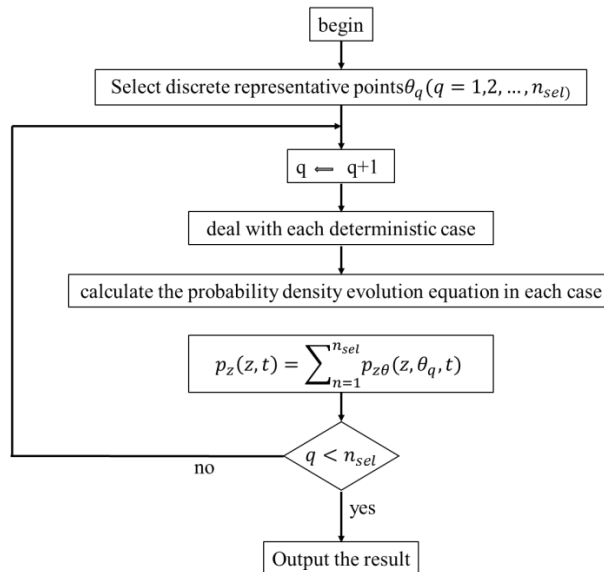


Figure 3. Basic steps of PDEM

The probability density evolution method is proposed by Li and Chen [14] based on the principle of probability conservation. After more than ten years of development, a relatively complete system of PDEM has been formed. The following is a brief introduction to the specific steps of probability density evolution analysis of structural random response. For more detailed contents of the method, please refer to the references [15].

The basic steps of stochastic response analysis of structures using PDEM are shown in Fig. 3, which are mainly divided into four steps. First are the discretization of probability space and the determination of the probability. Select a discrete set of representative points $\theta_q (q=1,2,\dots,n_{sel})$ in the random parameter space Ω_Θ , where n_{sel} represents the total number of the points. At the same time, according to the principle of selecting points, the probability of each representative point is determined. In this paper, we mainly study the vulnerable spots and failure mode of the structure under stochastic earthquake motion, which means the randomness comes from the input earthquake motion. Thus, in this paper, n_{sel} represents the number of stochastic earthquake motion samples.

The second step is to solve the deterministic dynamic system. For every representative point θ_q , carry out deterministic analysis on the dynamical system and then get the partial derivative of the concerned physical quantity to time $\dot{\mathbf{Z}}(\theta_q, t)$ ($q=1,2,\dots,n_{sel}$). The deterministic method used in this paper is the dynamic time history analysis method, that is, each stochastic earthquake motion sample is taken as input, the dynamic time history analysis of the structure is carried out to obtain the partial derivative of the concerned physical quantity to time, such as structural deformation and internal forces.

The third step is to solve the probability density evolution equation. The generalized probability density evolution equation can be expressed as Eq. (1).

$$\frac{\partial p_{Z\Theta}(z, \theta, t)}{\partial t} + \dot{\mathbf{Z}}(\theta, t) \frac{\partial p_{Z\Theta}(z, \theta, t)}{\partial z} = 0 \quad (1)$$

For each selected representative point $\theta_q (q=1,2,\dots,n_{sel})$, introduce $\dot{\mathbf{Z}}(\theta_q, t)$ into Eq. (1) and solve the equation under the corresponding initial conditions and boundary conditions using the finite difference method. Denote the solution as $p_{Z\Theta}(z, \theta_q, t)$. In this chapter, we use the finite difference method of Total Variation Diminishing (TVD) to solve the generalized probability density evolution equation.

The last step is to obtain the final solution through summation. Synthesize the solutions in previous step to obtain the numerical solution of PDEM equation, which is expressed as Eq. (2).

$$p_Z(z, t) = \sum_{q=1}^{n_{sel}} p_{Z\Theta}(z, \theta_q, t) \quad (2)$$

4.2 Generation of stochastic earthquake motions

In the numerical solution process of PDEM, the stochastic process of earthquake motion needs to be discrete in the probability space, and a certain number of stochastic earthquake motion samples are obtained. At present, among the various simulation methods of stochastic process, the spectral representation method, which is formally proposed by Shinozuka [16],

has been improved after more than 40 years of development. However, in order to ensure the appropriate accuracy, this method often requires a lot of random variables and a large amount of computation. In this paper, Spectral representation - Random function method proposed by Liu et al. [17] is used to simulate the random process of ground motion. This method reduces the number of random variables to 1-2. 254 representative points are selected in the probability space of the basic random variable space Ω_Θ , and 254 earthquake motion samples are generated for the dynamic time history analysis of the numerical model. According to “Chinese Code for Seismic Design of Buildings” [13], the station site classification is IV and the seismic fortification intensity is 7 (PGA of rare earthquake equals to 0.22g). The relevant parameters for generating the ground motion samples are determined according to “Chinese Code for Seismic Design of Buildings” [13]. Fig. 4 gives 3 typical ground motion samples.

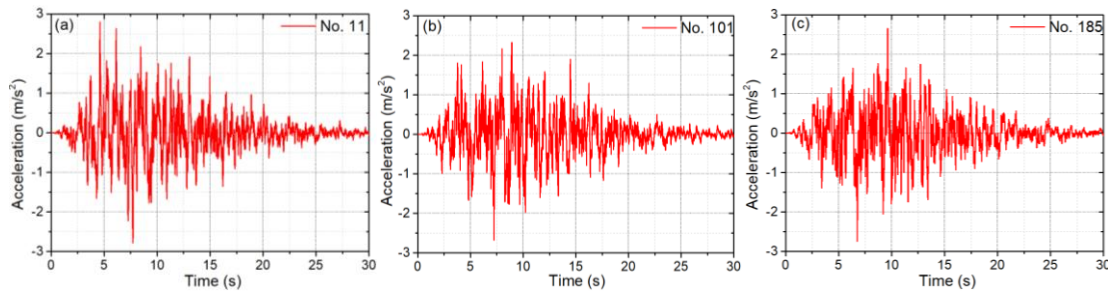


Figure 4. Typical acceleration time history samples

5 Result analysis

5.1 Structural reliability evaluation based on story drift

When evaluating the seismic performance of subway station structure, the maximum response value is often concerned. Therefore, based on PDEM and the equivalent extreme event thought, the corresponding equivalent extremum events are constructed from the deformation index to calculate the seismic reliability of the multi-story subway station structure under stochastic earthquake motion.

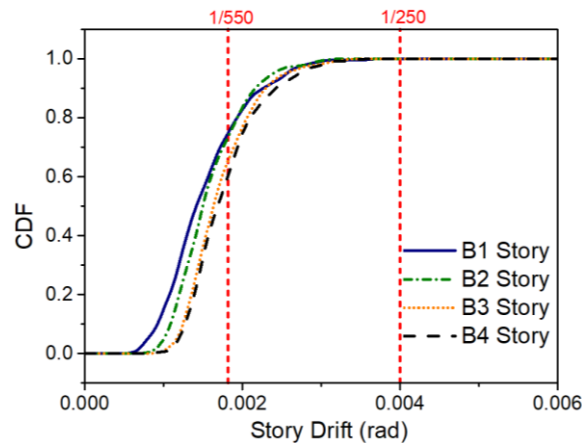


Figure 5. CDF of the extreme value of story drift

The story drift is an important index for evaluating the seismic performance of the structure from the point of view of deformation. Fig. 5 shows the cumulative probability distribution of story drift of four stories. “Chinese Code for Seismic Design of Buildings” [13] stipulates that

the elastic limit of story drift of the underground structure is $1/550$, and the elastoplastic limit of story drift is $1/250$. The story drift less than $1/550$ indicates that the structure has no damage or slight damage without repair; the story drift beyond the elastic limit means that the structural damage needs to be repaired, and the story drift over the elastoplastic limit means that the structural damage is serious. It can be obtained from Fig. 5 that the structure reliability of B1-B4 story is 0.739, 0.724, 0.641, 0.588, respectively, according to the elastic limit of story drift; and the reliability of the B1-B4 story is all 1 according to the elastoplastic limit. It can be seen that the reliability of the B1-B4 story decreases from top to bottom, indicating that the reliability of the structure decreases gradually along the soil depth. On the one hand, because of the load transfer, the axial pressure of the deep column increases obviously [18]; on the other hand, the depth increase makes the soil and water pressure of the deep side wall increase, which has a negative effect on the structure safety. In general, when the story drift is used as the deformation index to evaluate the structural reliability, the story drifts of four stories are less than the elastoplastic limit of story drift. This indicates that, the four stories of the structure are in two stages of slight damage and moderate damage, and there is no serious damage under the earthquake motion intensity of $\text{PGA}=0.22\text{g}$.

5.2 Failure probability analysis of structural components

The failure probability of structural components is calculated by strength index of components. According to the deterministic dynamic time history analysis, the internal force extremums of the component ends of the structure under 254 stochastic earthquake motions are obtained. Combined with the idea of equivalent extreme event, probability density evolution analysis is carried out, and the probability distribution of extreme value of internal force is obtained. The internal force bearing capacity (bending capacity and shear capacity) of the component can be used as a conservative limit of strength index, which is calculated according to “Chinese Code for Design of Concrete Structure” [12].

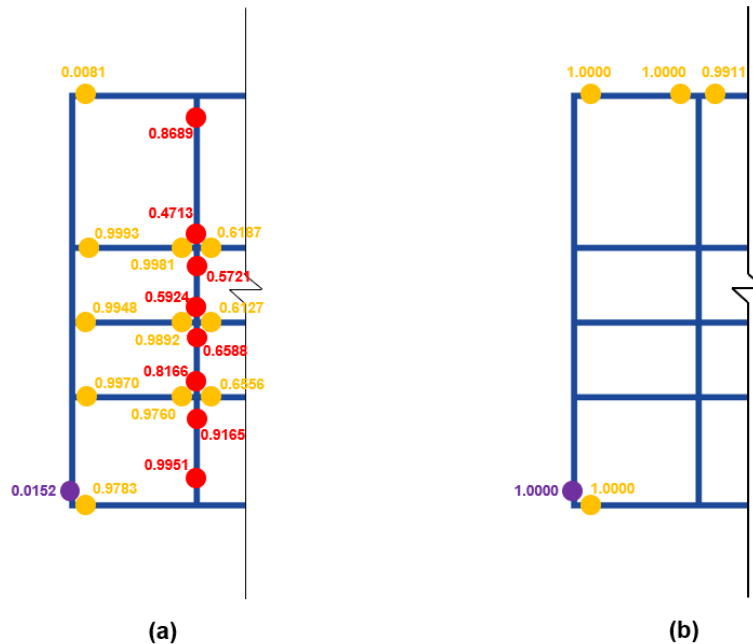


Figure 6. Failure probability of different positions (a) bending failure; (b) shear failure

Fig. 6 (a) (b) gives the failure probability of component ends according to the limit values of bending capacity and shear capacity. From Fig. 6 (a), it can be seen that the bending failure

probability of the vast majority of internal components, including the columns and plates in middle stories, is all more than 0.5, which indicates that these internal components have great possibility of bending failure under the earthquake motion intensity of $PGA=0.22g$. However, the bending failure probability of outer frame components such as top and bottom plates, side walls is very close to 0 except for the joint of bottom plate and sidewall. This is because, in order to bear the peripheral water and soil pressure, the size and stiffness of outer frame components is obviously larger than the internal components of the structure, which makes the bending deformation of the end of the internal components increase, thus the bending moment may exceed the bending capacity. In addition, we can see from Fig. 6 (a) that the side span of plates in middle stories is the largest failure probability component, and the failure probability of the component ends is all more than 0.95. From Fig. 6 (b), it can be seen that the components whose shear failure probability greater than 0 are mainly the outer frame components, such as side walls, top and bottom plates, and the failure probability of the internal components is all equals to 0, which means that the shear failure will only occur in the outer frame. This is because the structure depth of the multi-story subway station is much larger than the ordinary subway station, the water and soil pressure on the outer frame is obviously increased, so the outer frame of the structure is more prone to shear failure. From the perspective of components failure probability, the vulnerable spots of the multi-story subway station structure are mainly the columns (especially in bottom story), plates in middle stories, top plates, and joints of bottom plate and side wall. In general, although it can be seen from Fig. 5 that the story drifts do not exceed than the elastoplastic limit, but most of the components of the structure may exceed the limit of bearing capacity. Therefore, the limit of the story drifts of underground structure in “Chinese Code for Seismic Design of Buildings” [13] may need further study.

5.3 Analysis of structural failure modes

Four types of failure modes can be obtained after sorting out the failure components according to the analysis of 254 stochastic earthquake motion cases. Fig. 7 shows the failure location of four types of failure modes, and Table 1 gives the number of cases corresponding to each failure mode type and failure components. In Fig. 7, red points indicate bending failure, blue points indicate shear failure, and purple points indicate bending shear failure.

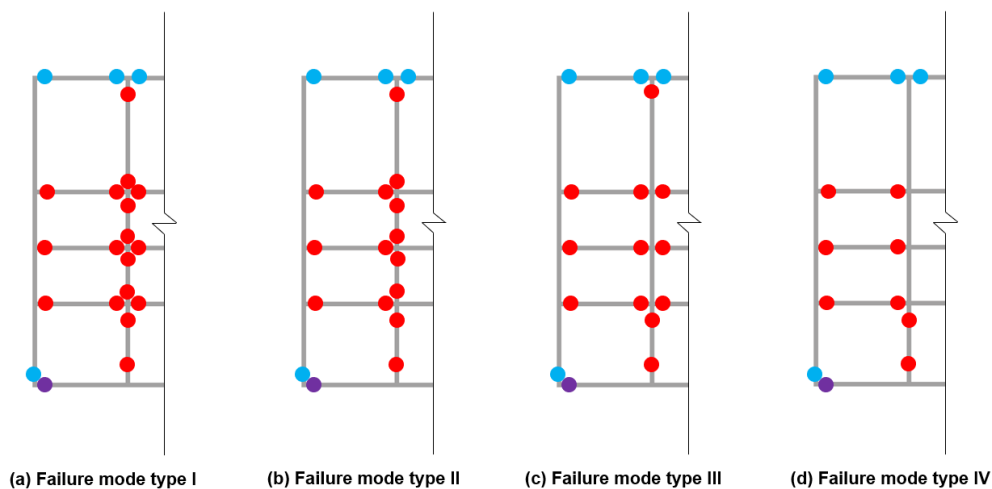


Figure 7. Classification of structural failure modes

From Table 1, we can see that the I-IV failure modes contain 123, 84, 37, and 10 cases, accounting for 48.4%, 33.1%, 14.6%, and 3.9% of the total cases respectively. In the four type

of failure modes, type I has the largest number of failure components. As shown in Fig. 7 (a), bending failure occurs at almost all the end of columns and plates in middle stories, and shear failure and bending shear failure occur respectively at the end of the top plates and the joints of bottom plate and sidewall. From Fig. 7 (b), it can be seen that compared with failure mode type I, failure components in type II do not include middle span of the plates in middle stories. From Fig. 7 (c), it can be seen that compared with failure mode type I, failure components in type III do not include columns in the middle stories. Type IV is one of the four failure modes with the least failure components. As shown in Fig. 7 (d), the failure components in type IV mainly include the side span plates in middle story, top plates, joints of bottom plate and sidewall, central columns in bottom story. In general, in the failure mode type I and II, most of failure appears at both vertical and horizontal components, while in the failure mode type III and IV most of failure occurs at horizontal components. The number of failure components and the number of the corresponding cases all decrease from the type I to type IV, which indicates that the structure tends to appear failure of most components under the rare earthquake intensity of 7 seismic fortification intensity.

Table 1. Classification and description of failure modes

Failure mode types	Failure location description	Total cases
I	Plates in middle story; top plates; joints of bottom plate and sidewall; central columns in most stories	123
II	Side span plates in middle story; top plates; joints of bottom plate and sidewall; central columns in most stories	84
III	Plates in middle story; top plates; joints of bottom plate and sidewall; central columns in bottom story	37
IV	Side span plates in middle story; top plates; joints of bottom plate and sidewall; central columns in bottom story	10

6 Conclusions

Based on PDEM and finite element elastoplastic dynamic time history analysis, the seismic reliability of a multi-story subway station structure under stochastic earthquake motion is studied in the presented paper with the evaluation index of story drift and component strength. The vulnerable spots and main failure modes are analyzed and summarized. The main conclusions are as follows:

- (1) Story drifts do not exceed than the elastoplastic limit, but most of the components of the structure may exceed the limit of bearing capacity, which means limit of the story drifts of underground structure in “Chinese Code for Seismic Design of Buildings” [13] may need further study.
- (2) From the perspective of failure probability of structural components, the vulnerable spots of the multi-story subway station structure are mainly the columns (especially in bottom story), plates in middle stories, top plates, joints of bottom plate and sidewall.
- (3) Failure of the internal components, such as central columns and plates in middle stories, are mainly bending failure, while failure of the outer frame components, such as the top and bottom plates and side walls, are mainly shear failure.
- (4) The failure modes of the multi-story subway station structure can be roughly divided into four types. The number of failure components and the number of the corresponding cases all decrease from the type I to type IV. In the failure mode type I and II, most of failure appears

at both vertical and horizontal components, while in the failure mode type III and IV most of failure occurs at horizontal components.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (Grant No. 41472246, 51778464), Key laboratory of Transportation Tunnel Engineering (TTE2014-01), and “Shuguang Program” supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission. All supports are gratefully acknowledged.

References

- [1] Nakamura, S., Yoshida, N., and Iwatate, T. (1996) Damage to Daikai subway station during the 1995 Hyogoken-Nambu earthquake and its investigation, *Japan Society of Civil Engineers, Committee of Earthquake Engineering*, 287-295.
- [2] Yamato, T., Umehara, T., Aoki, H., Nakamura, S., Ezaki, J., and Suetomi, I. Damage to Daikai subway station, Kobe rapid transit system and estimation of its reason, *Proceedings, technical conference on the great Hanshin-Awaji Earthquake*, JSCE, Tokyo, 1996, 247-254. (in Japanese)
- [3] Chen, G., Wang, Z., Zuo, X., Du, X., and Gao, H. (2013) Shaking table test on the seismic failure characteristics of a subway station structure on liquefiable ground, *Earthquake Engineering & Structural Dynamics* **42**(10), 1489-1507.
- [4] Khani, S., and Homami, P. (2014) Seismic performance of underground shallow subway stations in soft soil, *Journal of Engineering Geology* **8**(1), 1983-2002.
- [5] Parra-Montesinos, G. J., Bobet, A., and Ramirez, J. A. (2006) Evaluation of soil-structure interaction and structural collapse in Daikai subway station during Kobe earthquake, *ACI structural journal* **103**(1), 113.
- [6] Zhuang, H., Hu, Z., Wang, X., and Chen, G. (2015) Seismic responses of a large underground structure in liquefied soils by FEM numerical modelling, *Bulletin of Earthquake Engineering* **13**(12), 3645-3668.
- [7] Chen, G., Chen, S., Qi, C., Du, X., Wang, Z., and Chen, W. (2015) Shaking table tests on a three-arch type subway station structure in a liquefiable soil, *Bulletin of Earthquake Engineering* **13**(6), 1675-1701.
- [8] Chen, Z. Y., Chen, W. and Zhang, W. (2014) Seismic performance evaluation of multi-story subway structure based on pushover analysis, *Advances in Soil Dynamics and Foundation Engineering* 444-454, ASCE.
- [9] Chen, Z. Y., Chen, W., Zhang, W. and Lou, M. L. (2016) Effects of Axial Compression Ratio of Central Columns on Seismic Performance of a Multi-Story Underground Structure, *International Journal of Computational Methods* **13**(04), 1641014.
- [10] GB50909 (2014) *Code for Seismic Design of Urban Rail Transit Structures*, China Planning Press, Beijing, China.
- [11] ABAQUS, Users Manual V. 6.14. (2015) Dassault Systemes Simulia Corp., Providence, RI.
- [12] GB50010 (2010) *Code for Design of Concrete Structure*, China Architecture & Building Press, Beijing, China.
- [13] GB50011 (2010) *Code for Seismic Design of Buildings*, China Architecture & Building Press, Beijing, China.
- [14] Li, J. and Chen, J. (2008) The principle of preservation of probability and the generalized density evolution equation, *Structural Safety* **30**(1), 65-77.
- [15] Li, J. and Chen, J. (2009) *Stochastic dynamics of structures*, John Wiley & Sons.
- [16] Shinozuka, M. (1972) Monte Carlo solution of structural dynamics, *Computers & Structures* **2**(5-6), 855-874.
- [17] Liu Z.J., Zeng B. and Wu L.Q. (2015) Simulation of non-stationary ground motion by spectral representation and random functions, *Journal of Vibration Engineering* **28**(3), 411-417. (in Chinese)
- [18] Chen, Z., Chen, W., Li, Y. and Yuan, Y. (2016) Shaking table test of a multi-story subway station under pulse-like ground motions, *Soil Dynamics and Earthquake Engineering* **82**, 111-122.

Hydrodynamic characteristics of twin rudders

*Linfeng Chen¹, Xinyao Zhu¹ and †Lilan Zhou²

¹ School of Naval Architecture and Ocean Engineering, Jiangsu University of Science and Technology, Zhenjiang, China

² Key Laboratory of High Performance Ship Technology of Ministry of Education, Wuhan University of Technology, Wuhan, China

*Presenting author: chenlinfeng@just.edu.cn

†Corresponding author: zhoulilan1985@126.com

Abstract

Computational fluid dynamic (CFD) simulations by solving Reynolds-averaged Naviers-Stokes (RANS) equations with a SST two-equation model using OpenFOAM codes are performed. The difference of hydrodynamic forces between the single-rudder and the parallel twin rudders of NACA0012 blades are examined from 5° to 25° angle of attack. It is found that the interaction between the twin rudders becomes strong as the angle of attack increases. Lately, the impact of the distance between the two rudders on the hydrodynamic forces at 15° angle of attack is studied for the parallel twin rudders. It is noticed that the interaction between the twin rudders becomes weak with the lateral spacing increasing, and the lift to drag ratio of the twin rudders monotonously decreases with the lateral spacing. Finally, the effectiveness of the stopping performance of the twin rudders at the different lateral spacing is analyzed. As a result, the drag reaches the largest value as the lateral spacing equals to 1.3c.

Keywords: twin rudders, computational fluid dynamics, hydrodynamic characteristics, RANS

Introduction

The performance of ship rudders depends on the rudder hydrodynamic characteristics, and the rudder forces and moments are determined by the rudder area, angle of attach and incident flow velocity. In practical, the rudder area is sometimes limited due to shallow water. As a solution to solve such a problem, a twin-rudder configuration is commonly regarded.

By far, quite a few studies have been done for twin rudders, and hydrodynamic forces of twin rudders are mostly calculated by adding the corrected lift and drag profiles of the single rudder. Practically, the hydrodynamic characteristics of each rudder in twin rudders are interacted by each other such that the pressure profiles and the ambient flow of each rudder all vary and differ from those of the single rudder. Therefore, the hydrodynamic forces cannot be approximated as a duplication of the single rudder. The interaction between rudders is necessary to be taken into account when calculating hydrodynamic forces of the twin rudders.

For single-propeller twin-rudder ships, the asymmetrical behaviors would occur as the inflow is not parallel to ship's centerline. Hamamoto and Enomoto^[1] analytically and experimentally investigated the forces on a couple of rudders and the interaction between both rudders steered with rudder angle. Nagarajan *et al.*^[2] measured the rudder's axial force of a single-propeller twin-rudder ship, they proposed a prediction method to estimate the engine power of a ship installed with a special high lift twin-rudder system. The inflow characteristics to each one of the rudders of single-propeller twin-rudder system were investigated by Kang *et al.*^[3].

They proposed a method called “virtual zero rudder angle” arrangement to prevent asymmetric maneuvering characteristic of the ship, which were proved to be effective to improve ship’s propulsion performance.

Maneuverability and hydrodynamics of a twin-propeller twin-rudder ship were investigated by Yoshimura and Sakurai^[4]. They found that hydrodynamic characteristics of a twin-propeller twin-rudder are not so much different from those of a single-propeller single-rudder ship. Yoo *et al.*^[5] studied the maneuvering characteristics of a twin propeller/twin rudder ship during berthing and unberthing, and noticed that the tangential force acting on the rudder should be considered separately and that the bank suction effect between the hull and the quay developed an additional force. Khanfir *et al.*^[6] proposed a mathematical model for maneuverability and estimation of hydrodynamic coefficients of twin-propeller twin-rudder ship. Lately, captive model tests as well as free-running tests with a single-propeller twin-rudder ship and a twin-propeller twin-rudder ship were carried out by them^[7] to evaluate the effect of drift angle on the rudder forces and some peculiar phenomena concerning rudder normal force for twin-rudder ships.

As theoretical researches, Shcherbakov^[8] carried out a twin-rudder performance test for NACA 0018 in a water channel at $Re = 1.5 \times 10^4$ using PIV technique. The effect of the distance between two rudders on the hydrodynamics was examined, and the optimum distance was eventually determined. Liu and Hekkenberg^[9] implemented RANS simulations to present an initial study of the hydrodynamic characteristics of twin rudders at small attach angles.

A twin-rudder ship may help reduce the stopping distance by setting the two rudders outwards at 75° (called clam shell angles). Compared to a conventional reverse engine stopping, a twin-rudder at 75° may reduce the stopping distance by 50%^[10]. Hamamoto and Enomoto proposed analytical formulas of the ship speed drop and calculated the stopping time, the stopping distance as a ship stops at the clam shell angles. Hasegawa *et al.*^[11] carried out stopping test for a large vessel installed with a mariner type super VecTwin rudder. They simplified the complicated flow around the rudders to model the flow speed around the rudders and analyzed the outward rudder angles.

Although a few researches were carried out to investigate the maneuverability of some twin-rudder ships, the complicated flow around the rudders and the hydrodynamic characteristics are still paid little attention. In order to obtain accurate hydrodynamic coefficients of twin-rudder, computational fluid dynamic (CFD) simulations by solving Reynolds-averaged Naviers-Stokes (RANS) equations using a finite volume method with a shear-stress transport (SST) $k - \omega$ two-equation turbulence model in OpenFOAM are implemented to capture the complicated flow field in this study. The study puts a focus on the effect of the lateral spacing (the distance between the twin-rudder blades) on the hydrodynamic coefficients and on the drag coefficients at the clam shell angle.

Methodology

Flow model

The flow around the rudders are governed by the Reynolds-averaged equations for conservation of continuity and Navier-Stokes equations

$$\frac{\partial u_i}{\partial x_i} = 0$$

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j^2} - \frac{\partial \overline{u_i u_j}}{\partial x_j}$$

Where $i, j = 1, 2$. $\overline{u_i u_j}$ represents the Reynolds stress component, which is to be expressed in terms of a turbulent viscosity, ν_T and the mean flow gradients using Boussinesq approximation,

$$-\overline{u_i u_j} = \nu_T \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) - \frac{2}{3} k \delta_{ij}$$

where k is the turbulent kinetic energy and δ_{ij} is the Kronecker delta.

To solve the turbulent viscosity ν_T , a SST $k - \omega$ two-equation model proposed by Menter^[12] are used in this study. The two-equation model is given by the following

$$\begin{aligned} \frac{\partial}{\partial t}(\rho k) + u_j \frac{\partial}{\partial x_j}(\rho k u_i) &= \frac{\partial}{\partial x_j} \left[\left(\mu + \frac{\mu_T}{\sigma_k} \right) \frac{\partial k}{\partial x_j} \right] - \rho \overline{u_i u_j} \frac{\partial u_i}{\partial x_j} - \rho \beta^* k \omega \\ \frac{\partial}{\partial t}(\rho \omega) + \frac{\partial}{\partial x_i}(\rho \omega u_i) &= \frac{\partial}{\partial x_j} \left[\left(\mu + \frac{\mu_T}{\sigma_k} \right) \frac{\partial \omega}{\partial x_j} \right] - \frac{\alpha}{\mu_T} \overline{u_i u_j} \frac{\partial u_i}{\partial x_j} - \rho \beta \omega^2 + 2(1 - F_1) \rho \sigma_{\omega,2} \frac{\partial k}{\partial x_j} \frac{\partial \omega}{\partial x_j} \end{aligned}$$

where k is the turbulence kinetic energy, ω represents the specific dissipation rate and $\mu_T = \frac{\rho k}{\omega}$. For more details and the model coefficients, see Menter's paper^[12].

Simulation procedure and problem setup

The Reynolds-averaged equations for conservation of mass and momentum are discretized with a finite volume method, and solved by a PISO solver of OpenFOAM, in conjunction with a SST $k - \omega$ two-equation model.

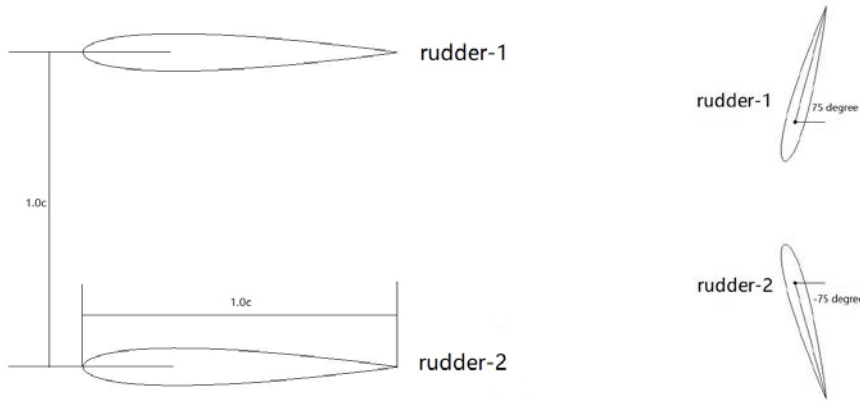


Figure 1. Parallel twin-rudder system of NACA0012 and twin rudders at a clam shell angle.

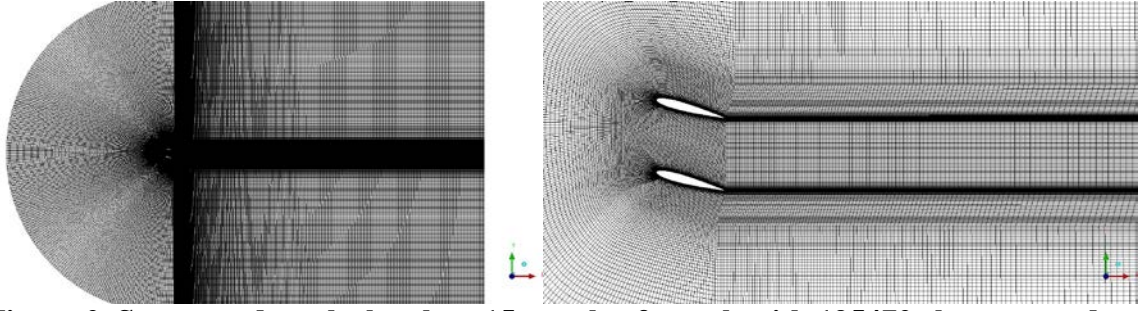


Figure 2. Structured mesh sketch at 150 angle of attack with 195472 elements and 194094 nodes for $d=1.0c$.

In this study, two NACA0012 blades are selected as the targeted parallel twin-rudder system (Figure 1). The computational domain includes a half circular of a radius of $20c$ (20 times rudder chord) and a $40c \times 40c$ rectangular (Figure 2). The size of the computational domain is sufficient to neglect the far-field effects.

Body-fitted, structured O-meshes for the computational domain are computed by ICEM CFD. A sketch of the mesh for a case with the lateral spacing $d=1.0c$ is illustrated in Fig. 2. For all simulations, the size of the elements near the blades is set such that the dimensionless off-wall distance y^+ is smaller than 1, where y^+ is defined as $y^+ = u_\tau y / \nu$, u_τ denotes the friction velocity near the blade, and y denotes the normal distance from the blade.

Result and discussion

Liu and Hekkenberg^[9] performed an initial study of the hydrodynamic characteristics of twin rudders, and analyzed the variation of the hydrodynamic coefficients with the attack angles. The present study extends Liu's work and examines the effect of the lateral spacing between the two rudders on the hydrodynamic coefficients. The discussion as follows contains the effect of angle of attack, the lateral spacing on the hydrodynamic coefficients of the parallel rudders, and the effect of the lateral spacing on the drag coefficients of the rudders at the clam shell angles.

Difference between single-rudder and twin-rudder

The first set of cases start with the twin rudders with $1.0c$ lateral spacing. Figure 3 shows the drag coefficients of the single-rudder and the twin rudders versus angle of attack. It can be seen that the drag coefficients of both system increases with the attack angle increasing. The drag coefficients of the twin rudders over all the attack angles are larger than those of the single rudder. It is noticed that rudder 1 causes the similar drag to the single rudder, while that the drag of rudder 2 decrease much due to the hydrodynamic interaction between the two rudders.

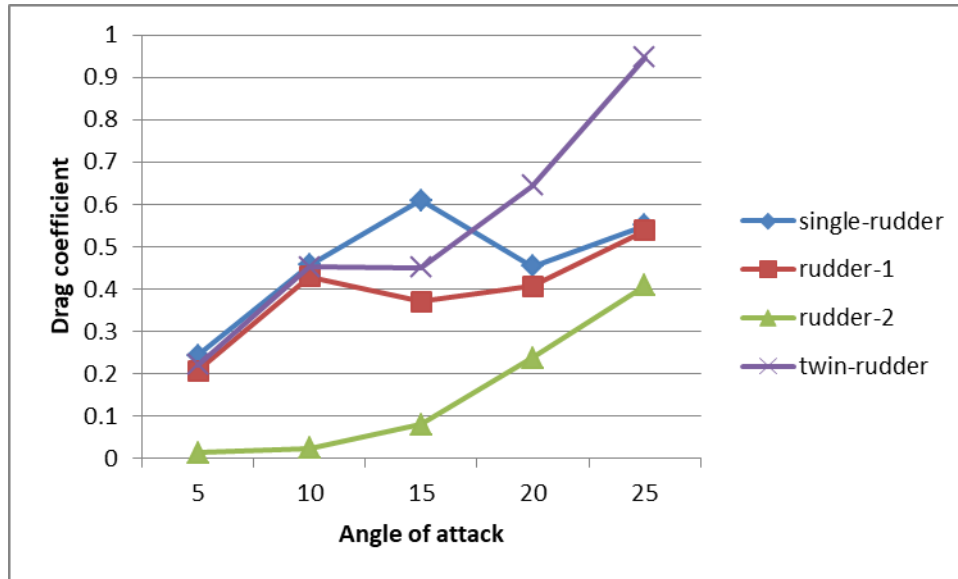


Figure 3 Drag coefficients versus angle of attack.

Figure 4 shows the lift coefficients of the single-rudder and the twin rudders versus angle of attack. For both systems, the lift coefficients reach the highest value at 15° . And the twin-rudder give rise to much larger lift coefficients than the single rudder. For the lift of the two rudders, it is noticed that rudder 2 causes the larger value than rudder 1 as the attack angle increases, and that the total lift is not just twice the lift of the single rudder, which is because that the hydrodynamic interaction of the two rudders becomes strong when the attack angle increase.

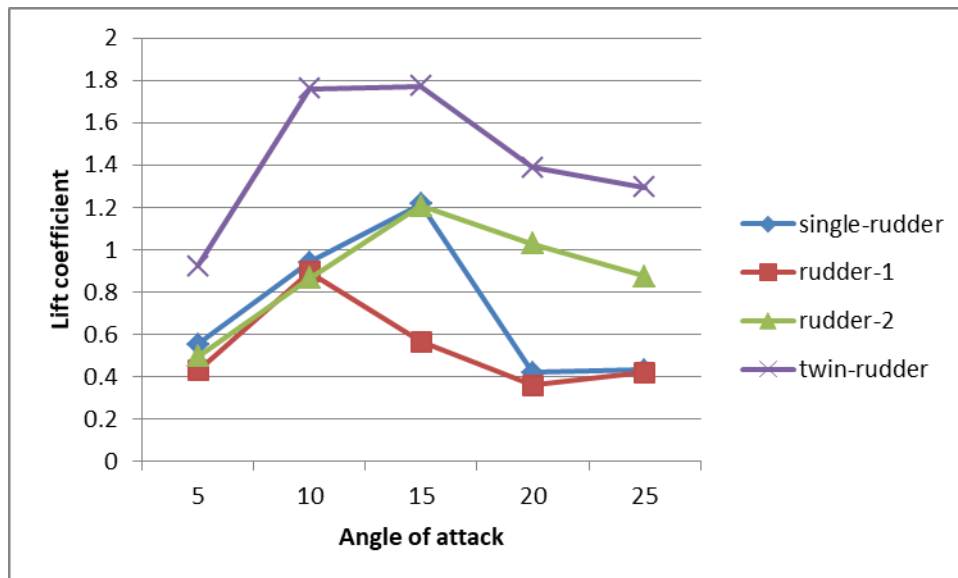


Figure 4 Lift coefficients versus angle of attack.

Figure 5 shows the ratio of the lift to the drag of the single-rudder and the twin rudders versus angle of attack. It can be seen that, for NACA0012, the value of the lift to drag ratio decreases with the angle of attack increasing for both systems. And it is found that the lift to drag ratio of the twin rudders is much better than the single rudder, which implies that the twin rudders can be used to improve the hydrodynamic performance when the single rudder is limited.

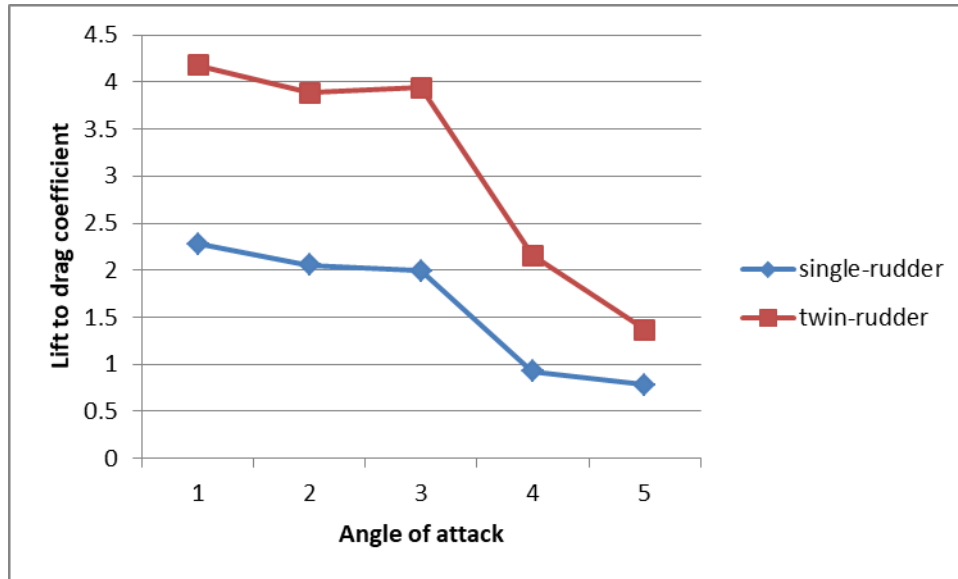


Figure 5 Lift to drag ratios versus angle of attack.

When performing the simulation of the flow around the rudders, it is found that the drag and lift coefficients of the two rudders start vibrating periodically (Figure 6) as the angle of attack is larger than 15° . While, at all angles of attack, the drag and lift coefficients of the single rudder normally converge to a stable value. This is possibly because that the flow interaction between the two rudders leads to more unstable vortices downstream so that the pressure downstream of the rudders becomes periodically unstable.

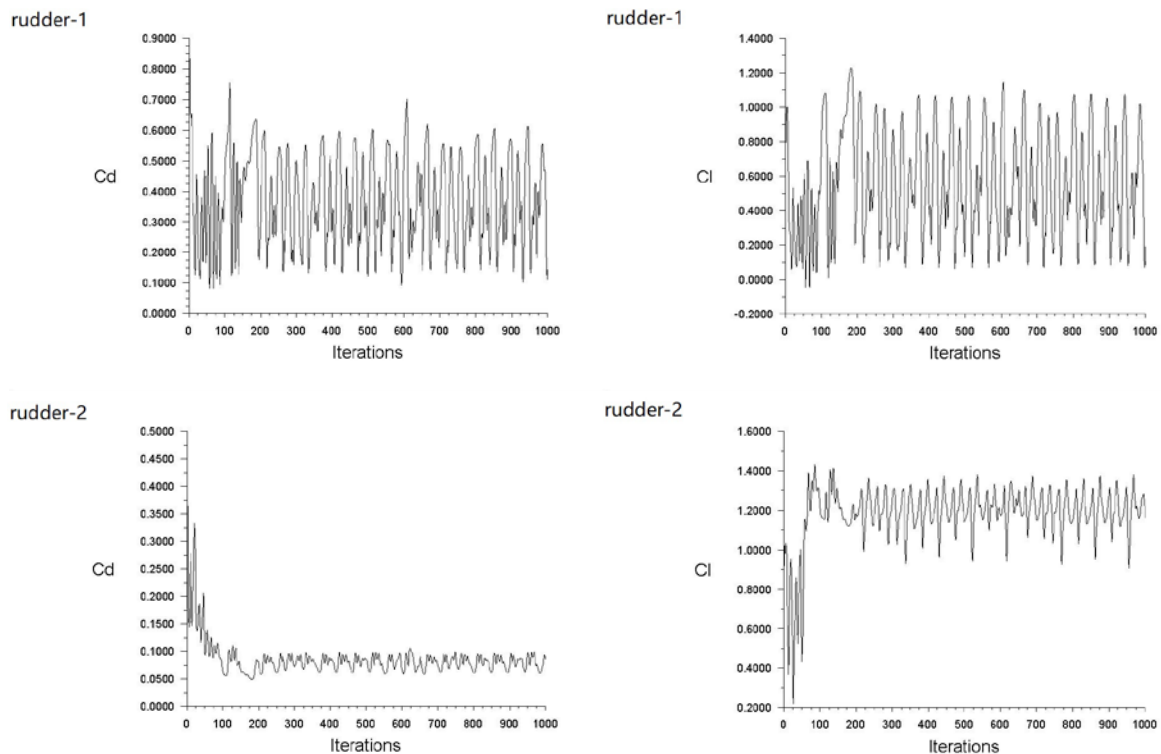
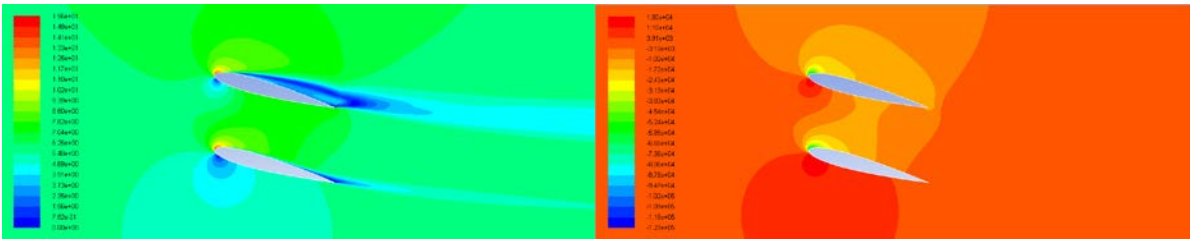


Figure 5 Evolution of the drag and lift coefficients of the two rudders.

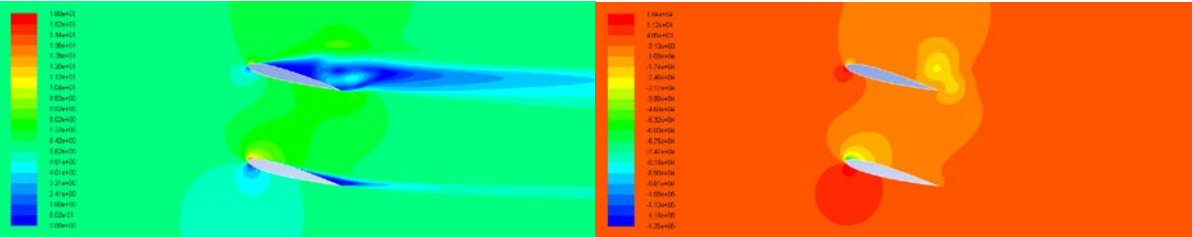
Impact of the lateral spacing on the hydrodynamics

The distance between the two rudder blades (d , lateral distance) is one of the most important factors which directly affects the flow interaction between the twin rudders. As rudders are mostly performed at small angles of attack within 15 degrees of course keeping and initial turning, here the twin rudders with the different lateral spacing at 15° angle of attack are considered. Figure 6 shows the velocity and static pressure contours around the twin rudders of the cases with the different lateral spacing. It can be seen from the velocity contours that the small distance causes an velocity increase between the two rudders as $d=0.6c$, such that the velocity at the leading edge of rudder 2 decreases and the vortex downstream of rudder 2 declines. As the lateral spacing increases, the interaction between the two rudders becomes relatively weak. Therefore the flow around the two rudders approaches to be similar. Similarly, the same phenomena occurs to the static pressure around the two rudders.

$d=0.6c$



$d=1.0c$



$d=1.4c$

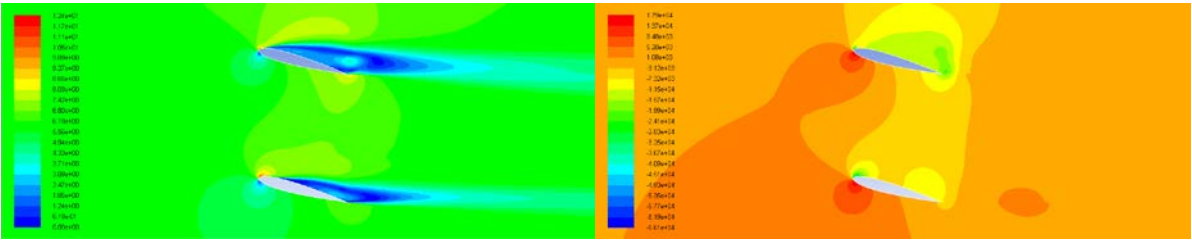


Figure 6. Velocity and static pressure contours of the cases with the different lateral spacing.

Figure 7 shows the evolution of the drag coefficient of the twin rudders at 15° versus the lateral spacing. It is noticed that the drag of rudder 1 dominates between the two rudders, and its value increases as the lateral spacing increases. The drag of rudder 2 is relatively smaller than that of rudder 1 and reaches the smallest value at $d=1.2c$. The total drag of the twin rudders increase with the lateral spacing increasing.

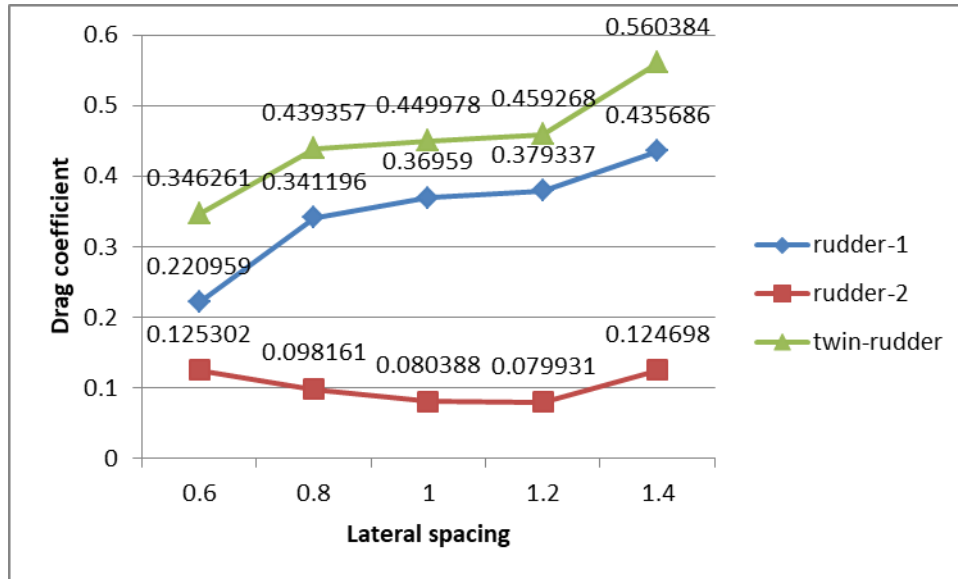


Figure 7. Drag coefficients of the twin rudders versus the lateral spacing.

Figure 8 shows the variation of the lift coefficients of the twin rudders at 15° versus the lateral spacing. Instead, for the lift, rudder 2 give rise to the larger values than rudder 1. As the lateral spacing increases, the lift of rudder 1 increases and the lift of rudder 2 decreases. The total lift coefficients of the twin rudders reaches the highest value at $d=1.0c$. Additionally, it is noticed that the value of the lift coefficients of the two rudders approaches to each other as the lateral spacing is 1.4, and the total lift is approximate to twice the lift of the single rudder, which indicates that the effect of the interaction between the two rudders on the lift coefficients becomes weak when the lateral spacing is large.

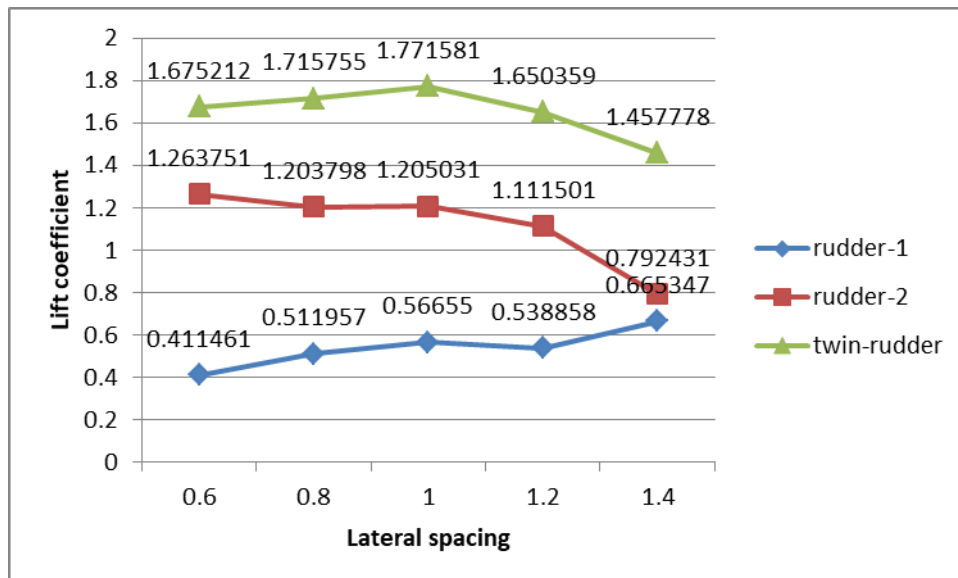


Figure 8. Lift coefficients of the twin rudders versus the lateral spacing.

The profile of the lift to drag ratio versus the lateral spacing is shown in Figure 9. It is clear that the lift to drag ratio of the twin rudders at 15° monotonously decreases with the lateral spacing increasing. According to the variation of the lift to drag ratio, it is suggested to have a reasonably small lateral spacing for a high lift to drag ratio in practical.

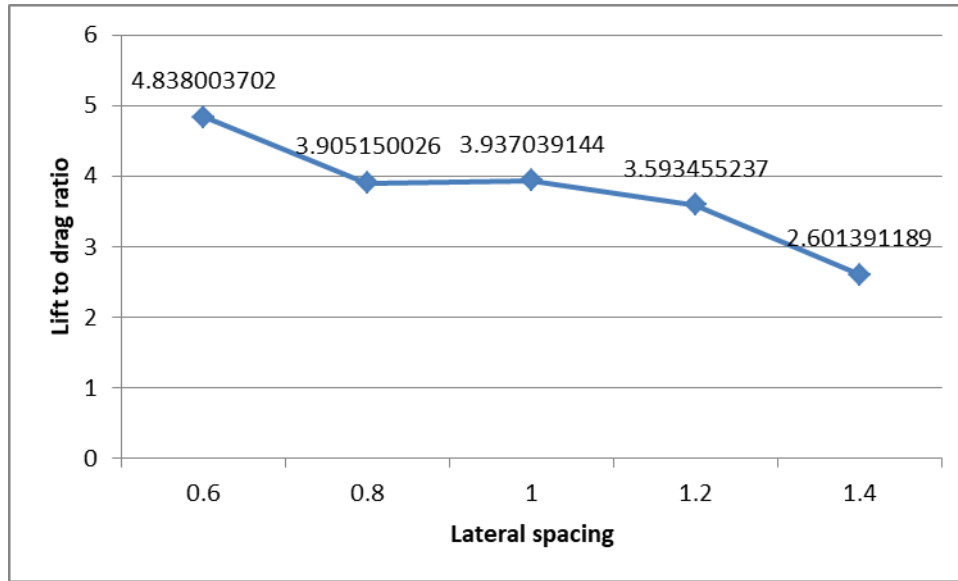


Figure 9. Lift to drag ratios of the twin rudders versus the lateral spacing.

For more details, the static pressure coefficients around the rudders are shown in Figure 10. It is shown that, for $d=0.6c$, the pressures in the front of the rudders are quite close, while the ones in the back of rudder 1 is much smaller than those in the back of rudder 2, which is due to the strong interaction between the two rudders. For $d=1.0c$, the pressure around the two rudders are still much different, which means that the interaction still makes much sense. As the lateral spacing approaches to $1.4c$, the difference of the pressure around the two rudders becomes small, which agrees to a statement that the interaction weakens as the lateral spacing increases.

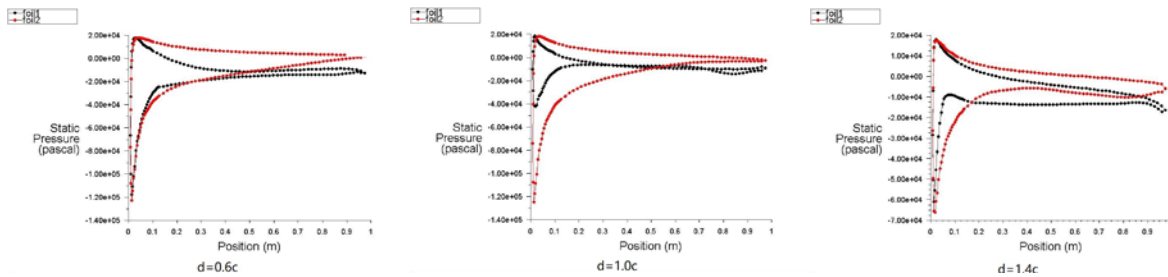


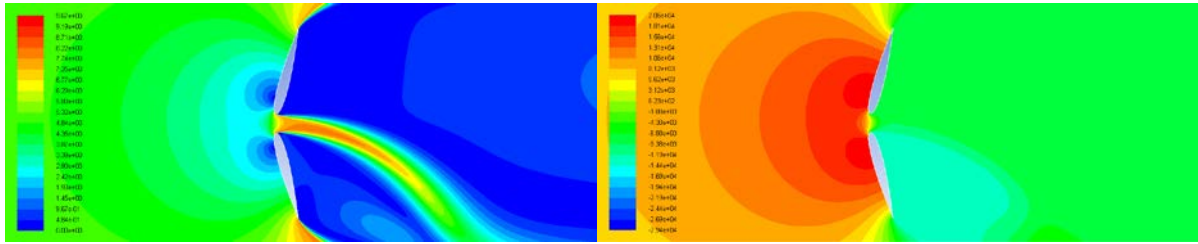
Figure 10. Static pressure coefficients around the rudders with the different lateral spacing.

Impact of the lateral spacing on stopping performance

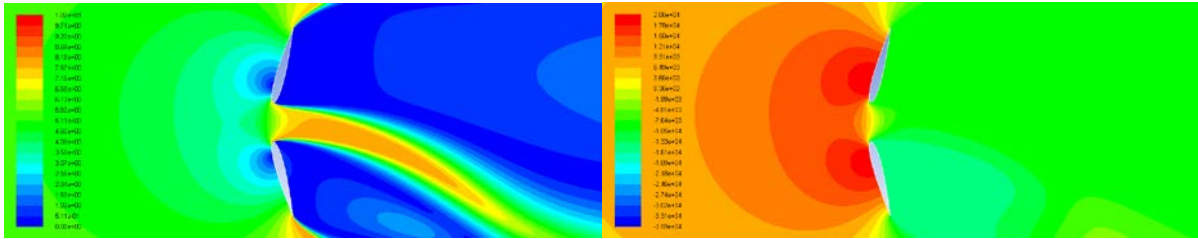
As the rudders are simultaneously set at a clam shell angle (75° outward) for a stopping performance, the gap between the leading edge of the twin rudders, which is determined by the lateral spacing, may affect the drag of the rudders as well as the stopping distance. Here flow around the twin rudders with the different lateral spacing at a clam shell angle are simulated. Figure 11 shows the flow and static pressure contours for $d=0.6c$, $1.0c$ and $1.4c$. It is noticed that the flow field downstream of the rudders are not symmetrical. For $d=0.6c$ and $1.0c$, the flow cross the gap with high velocity travels downwards downstream of the rudders, which causes that the pressure downstream of rudder 1 is larger than that downstream of rudder 2. As $d=1.4c$, the flow cross the gap travels upwards downstream of the rudders, namely, the pressure downstream of rudder 1 is smaller than that downstream of rudder 2. The flow cross

the gap would instantaneously vibrate. The asymmetry of the flow downstream of the rudders is likely due to the limitation of the numerical method used in the present study.

$d=0.6c$



$d=1.0c$



$d=1.4c$

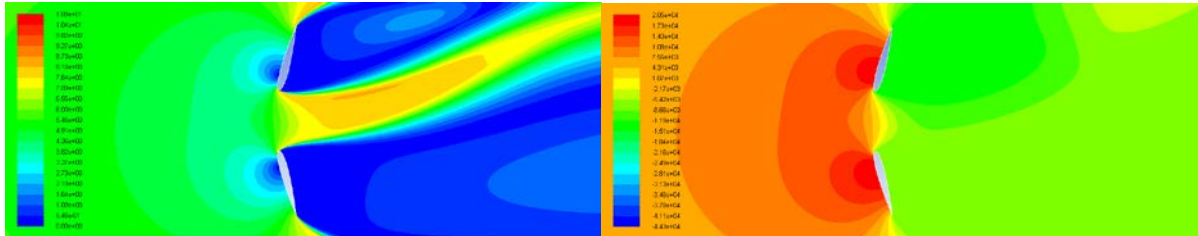


Figure 11. Flow and static pressure contours around the twin rudders with the different lateral spacing at a clam shell angle.

To examine the stopping performance of the twin rudders with the different lateral spacing, the drag coefficients of the twin rudders at a clam shell angle are tested. Figure 12 shows the variation of the drag coefficients of the twin rudders at outward versus the lateral spacing. It can be seen that the drag coefficients of the two rudders are different due to the flow asymmetry downstream of the rudders. The drag of rudder 1 is smaller than that of rudder 2, which agrees to the statement about pressure distribution addressed above. The drag of rudder 1 slowly increases as d increases, and reaches to the largest value at $d=1.3c$. Instead the drag of rudder 2 decreases as d increases, and reaches to the lowest point at $d=1.3c$. The drags of the two rudders approaches to each other at $d=1.4c$. This is because that a larger gap between the leading edge weakens the interaction of the flow downstream. The value of the total drags of the twin rudders are almost twice that of the single rudder, which indicates that the twin rudders make better sense than the single rudder during the stopping performance. It is shown that the variation of the total drag with the lateral spacing is small. It reaches the largest value at $d=1.3c$, and decreases as d is larger than 1.3.

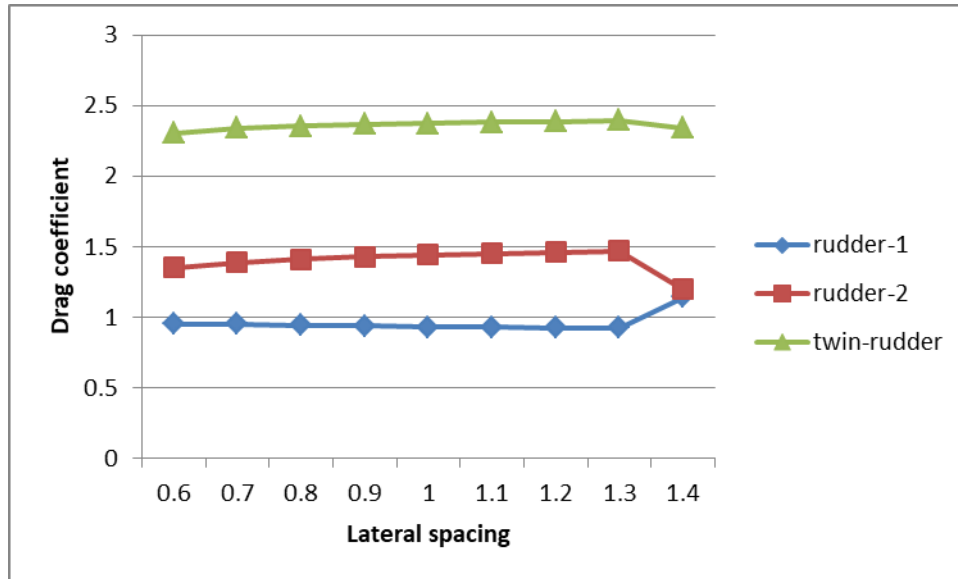


Figure 12. Drag coefficients of the twin rudders at a clam shell angle versus the lateral spacing.

Conclusions

Computational fluid dynamic simulations of the twin rudders are performed by solving Reynolds-averaged Navier-Stokes equations with a SST $k-\omega$ two-equation model using OpenFOAM codes. The difference of hydrodynamic forces between the single-rudder and the parallel twin rudders of NACA0012 blades are examined from 5° to 25° angle of attack. The impact of the distance between the two rudders on the hydrodynamic forces at 15° angle of attack is studied for the parallel twin rudders. Finally the effectiveness of the stopping performance of the twin rudders at the different lateral spacing is analyzed. According to the results and discussions, conclusions of the present study can be summarized as follows

- (1) The hydrodynamic forces of the twin rudders at all tested angles of attack are larger than those of the single rudder. The interaction between the twin rudders becomes strong as the angle of attack increases.
- (2) The interaction between the twin rudders becomes weak with the lateral spacing increasing. The drag of the twin rudders increases with the lateral spacing increasing. The lift reaches the largest value at $d=1.0c$. And the lift to drag ratio monotonously decreases with the lateral spacing.
- (3) The total drag of the twin rudders is almost close to twice that of the single rudder. The drag slowly increases with the lateral spacing, and decreases as the lateral spacing is larger than $1.3c$.

This study present some limited results due to the limitation of the numerical method used. To obtain more accurate flow data around the twin rudders, it is requested to use a transient solver or a large eddy simulation, which would be the further research of the authors.

References

- [1] Hamamoto, M., & Enomoto, T. (2009). Maneuvering performance of a ship with vectwin rudder system : *Journal of the Japan Society of Naval Architects & Ocean Engineers*, 1997(181), 197-204.

- [2] Nagarajan, V., Kang, D. H., Hasegawa, K., Nabeshima, K., & Arii, T. (2009). A proposal for propulsion performance prediction of a single-propeller twin-rudder ship. *Journal of Marine Science & Technology*, 14(3), 296-309.
- [3] Kang, D., Nagarajan, V., Gonno, Y., Uematsu, Y., Hasegawa, K., & Shin, S. C. (2011). Installing single-propeller twin-rudder system with less asymmetric maneuvering motions. *Ocean Engineering*, 38(10), 1184-1196.
- [4] Yoshimura, Y., & Sakurai, H. (1989). Mathematical model for the manoeuvring ship motion in shallow water (3rd report) : manoeuvrability of a twin-propeller twin rudder ship. *Journal of the Kansai Society of Naval Architects Japan*, 115-126.
- [5] Yoo, W. J., Yoo, B. Y., & Rhee, K. P. (2006). An experimental study on the maneuvering characteristics of a twin propeller/twin rudder ship during berthing and unberthing. *Ships & Offshore Structures*, 1(3), 191-198.
- [6] Khanfir, S., Hasegawa, K., Lee, S. K., Jang, T. S., Lee, J. H., & Cheon, S. J. (2008). 2008K-G4-3 Mathematical Model for Maneuverability and Estimation of Hydrodynamic Coefficients of Twin-Propeller Twin-Rudder Ship. Conference Proceedings *The Japan Society of Naval Architects and Ocean Engineers* (Vol.7, pp.57-60). The Japan Society of Naval Architects and Ocean Engineers.
- [7] Khanfir, S., Nagarajan, V., Shouji, K., & Lee, S. K. (2011). Manoeuvring characteristics of twin-rudder systems: rudder-hull interaction effect on the manoeuvrability of twin-rudder ships. *Journal of Marine Science & Technology*, 16(4), 472-490.
- [8] Shcherbakov, A. V. (2013). Assessment of flow characteristics around twin rudder with various gaps using piv analysis in uniform flow. *Ocean Engineering*, 66(3), 1-11.
- [9] Liu, J., & Hekkenberg, R. (2015). Hydrodynamic characteristics of twin-Rudders at small attack angles. *The International Marine Design Conference*.
- [10] Baudu H. (2014) Ship handling. 1st ed. Enkhuizen (The Netherlands): DOKMAR Maritime Publishers BV.
- [11] Hasegawa, K., Kang, D., Sano, M., & Nabeshima, K. (2006). Study on the maneuverability of a large vessel installed with a mariner type super vectwin rudder. *J Marine Science & Technology*, 11(2), 88-99.
- [12] Menter, F. R. (1994). Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA Journal*, 32(8), 1598-1605.

Winkler model for seismic responses of shafts under stochastic earthquakes

***Bu Zhang¹ and †Zhiyi Chen^{1,2}**

¹Department of Geotechnical Engineering, Tongji University, Shanghai 200092, China

²Key Laboratory of Geotechnical and Underground Engineering of Ministry of Education, Shanghai 200092, China

*Presenting author: zhangbu@tongji.edu.cn

†Corresponding author: zhiyichen@tongji.edu.cn

Abstract

A simplified method for calculating the seismic responses of the shaft is proposed in this paper. First, based on the theory of Winkler elastic foundation beam, the urban shaft is simplified as a vertical beam. Secondly, the horizontal soil reaction and vertical shear tractions between the shaft circumference and the surrounding soils are considered through horizontal springs and rotating springs on the sidewall of the shaft. The translation and rocking motion of the shaft are considered through horizontal springs and rotating springs at the bottom of the shaft. Then, the dynamic analysis model of the shafts under seismic motion is established, and the control equation of the dynamic response of the shaft in frequency domain is deduced. Finally, the analytical solution of the steady state response of the shaft is obtained. Considering the randomness of the earthquake motion, this method can get the shaft responses under different ground motions efficiently. At the same time, the influence of ground motion frequency on the dynamic response of shaft can be observed.

Keywords: Shaft, Winkler model, Seismic responses, Random

1. Introduction

As a subsidiary structure connecting the ground and underground structures, the vertical shaft has been widely constructed in the areas of underground transportation system, power system and utility tunnel system. At present, round and square are mainly cross section shape of existing shafts, which are usually of small cross-section and shallow depth within 40 meters. With the exploitation and utilization of deep underground space in urban cities, large-depth (more than 40 meters) shafts are now widely used in deep urban drainage systems. Such as the Metropolitan Area Outer Underground Discharge Channel in Tokyo and the Deep Storage and Drainage Pipe System in Shanghai. Round is the mainly cross section shape of the deep shaft. The dynamic responses of the shafts are being studied.

In order to know the shaft dynamic responses, there are about two methods to calculate the dynamic responses: the quasi-static method and the three dimensional dynamic time history analysis method. In the quasi-static methods the shafts are usually treated as a vertical beam

embedded in the soil [1]-[3]. In the dynamic time history analysis the results are more accurate. The main differences between the two methods is that the quasi-static method could not considering the shaft responses at every moment under the excitation of the earthquake motion, while the numerical methods such as three dimensional finite element dynamic time history analysis method and so on are usually with low computational efficiency and could not explain the dynamic responses from the perspective of mechanical mechanism for engineering design. Due to the uncertainty and complexity of the earthquake motion, the ground motion is a complicated time process which should be carefully considered when implementing the shaft seismic design. Therefore it is essential to propose a simplified method which can not only capture the dynamic responses of the shafts but also can calculate efficiently.

In this paper, a dynamic winkle beam physical model for the shaft is purposed and established with considering the horizontal reaction and vertical shear tractions from the surrounding soil and horizontal traction springs and rotating springs at the bottom of the shaft for the kinematic motion based on the existing quasit-static method. The dynamic equilibrium equation of the shaft is derived and the analytic solution of the kinematic equilibrium is solved and presented in frequency domain. The dynamic responses of the shafts under stochastic earthquakes excitation in time domain would be obtained efficiently through the FFT and iFFT with this simplified method.

2. Physical model

In the physical model of quasi-static method [3], horizontal reactions and vertical shear tractions distributed along the shaft wall are represent by the horizontal spring and the rotation spring. In order to reflect the translation motion and rotational motion, the shear spring and rotation spring are stetted at the shaft bottom.

Gerolymos and Gazetas [4]-[6] purposed a winkle four spring model for lateral response of rigid caisson foundations in liner soil. Due to the structural and functional characteristics of the caisson foundation, the caissons are usually simplified as a rigid body due to its great structural stiffness relative to the surrounding soil, while the pile foundations are usually simplified as a beam due to the small structural stiffness relative to the surrounding soil. There are many differences between the caisson foundation and the deep shaft, especially the structural stiffness, the underground deep shafts are hollow structures while the caisson foundations are solid one. Chen and Zhang [7] and Mayoral [8] concluded that the large-depth shaft dynamic responses in soft soil approximately like a rigid body with translation motion, rotational motion and small bending deformation. Considering the structural and functional characteristic of the shaft and effect of the soil-structure relative stiffness on the dynamic responses of the underground structure. Finally, the shaft is simplified as a beam. In this paper the simplified dynamic analysis method for the shaft is proposed with four springs and dashpots and the shaft is simplified a beam, as shown in Fig. 1

A circle in plain shaft embedded in homogeneous isotropic and viscoelastic soil was illustrated in Fig. 1. The assumption of model are as follows: the shaft is assumed to be with uniform wall thickness and the shaft with linear deformation under the seismic motion; the shaft is perfectly

glued to the surrounding soils, indicating that there is no slippage or separation along the shaft-soil interface.

The parameters of the shaft are as follows: the shaft depth is L , the external diameter is D , the inner diameter is d , the Young's modulus is E , and the density is ρ . The parameters of the soil are as follows: the Young's modulus is E_s , the shear modulus is G_s , the density is ρ_s , the shear wave velocity is V_s . The four-spring coefficients purposed by Gazetas [9] is adopted for simulating the soil-caisson interaction here and will be introduced in the next session. The distributed lateral springs k_x and dashpots c_x and rotation springs k_θ and dashpots c_θ are the simplified horizontal soil reactions and vertical shear tractions, while the shear translation spring k_{bx} and dashpot c_{bx} and base rotation spring $k_{b\theta}$ and dashpot $c_{b\theta}$ are the simplified horizontal shearing force and the moment produced by the base of the shaft.

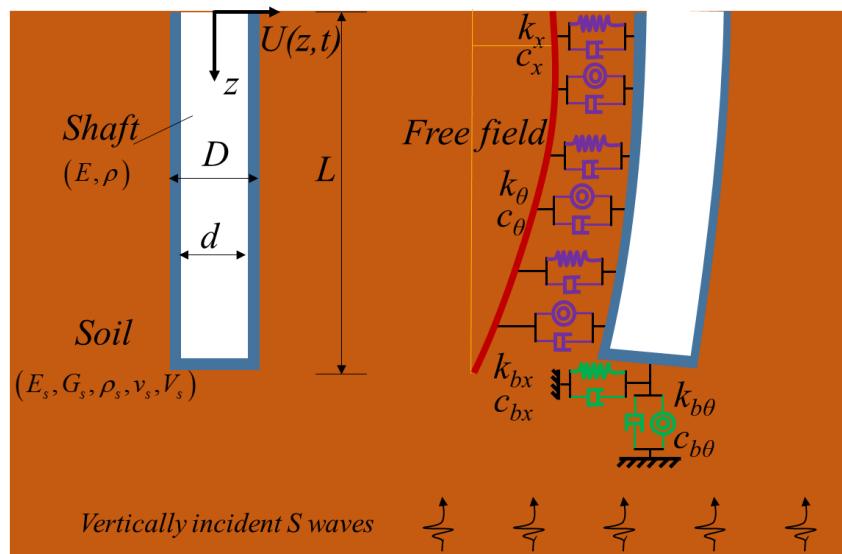


Figure 1. Schematic diagram of shaft under vertically incident S waves

3. Explicit representation of the model

3.1 Shaft kinematic responses equation

In order to establish the differential equation of the shaft kinematic responses, a shaft element is chosen from the physical model in Fig. 1. Fig. 2 illustrates the state of the beam element in a viscoelastic soil under the seismic loads. The main loads in horizontal direction are inertia force, the soil reaction force and the shear forces from the connecting shaft elements. The main moment loads are inertia moment, the moment from the soil vertical tractions, the soil reaction moment and the moment from the connecting shaft elements.

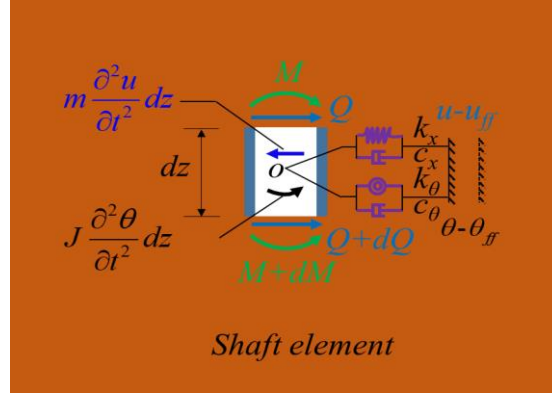


Figure 2. Schematic diagram of force acting on the shaft element under seismic motion

According to the dynamic equilibrium of the transverse forces and the dynamic moment equilibrium with respect to the central point O of the shaft element. The two differential governing equations for the shaft kinematic responses element in time domain can be expressed as

$$m \frac{\partial^2 u}{\partial t^2} dz + k_x (u - u_{ff}) dz + c_x \frac{\partial (u - u_{ff})}{\partial t} dz = -dQ \quad (1)$$

$$J \frac{\partial^2 Q}{\partial t^2} dz + k_\theta (\theta - \theta_{ff}) dz + c_\theta \frac{\partial (\theta - \theta_{ff})}{\partial t} dz = -dM + Q dz - \tau \frac{\pi D^2}{4} dz \quad (2)$$

where m is the mass of the shaft per length. J is the moment of inertia of the shaft per length. $u(z, t)$ is the displacement of the shaft central. dz is the length of the shaft element. u_{ff} is the displacement of the free field along the depth. The formulate of u_{ff} in frequency domain can be expressed as

$$u_{ff}(w, z) = u_{ff0} \cos\left(\frac{w}{V_s} z\right) \quad (3)$$

where u_{ff0} is the displacement of the soil surface. w is the circular frequency. z is the vertical coordinate starting from the top central of the shaft. θ_{ff} is the rotation angle of the free field. The formulate of θ_{ff} in frequency domain can be expressed as

$$\theta_{ff}(w, z) = -\frac{du_{ff}}{dz} = u_{ff0} \left(\frac{w}{V_s}\right) \sin\left(\frac{w}{V_s} z\right) \quad (4)$$

The bending moment M at the shaft cross section can be expressed as

$$M = -EI \frac{d^2 u}{dz^2} \quad (5)$$

where I is the area moment of inertia of the shaft cross section. Q is the shear force at the shaft cross section

$$Q = -EI \frac{d^3 u}{dz^3} \quad (6)$$

θ is the rotation angle of the shaft along the shaft depth

$$\theta = -\frac{du}{dz} \quad (7)$$

Submitting Eq. (1) in Eq. (2), one obtains the final dynamic equilibrium equation of the shaft element in frequency domain

$$EI \frac{d^4 u}{dz^4} + (J\omega^2 + i\omega c_\theta + k_\theta) \frac{d^2 u}{dz^2} + (-m\omega^2 - i\omega c_x - k_x)u = \begin{pmatrix} (i\omega c_\theta + k_\theta) \frac{d^2 u_{ff}}{dz^2} \\ +(-i\omega c_x - k_x)u_{ff} \end{pmatrix} \quad (8)$$

For express simplification, then name the equation coefficients as A_c , B_c , C_c and D_c .

$$A_c = EI \quad (9)$$

$$B_c = J\omega^2 + i\omega c_\theta + k_\theta \quad (10)$$

$$C_c = -m\omega^2 - i\omega c_x - k_x \quad (11)$$

$$D_c = \left[(i\omega c_\theta + k_\theta)u_{ff0} \left(-\frac{\omega^2}{V_s^2} \right) + (-i\omega c_x - k_x)u_{ff0} \right] \cos\left(\frac{\omega}{V_s} z\right) \quad (12)$$

Then the dynamic equilibrium Eq. (8) could be simplified as

$$A_c \frac{d^4 u}{dz^4} + B_c \frac{d^2 u}{dz^2} + C_c u = D_c \quad (13)$$

The analytic solution the Eq. (8) in frequency domain can be expressed as

$$u(w, z) = C_1 e^{r_1 z} + C_2 e^{r_2 z} + C_3 e^{r_3 z} + C_4 e^{r_4 z} + E_c \cos\left(\frac{\omega}{V_s} z\right) \quad (14)$$

Where C_1 , C_2 , C_3 and C_4 are the four underdetermined parameters, which can be obtained through the boundary conditions. The parameters of r can be obtained as

$$\begin{cases} r_{1,2} = \pm \sqrt{\frac{-B_c - \sqrt{B_c^2 - 4A_c C_c}}{2A_c}} \\ r_{3,4} = \pm \sqrt{\frac{-B_c + \sqrt{B_c^2 - 4A_c C_c}}{2A_c}} \end{cases} \quad (15)$$

The coefficients of the specific solution E_c can be expressed as

$$E_c = \frac{D_c}{C_c - B_c \left(\frac{w}{V_s}\right)^2 + \left(\frac{w}{V_s}\right)^4} \quad (16)$$

The boundary condition about the shaft top is free at the shaft head and constraint at the bottom, then the boundary conditions can be formulated as follows

$$\begin{cases} M|_{z=0} = -EI \frac{d^2 u}{dz^2} \Big|_{z=0} = 0 \\ Q|_{z=0} = -EI \frac{d^3 u}{dz^3} \Big|_{z=0} = 0 \end{cases} \quad (17)$$

$$\begin{cases} M|_{z=l} = -EI \frac{d^2 u}{dz^2} \Big|_{z=l} = (k_{b\theta} + iwc_{b\theta}) \left(-\frac{du}{dz} \Big|_{z=l} \right) \\ Q|_{z=l} = -EI \frac{d^3 u}{dz^3} \Big|_{z=l} = (k_{bx} + iwc_{bx}) u|_{z=l} \end{cases} \quad (18)$$

Submitting the Eq. (14) into the Eq. (17) and (18), then the four undetermined parameters C_1 , C_2 , C_3 and C_4 can be obtained by the matrix as follows

$$\begin{aligned}
 & \begin{bmatrix} r_1^2 & r_2^2 & r_3^2 & r_4^2 \\ r_1^3 & r_2^3 & r_3^3 & r_4^3 \\ \left(r_1^2 - \frac{k_r + iwc_r}{EI}\right)e^{r_1 l} & \left(r_2^2 - \frac{k_r + iwc_r}{EI}\right)e^{r_2 l} & \left(r_3^2 - \frac{k_r + iwc_r}{EI}\right)e^{r_3 l} & \left(r_4^2 - \frac{k_r + iwc_r}{EI}\right)e^{r_4 l} \\ \left(r_1^3 + \frac{k_h + iwc_h}{EI}\right)e^{r_1 l} & \left(r_2^3 + \frac{k_h + iwc_h}{EI}\right)e^{r_2 l} & \left(r_3^3 + \frac{k_h + iwc_h}{EI}\right)e^{r_3 l} & \left(r_4^3 + \frac{k_h + iwc_h}{EI}\right)e^{r_4 l} \end{bmatrix} \\
 & \begin{bmatrix} E_c \left(\frac{w^2}{V_s^2}\right) \\ 0 \\ E_c \left[\left(\frac{k_r + iwc_r}{EI}\right) \left(\frac{-w}{V_s}\right) \sin\left(\frac{w}{V_s} L\right) - \left(\frac{-w^2}{V_s^2}\right) \cos\left(\frac{w}{V_s} L\right) \right] \\ -E_c \left[\left(\frac{k_h + iwc_h}{EI}\right) \cos\left(\frac{w}{V_s} L\right) + \left(\frac{w^3}{V_s^3}\right) \sin\left(\frac{w}{V_s} L\right) \right] \end{bmatrix} \quad (19)
 \end{aligned}$$

Finally, according to the Eqs. (5), (6), (7) and (14), the rotation angle, the bending moment and the shear force of the shaft along the depth in frequency domain can be obtained as follows

$$\theta(w, z) = - \left[C_1 r_1 e^{r_1 z} + C_2 r_2 e^{r_2 z} + C_3 r_3 e^{r_3 z} + C_4 r_4 e^{r_4 z} + E_c \left(-\frac{w}{V_s} \right) \sin\left(\frac{w}{V_s} z\right) \right] \quad (20)$$

$$M(w, z) = -EI \left[C_1 r_1^2 e^{r_1 z} + C_2 r_2^2 e^{r_2 z} + C_3 r_3^2 e^{r_3 z} + C_4 r_4^2 e^{r_4 z} + E_c \left(-\frac{w^2}{V_s^2} \right) \cos\left(\frac{w}{V_s} z\right) \right] \quad (21)$$

$$Q(w, z) = -EI \left[C_1 r_1^3 e^{r_1 z} + C_2 r_2^3 e^{r_2 z} + C_3 r_3^3 e^{r_3 z} + C_4 r_4^3 e^{r_4 z} + E_c \left(\frac{w^3}{V_s^3} \right) \sin\left(\frac{w}{V_s} z\right) \right] \quad (22)$$

3.2 Spring and dashpot coefficients

The four spring and dashpot coefficients are adopted from the coefficients which are proposed by Gazetas [9] and Gerolymos [6] and revised by Zhong [10]. They have done the calibration of the spring and dashpot coefficients with Novak, Varun and Wolf and verified that the coefficients perform well. This coefficients are related with the soil parameters and geometric parameters of shaft. The expression of the lateral horizontal spring coefficients k_x is as follows

$$k_x = \left(\frac{I_{tw} \chi_{emb} - 1}{L} \right) \frac{2.02 E_s D}{(2 - \nu_s)(1 - \nu_s)} \quad (23)$$

where I_{tw} is the horizontal embedment factor of a cylindrical shaft, χ_{emb} is the dynamic coefficient.

$$I_{tw} = 1 + 0.21 \left(\frac{L}{D} \right)^{0.5} + 1.43 \left(\frac{L}{D} \right)^{0.8} + 0.3 \left(\frac{L}{D} \right)^{1.3} \quad (24)$$

$$\chi_{emb} = 1 + \left(\frac{wD}{2V_s} \right) \left(\frac{L}{D} \right) \left[\left(0.08 - 0.0074 \frac{L}{D} \right) \left(\frac{wD}{2V_s} \right)^2 - \left(0.31 - 0.0416 \frac{L}{D} \right) \left(\frac{wD}{2V_s} \right) \right] - 0.442 \frac{L}{D} + 0.14 \quad (25)$$

The expression of the lateral horizontal dashpot coefficients c_x is as follows

$$c_x = \left(2 + \frac{2.16}{1 - \nu_s} \right) \frac{E_s}{1 + \nu_s} \left(\frac{D}{2V_s} \right) \quad (26)$$

The expression of the lateral rotation spring coefficients k_θ is as follows

$$k_\theta = \frac{(\Gamma_w - 1)(1 - 0.3a_0)}{L} \frac{0.16E_s D^3}{(1 - \nu_s)(1 + \nu_s)} - \frac{1}{3} L^2 \frac{(I_{tw} \chi_{emb} - 1)}{L} \frac{2.02E_s D}{(2 - \nu_s)(1 + \nu_s)} \quad (27)$$

where Γ_w is the rocking embedment factor of a cylindrical shaft.

$$\Gamma_w = 1 + 2.25 \left(\frac{L}{D} \right)^{0.6} + 7.01 \left(\frac{L}{D} \right)^{2.5} \quad (28)$$

The expression of the lateral rotation dashpot coefficients c_θ is as follows

$$c_\theta = \left[\left[\left(\frac{2}{3} + \frac{0.72}{1 - \nu_s} \right) L^2 + \frac{1}{2} D^2 \right] \frac{E_s}{1 + \nu_s} \left[0.25 + 0.65 \left(\frac{L}{D} \right)^{-0.25} \sqrt{\frac{wD}{2V_s}} \right] - \frac{1}{3} L^2 \left(2 + \frac{2.16}{1 - \nu_s} \right) \frac{E_s}{(1 + \nu_s)} \right] \left(\frac{D}{2V_s} \right) \quad (29)$$

The expression of the base shear spring coefficient k_{bx} is as follows

$$k_{bx} = \frac{2.02E_s D}{(2 - \nu_s)(1 - \nu_s)} \quad (30)$$

The expression of the base shear dashpot coefficient c_{bx} is as follows

$$c_{bx} = \frac{0.79E_s D}{1 + \nu_s} \left(\frac{D}{2V_s} \right) \quad (31)$$

The expression of the base shear spring coefficient $k_{b\theta}$ is as follows

$$k_{b\theta} = \frac{0.16E_s D^3}{(1 - \nu_s)(1 + \nu_s)} \left(1 - 0.3 \frac{wD}{2V_s} \right) \quad (32)$$

The expression of the base rotation dashpot coefficient $c_{b\theta}$ is as follows

$$c_{b\theta} = \frac{0.05E_s D^3}{(1 - \nu_s)(1 + \nu_s)} \left(\frac{D}{2V_s} \right) \quad (33)$$

4. Solution technique

Inspect from the above analytical equations, the input parameters u_{ff0} should be in frequency domain. In order to know the transient seismic responses of the shaft, the input parameters u_{ff0} should be transformed into frequency domain which can be achieved by conducting FFT and the analytical results u and so on should be transformed into time domain which can be achieved by conducting iFFT. There are total three steps to obtain the dynamic responses of the shaft. The flowchart of the solve procedure is shown in Fig. 3.

The first step is to transform the time history of the input ground motion u_{ff0} into frequency domain through the FFT method. The ground motion u_{ff0} in frequency domain will be obtained with the corresponding frequency w . From the matrix Eq. (19) the coefficients C_1 , C_2 , C_3 and C_4 can be obtained under the corresponding frequency w .

The second step is to obtain the shaft dynamic responses along the depth, such as: shaft displacement, rotational angle, bending moment and shear force in frequency domain through the Eq. (14), (20), (21) and (22). At the same time, the influence of ground motion frequency on the dynamic response of shaft can be observed.

The third step is to transform the shaft's frequency domain dynamic response parameters into time domain by implementing iFFT.

Repeat the above operations, then the shaft dynamic responses under stochastic earthquakes can be achieved with this simplified method. All this procedures can be implemented by MATLAB software efficiently.

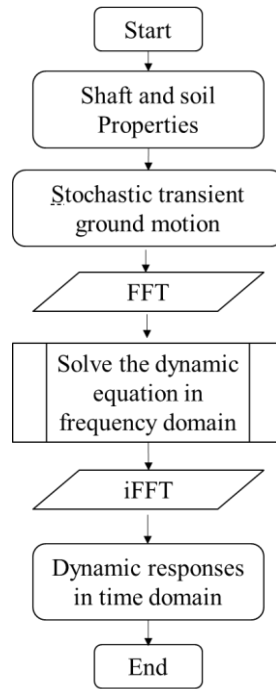


Figure 3. Flowchart of the solve procedure

5. Conclusion

Based on the theory of dynamic Winkle beam on elastic foundation, the shaft dynamic model is proposed and the closed-form solution for the dynamic responses of the shaft is established in frequency domain with explicit expression. A simplified model for calculating the seismic responses of the shaft under stochastic earthquakes is proposed and established. The seismic responses of shaft under stochastic earthquakes would be easily solved with the FFT and iFFT method between the frequency domain and the time domain. The simplified dynamic Winkle beam model could capture and properly reflect the shaft's dynamic kinematic responses, translation, rotational motion and bending curve along the depth with low computational cost compared with the three dimensional dynamic time history analysis.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (Grant No. 41472246, 51778464), and “Shuguang Program” supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission. All supports are gratefully acknowledged.

References

- [1] Kunichika, M. and Sato, T. (2004) An earthquake resistant design methodology for divisional shaft based on the response displacement method, *Japanese Society for Natural Disaster Science* **23**, 381-390.
- [2] Kunichika, M. and Sato, T. (2004) Evaluation of seismic capacities of divisional shafts with different assembly methods, *Japanese Society for Natural Disaster Science* **23**, 391-403.
- [3] Kawashima, K., (1994) *Seismic design of underground structures*, Deer Island Press. (In Japanese)

- [4] Gerolymos, N. and Gazetas, G. (2006) Development of Winkler model for static and dynamic response of caisson foundations with soil and interface nonlinearities, *Soil Dynamics and Earthquake Engineering* **26**, 363–76.
- [5] Gerolymos, N. and Gazetas, G. (2006) Static and dynamic response of massive caisson foundations with soil and interface nonlinearities—validation and results. *Soil Dynamics and Earthquake Engineering* **26**, 377–94.
- [6] Gerolymos, N. and Gazetas, G. (2006) Winkler model for lateral response of rigid caisson foundations in linear soil, *Soil Dynamics and Earthquake Engineering*; **26**, 347–61.
- [7] Chen, Z.Y. and Zhang, B. Seismic Responses of the Large-Scale Deep Shaft in Shanghai Soft Soils, *Proceedings of GeoShanghai 2018 International Conference*, Shanghai, China, 2018, Advances in Soil Dynamics and Foundation Engineering. GSIC 2018. Springer, Singapore
- [8] Mayoral, J.M., Argyroudis, S. and Castañón, E. (2016) Vulnerability of floating tunnel shafts for increasing earthquake loading, *Soil Dynamics & Earthquake Engineering* **80**, 1-10.
- [9] Gazetas, G. (1991) Formulas and Charts for Impedances of Surface and Embedded Foundations, *Journal of Geotechnical Engineering* **117**, 1363-1381.
- [10] Zhong, R. and Huang, M. (2014) Winkler model for dynamic response of composite caisson–piles foundations Seismic response, *Soil Dynamics & Earthquake Engineering* **66**, 241-251.

Optimization of Left Ventricle Pace Maker Location Using Echo-Based Fluid-Structure Interaction Models

*§ Longling Fan¹, §Jing Yao², Chun Yang³, Di Xu², † Dalin Tang^{1,4}

¹School of Mathematics, Southeast University, Nanjing, 210096, China.

² Department of Cardiology, First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China.

³ China Information Tech. Designing & Consulting Institute Co., Ltd., Beijing, 100048, China.

⁴ Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609.

*Presenting author: fan_longling2008@126.com

§ These authors contributed equally to this work,

†Corresponding authors. Dalin Tang, email: dtang@wpi.edu

Abstract

Cardiac pacing has been an effective treatment in the management of patients with bradyarrhythmia and tachyarrhythmia. Different pacemaker location has different responses, and pacemaker effectiveness to each individual can also be different. A novel image-based ventricle animal modeling approach was proposed to optimize ventricular pacemaker site for better cardiac outcome. One health female adult pig weight 42.5kg was used to make pacing animal model with different ventricle pacing locations. Pig health status was assessed before undergoing experimental procedures. Ventricle surface electric signal, blood pressure and echo image were acquired 15 minutes after the pacemaker was implanted. Echo-based left ventricle (LV) fluid-structure interaction (FSI) models were constructed to perform ventricle function analysis and investigate impact of pacemaker location on cardiac outcome. The nonlinear Mooney-Rivlin model was used for ventricle tissue material model. With the measured electric signal map from the pig associated with the actual pacemaker site, electric potential conduction of myocardium was modeled by material stiffening and softening in our model, with stiffening simulating contraction and softening simulating relaxation. Material stiffness parameters were adjusted in a cardiac cycle to match Echo-measured LV deformation and volume variations. Mapping between material stiffness and ventricle electric signal was quantified using data measured from the animal with pacemaker applied. Ventricle model without pacemaker and three ventricle models with the following pacemaker locations were simulated: right ventricular apex (RVA), posterior interventricular septum (PIVS) and right ventricular outflow tract (RVOT). Data for ventricle volume change, ejection fraction, stress and strain, flow velocity and shear stress data were collected for comparisons. Our results demonstrating that PIVS pacing model had higher peak flow velocity and stress/strain. It indicated PIVS pacemaker site may be the best location. This modeling approach could be used as “virtual surgery” to try various pacemaker locations and avoid risky and dangerous surgical experiments on real patients.

Keywords: Fluid-structure interaction model, pacemaker electrical conduction, Fluid dynamic, ventricle material properties, ventricle mechanics.

Introduction

In recent decades, rapid development of cardiac pacing has become the only effective treatment for slow cardiac arrhythmia. According to some statistics, between 1993 and 2009, 2.9 million patients received permanent pacemakers in the United States [1]. China Ministry

of Health Online Registration indicated that pacemaker implants were placed in 70,000 patients in 2016, and the number has been increasing year by year [2]. Right ventricular apex (RVA) has been the conventional location for pacemaker lead placement. However, RVA pacing is associated with abnormal myocardial contractile pattern, hemodynamic disorder, and histologic remodeling [3]. The review by Tops et al. provided a contemporary overview of the available evidence on the detrimental effects of RVA pacing [4]. So optimization of right ventricular pacing site becomes an important object of pacing electrophysiology. In recent years, the concept of physiological pacing has been proposed in the field of electrophysiological, and the study of the selection of pacing sites has received great attention [2] [5]-[6]. Singh et al. assessed left ventricular (LV) function and dyssynchrony in patients with right ventricle outflow tract (RVOT) pacing and conventional RVA pacing using equilibrium radionuclide angiography. Their results indicated RVOT pacing may lead to better preservation of LV function on longer follow-up [7]. Kronborg et al. showed that His or para-His pacing preserves LV ejection fraction and mechanical synchrony compared with right ventricular (RV) septal pacing in patient with atrioventricular block and may be a future pacing strategy to prevent pacing-induced heart failure in selected pacemaker patients [8]. Here His indicates His bundle which is a collection of heart muscle cells specialized for electrical conduction. As part of the electrical conduction system of the heart, it transmits the electrical impulses from the atrioventricular node (located between the atria and the ventricles) to the point of the apex of the fascicular branches via the bundle branches. Sharma et al. assessed the safety, feasibility, and success rates of His-bundle pacing in unselected patients without the use of a mapping catheter or a backup RV lead as compared to RVA pacing [9]. Zanon et al. systematically investigated the hemodynamic benefit of multipoint pacing performed at many pacing sites per heart and related hemodynamic effect to both LV electrical delay and the reduction in QRS duration [10].

Recent advances in computational modeling, methods and computer technology have made it possible for computer-simulated procedures to be used in clinical decision-making for diseased hearts. In our previous studies, we introduced patient-specific cardiac magnetic resonance (CMR)-based LV/RV models with fluid-structure interactions (FSI) with various surgical design and potential applications [11]-[14]. Echo-based 3D LV FSI models were introduced to perform ventricle mechanical analysis and investigate flow behaviors [15].

This paper will integrate echocardiography images, propagating dynamic electric potential on ventricle surface induced by pacemaker, and computational models with fluid-structure interactions to perform myocardial function and intra-cardiac flow assessment. The models will be used to evaluate and optimize pacemaker location.

Methods

3D echo data acquisition

The animal study was conducted at the First Affiliated Hospital of Nanjing Medical University, Nanjing, China. A health female adult pig weight 42.5kg was intubated and mechanically ventilated. Anesthesia was maintained using isoflurane. The pig was placed on an operating table in the semi-left lateral position with upright tilt, suitable for echocardiographic examination. Electric potential data recording and image acquisition were started 15 minutes after different pacemaker was implanted. Pacemaker locations included RVA, posterior interventricular septum (PIVS) and RVOT. Standard echocardiograms were obtained using an ultrasound machine (E9, GE Mechanical Systems, Milwaukee, Wisconsin) with a 3V probe. Electrophysiological recorder records body surface 12-lead electrocardiogram and intracardiac electrogram. Meantime, the pressure gauge catheter was

connected to the Medtronic Lifepark12 monitor. The left ventricular pressure curve was measured before and during the time period when the pacemaker was implanted. Table 1 gives basic information including ventricular pacing location, pressure and volume data.

Table 1. Ventricular pacing location and volume data

Pacemaker location	Non-pacemaker		RVA		RVOT		PIVS	
Pressure (mmHg)	Min= 9	Max= 102	Min= 10	Max= 119	Min= 8	Max= 90	Min= 7	Max= 83
Echo Vol (ml)	Min= 25	Max= 54	Min= 27	Max= 55	Min= 26	Max= 50	Min= 19	Max= 44
Echo EF (%)	53.70		50.91		48.00		56.82	
Model Vol (ml)	Min= 24.96	Max= 54.01	Min= 27.07	Max= 54.96	Min= 26.05	Max= 50.16	Min= 18.98	Max= 44.02
Model EF (%)	53.79		50.75		48.07		56.88	

The fluid-structure interaction model of LV

Blood flow in the left ventricle was assumed to be laminar, Newtonian, viscous and incompressible. The Navier-Stokes equations with arbitrary Lagrangian-Eulerian (ALE) formulation were used as the governing equations. When the inlet or outlet were closed, flow velocity was set to zero and pressure was left unspecified. When the inlet or outlet was open, flow velocity was left unspecified and pressure was prescribed. No-slip boundary conditions and natural force boundary conditions were specified at all interfaces to couple fluid and structure models together [11][16]. Standard governing equations and boundary conditions for the LV model were given:

$$\rho(\partial \mathbf{u} / \partial t + ((\mathbf{u} - \mathbf{u}_g) \cdot \nabla) \mathbf{u}) = -\nabla p + \mu \nabla^2 \mathbf{u}, \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (2)$$

$$\mathbf{u}|_{\Gamma} = \partial \mathbf{x} / \partial t, \quad (3)$$

$$P|_{\text{inlet}} = p_{\text{in}}(t), \partial \mathbf{u} / \partial n|_{\text{inlet}} = 0, \mathbf{u}|_{\text{outlet}} = 0, \text{ (filling phase)}, \quad (4)$$

$$P|_{\text{outlet}} = p_{\text{out}}(t), \partial \mathbf{u} / \partial n|_{\text{outlet}} = 0, \mathbf{u}|_{\text{inlet}} = 0, \text{ (ejection phase)}, \quad (5)$$

$$\sigma_{ij} \cdot \mathbf{n}_j|_{\text{out_wall}} = 0, \quad (6)$$

$$\sigma_{ij}^r \cdot \mathbf{n}_j^r|_{\text{interface}} = \sigma_{ij}^s \cdot \mathbf{n}_j^s|_{\text{interface}}, \quad (7)$$

where \mathbf{u} and p are flow velocity and pressure, \mathbf{u}_g is mesh velocity, μ is the viscosity of blood. Γ stands for LV inner wall, $f_{\bullet,j}$ stands for derivative of f with respect to the j th variable (or time t), σ^r and σ^s are fluid and structure stress tensors, and \mathbf{n}^r and \mathbf{n}^s are their outward normal directions, respectively.

The ventricle material tissue was assumed to be hyperelastic, anisotropic, homogeneous and nearly-incompressible. The governing equations for the LV structure model were:

$$\rho v_{i,tt} = \sigma_{ij,j}, i, j = 1, 2, 3; \text{ sum over } j, \quad (8)$$

$$\varepsilon_{ij} = (v_{i,j} + v_{j,i} + v_{\alpha,i} v_{\alpha,j}) / 2, i, j, \alpha = 1, 2, 3, \quad (9)$$

where σ is the stress tensor, ε is the strain tensor, v is displacement, and ρ is material density. The normal stress was assumed to be zero on the outer (epicardial) LV surface and equal to the normal stress imposed by fluid forces on the inner (endocardial) LV surface as specified by Eq.(7).

The nonlinear Mooney-Rivlin model was used to describe the nonlinear anisotropic material properties. The strain energy function for the anisotropic modified Mooney-Rivlin model is given:

$$W = c_1(I_1 - 3) + c_2(I_2 - 3) + D_1[\exp(D_2(I_1 - 3)) - 1] + (K_1/K_2)\exp[K_2(I_4 - 1)^2 - 1], \quad (10)$$

where I_1 and I_2 are the first and second strain invariants given by,

$$I_1 = \sum C_{ii}, \quad I_2 = \frac{1}{2}[I_1^2 - C_{ij}C_{ij}], \quad I_4 = C_{ij}(\mathbf{n}_f)_i(\mathbf{n}_f)_j \quad (11)$$

$C = [C_{ij}] = \mathbf{X}^T \mathbf{X}$ is the right Cauchy-Green deformation tensor, $\mathbf{X} = [X_{ij}] = [\partial x_i / \partial a_j]$, (x_i) is the current position, (a_i) is the original position, \mathbf{n}_f is the fiber direction, c_i , D_i and K_i are material parameters chosen to match experimental measurements [12][15]. With parameters properly chosen, it was shown that stress-strain curves derived from Eq. (10) agreed very well with the stress-strain curves from the anisotropic (transversely isotropic) strain-energy function with respect to the local fiber direction given in McCulloch et al.[14]:

$$W = \frac{C}{2}(e^Q - 1), \quad (12)$$

$$Q = b_1 E_{ff}^2 + b_2 (E_{cc}^2 + E_{rr}^2 + E_{cr}^2 + E_{rc}^2) + b_3 (E_{fc}^2 + E_{cf}^2 + E_{fr}^2 + E_{rf}^2), \quad (13)$$

where E_{ff} is fiber strain, E_{cc} is cross-fiber in-plane strain, E_{rr} is radial strain, and E_{cr} , E_{fr} and E_{fc} are the shear components in their respective coordinate planes, C , b_1 , b_2 , and b_3 are parameters to be chosen to fit experimental data. For simplicity, we set $b_1=0.8552$, $b_2=1.7005$, $b_3=0.7742$ in Eq. (12) so that we can have a single parameter C for comparison. The least-squares method was used to find the equivalent Young's moduli (YM) for the material curves for easy comparison.

As patient-specific fiber orientation data was not available from these patients, we chose to construct a two-layer LV model and set fiber orientation angles using fiber angles given in Axel [15]. Fiber orientation angles were set at -60 degree and 80 degree for epicardium (outer layer) and endocardium (inner layer), respectively. Fiber orientation can be adjusted when patient-specific data becomes available [11].

A pre-shrink process and geometry-fitting technique for mesh generation

Under in vivo condition, ventricles are pressurized and the zero-stress ventricular geometries are not known. In our model construction process, a pre-shrink process was applied to in vivo end-systolic ventricular geometries to generate the starting shape for the computational simulation [15]. A geometry-fitting mesh generation technique was also used to generate mesh for our models [13]. Mesh analysis was performed by decreasing mesh size by 10% (in each dimension) until solution differences were less than 2%. The mesh was then chosen for our simulations.

Solution methods and Data collection for Statistical analysis

The Echo-based anisotropic LV models were constructed for the three patients and the models were solved by ADINA (ADINA R&D, Watertown, MA, USA) using unstructured finite elements and the Newton-Raphson iteration method. The "Re-Start" feature in ADINA was used to adjust material parameters at each numerical time step to implement the potential conduction of myocardium. Flow velocity and stress/strain distributions were obtained for analysis. Because stress and strain are tensors, for simplicity, maximum principal stress (Stress- P_1) and strain (Strain- P_1) were used and referred to as stress and strain in this paper.

Results and Discussion

It is common to use selected cut-surfaces and critical time points (begin-filling, peak velocity during filling, begin-ejection, peak velocity during ejection, etc.) to demonstrate and compare

solution behaviors. For our modeling set-up, the time points for begin-filling and end-ejection are connection points of systole and diastole phases. The same is true for end-filling and before-ejection time points. This explanation should be helpful to understand why we mainly used end-filling and end-ejection in our comparative analyses.

Table 2 gives the maximum velocity values over the whole LV flow domain and the average flow shear stress (FSS) on LV inner surface at selected time points from the four models studied. Using the No-Pacemaker (NP) model as baseline, at the peak of filling, velocity magnitude for RVA and PIVS pacing models were 7% and 33% higher than that of the NP, respectively. Velocity magnitude for RVOT pacing model was 5% lower than that of the NP model. At the peak of ejection, velocity magnitude for RVA and PIVS pacing models were 29% and 45% higher than that of the NP model, while velocity magnitude for RVOT pacing model was 24% lower than that of the NP model.

Table 2. Velocity and flow shear stress (FSS) of pacing models

	Begin-filling		Peak of filling		Begin-ejection		Peak of ejection	
	Velocity (cm/s)	FSS (dyn/cm ²)	Velocity (cm/s)	FSS (dyn/cm ²)	Velocity (cm/s)	FSS (dyn/cm ²)	Velocity (cm/s)	FSS (dyn/cm ²)
NP	17.60	0.2142	109.9	1.359	39.87	1.096	183.1	1.970
RVA	26.20	0.7010	117.5	1.771	52.41	1.159	235.8	3.205
RVOT	19.08	0.2531	104.9	1.331	32.96	0.7831	139.9	1.616
PIVS	24.67	0.2981	146.5	1.738	44.45	1.573	266.0	3.866

Ventricle stress and strain are good measure about how hard ventricle muscle is working. It is of interest to calculate LV stress/strain conditions for comparisons. Comparison of average stress and strain values on LV inner contours of four models were given in Table 3. Using NP model as baseline, at the peak of filling, stress of RVA model was 9% higher than that of NP model. Stress of RVOT and PIVS models were 19% and 5% lower than that of NP model, respectively. Meanwhile, strain of RVA and RVOT models were 4% and 5% lower than that of NP model, respectively. Strain of PIVS model was 11% higher than that of NP model. At the peak of ejection, stress of RVA, RVOT and PIVS models were 36%, 34% and 120% higher than that of NP model, respectively. Moreover, strain of RVA pacing model was close to NP model, while strain of RVOT and PIVS pacing models were 11% and 47% higher than that of NP model, respectively.

Table 3. Stress and Strain comparison of pacing models

	Begin-filling		Peak of filling		Begin-ejection		Peak of ejection	
	Stress (kPa)	Strain	Stress (kPa)	Strain	Stress (kPa)	Strain	Stress (kPa)	Strain
NP	2.779	0.0882	81.21	0.5756	135.9	0.6979	23.46	0.4214
RVA	3.604	0.0897	88.31	0.5506	153.5	0.6709	31.90	0.4181
RVOT	2.371	0.0779	65.70	0.5482	114.3	0.6691	31.52	0.4670
PIVS	2.270	0.1062	76.78	0.6394	135.2	0.7713	51.56	0.6204

Conclusions

Correct ventricle flow characteristics and stress/strain calculations are of fundamental importance for many cardiovascular research where mechanical forces play a role in disease initiation, progression and treatment strategy selections. Ventricle remodeling, disease development, tissue regeneration, patient recovery after surgery and many other cell biological activities are closely associated with ventricle mechanical conditions. FSI models provide complete mechanical analysis including both flow forces and structural stress/strain conditions and fluid structure interaction. The existence of alternatives to existing leads and pacing methods may permit improvement in long-term outcomes with chronic pacemaker therapy while also making therapies such as synchronous pacing available to a wider array of patients with clinical situations. Direct comparison studies between pacing options will be needed to better understand the electromechanical associations and how these correlate with long-term morbidity, mortality, and quality of life. Studies concentrating on the therapeutic benefits of existing experimental therapies will also allow for the development of parameters that may permit correlation of findings during acute animal studies with long-term clinical outcomes. Further research needs to be done into options for alternative pacing methods, such as RVOT pacing, PVIS pacing, and how they correlate with long-term clinical outcomes. Lack of in vivo data and model construction cost are also considerations. Data from the literature or from ex vivo experiments have to be used to complete the computational models. We are in need of patient-specific data such as fiber orientation, sarcomere length contraction rate, regional material properties, etc.

Acknowledgement. This research was supported in part by National Sciences Foundation of China grants 11672001, 81571691, 81771844. Longling Fan's research is supported in part by the Fundamental Research Funds for the Central Universities (KYLX15_0110) and the Scientific Research Foundation of Graduate School of Southeast University (YBJJ1617).

References

- [1] Sohinki, D. and Obel, O. A. (2014) Newer algorithms in bradycardia management. *Cardiol Clin.* **32**(2), 283-292.
- [2] Liu, H. S. (2017) Mainland China Congenital Heart Disease Intervention Registration Data Sharing in 2016. *China Medical News.* **10**, 18.
- [3] Kapa, S. and Bruce, C. J. (2010) Friedman PA, Asirvatham SJ. Advances in cardiac pacing: beyond the transvenous right ventricular apical lead. *Cardiovasc Ther.* **28**(6), 369-379.
- [4] Tops, L. F., Schali, M. J. and Bax, J. J. (2009) The effects of right ventricular apical pacing on ventricular function and dyssynchrony implications for therapy. *J Am Coll Cardiol.* **54**(9), 764-776.
- [5] Das, A. and Kahali, D. (2016) Physiological cardiac pacing: Current status. *Indian Heart J.* **68**(4), 552-558.
- [6] Coppola, G., La Greca, C., Corrado, E., Ajello, L., Nogara, A., Ciaramitaro, G., Augugliaro, S., Novo, G., Novo, S. and Assennato, P. (2015) Selective pacing sites. *Minerva Cardioangiol.* **63**(2), 151-160.
- [7] Singh, H., Patel, C. D., Sharma, G. and Naik, N. (2015) Comparison of left ventricular systolic function and mechanical dyssynchrony using equilibrium radionuclide angiography in patients with right ventricular outflow tract versus right ventricular apical pacing: A prospective single-center study. *J Nucl Cardiol.* **22**(5), 903-911.
- [8] Kronborg, M. B., Mortensen, P. T., Poulsen, S. H., Gerdes, J. C., Jensen, H. K. and Nielsen, J. C. (2014) His or para-His pacing preserves left ventricular function in atrioventricular block: a double-blind, randomized, crossover study. *Europace.* **16**(8), 1189-1196.
- [9] Sharma, P. S., Dandamudi, G., Naperkowski, A., Oren, J. W., Storm, R. H., Ellenbogen, K. A. and Vijayaraman, P. (2015) Permanent His-bundle pacing is feasible, safe, and superior to right ventricular pacing in routine clinical practice. *Heart Rhythm.* **12**(2), 305-312.
- [10] Zanon, F., Baracca, E., Pastore, G., Marcantoni, L., Fraccaro, C., Lanza, D., Picariello, C., Aggio, S., Roncon, L., Dell'Avvocata, F., Rigatelli, G., Pacetta, D., Noventa, F. and Prinzen, FW. (2015) Multipoint pacing by a left ventricular quadripolar lead improves the acute hemodynamic response to CRT compared with conventional biventricular pacing at any site. *Heart Rhythm.* **12**(5), 975-981.
- [11] Tang, D., Yang, C., Geva, T. and del Nido, P. J. (2008) Patient-specific MRI-based 3D FSI RV/LV/Patch models for pulmonary valve replacement surgery and patch optimization, *J. of Biomech. Eng.* **130**, 041010.

- [12] Tang, D., Yang, C., Geva, T. and del Nido, P. J. (2010) Image-Based Patient-Specific Ventricle Models with Fluid-Structure Interaction for Cardiac Function Assessment and Surgical Design Optimization, *Progress in Pediatric Cardiology* **30**, 51-62.
- [13] Tang, D., Yang, C., Geva, T., Gaudette, G. and del Nido, P. J. (2011) Multi-physics MRI-based two-layer fluid-structure interaction anisotropic models of human right and left ventricles with different patch materials: cardiac function assessment and mechanical stress analysis, *Computers & Structures* **89**, 1059-1068.
- [14] Tang, D., del Nido, P. J., Yang, C., Zuo, H., Huang, X. Y., Rathod, R. H., Tang, A., Wu, Z. Y., Billiar, K. L. and Geva, T. (2016) Patient-Specific MRI-Based Right Ventricle Models Using Different Zero-Load Diastole and Systole Geometries for Better Cardiac Stress and Strain Calculations and Pulmonary Valve Replacement Surgical Outcome Predictions, *PLOS One* **11**, e0162986.
- [15] Fan, L. L., Yao, J., Yang, C., Tang, D. and Xu, D. (2018) Patient-Specific Echo-Based Fluid-Structure Interaction Modeling Study of Blood Flow in the Left Ventricle with Infarction and Hypertension, *CMES*. **114**, 221-237.
- [16] McCulloch, A. D., Waldman, L., Rogers, J. and Guccione, J. M. (1992) Large-scale finite element analysis of the beating heart, *Critical Rev. in Biomed. Eng.* **20**, 427-449.
- [17] Axel, L. (2002) Biomechanical dynamics of the heart with MRI, *Annu. Rev. Biomed. Eng.* **4**, 321-347.

Two parameters modelling of clay brick masonry confinement

***Giancarlo Ramaglia^{1,2}, Francesco Russo Spena², †Gian Piero Lignola², and Andrea Prota²**

¹Department of Engineering, Telematic University Pegaso, Italy.

²Department of Structures for Engineering and Architecture, University of Naples Federico II, Italy

*Presenting author: giancarlo.ramaglia@unina.it

†Corresponding author: glignola@unina.it

Abstract

Confinement is a well-known structural application since ancient times. Its early applications involved mainly masonry elements or structures, however in recent times a lot of research has been performed experimentally on confinement of concrete columns, either cylindrical or prismatic. Only recently the research differentiated the behavior of plain concrete from reinforced concrete, and the number of available confinement models increased rapidly. Predictive models are usually quite different in nature; earlier developments involved sound mechanically based approaches, based on classical failure criteria, while moving from those outcomes, proposals deviated on best fitting and empirical approaches, up to recent neural network approaches. In this framework, even if masonry confinement was the pioneer application, masonry confinement modelling has been usually borrowed from concrete confinement, which was vastly tested in the last decades.

However concrete and masonry have some crucial differences in their behavior, mainly related to their nature. Masonry is characterized by non-isotropic and non-linear behavior also for reduced strain levels. The behavior can vary significantly from masonry to masonry depending on its composition, i.e. the type and aggregation of the artificial or natural resistant elements and the type of mortar. Under uniaxial loading, masonry material exhibits a brittle behavior characterized by tensile strength far lower than compressive strength. If this is similar to concrete, the variability of the ratios between tensile and compressive behavior is notably wider for masonry. In fact ordinary concrete performance can be usually fully defined by the cylindrical compressive strength, as it is the only parameter used to individuate the confinement performance of concrete after the lateral confining pressure is known.

Authors are working on theoretical modelling of masonry confinement aiming at include other features characterizing the masonry behavior on a solid mechanics base (e.g. recently CNR guidelines added empirically the specific weight of masonry as an index for the confinement efficiency). In the present work, a failure criterion is considered containing the mean, or hydrostatic stress, able to promote the difference between compressive and tensile strength. This criterion is defined in the principal stress field and the mean stress (or first invariant) is crucial to the failure in brittle and compacting porous materials. Criteria of this kind are particularly useful not only to introduce non-uniform stress states, as those developed in non-axisymmetric confined elements, but also to be implemented in finite elements applications.

The validity of the adopted failure criterion has been checked against actively confined clay brick masonry experimentally tested under accurately known lateral pressure levels.

Keywords: Clay Brick masonry, Confinement, Failure criterion, Triaxial Hoek cell

Introduction

The confinement is the application of a wrapping around an object with the aim of limiting or preventing the lateral deformations and the failure. In the construction industry, the application of confinement is used both to confine individual structural elements either entire buildings or parts of them. Since ancient times the confinement was well-known and adopted in various fields, for instance for the vaults, the columns and the domes.

The research on the physical-mechanical characterization of materials has led to the evolution of the confinement techniques, refining methodologies and sizing. The basic idea is that, by increasing the lateral compression in an axially loaded element, a three-dimensional stress state is obtained, beneficial in terms of the ultimate load, as it is well known from the application of the classical failure criteria to the building materials.

Despite masonry confinement applications have been among the first to be developed, the wider part of confinement research of the last decades was focused on ordinary concrete elements. The two materials share the quasi-brittle nature and the relevant difference between compressive and tensile strength. However concrete and masonry have some crucial differences in their behavior, mainly related to their nature. The behavior can vary significantly from masonry to masonry depending on its composition. Masonry is characterized by non-isotropic behavior due to the type and aggregation of the bricks/blocks and the type of mortar/joints. Confinement also has unquestionable advantages in the case where the heterogeneity of the materials (such as in masonry with alternating mortar and bricks) induces tensile stresses in one of the components.

Outlining the differences between masonry and concrete is out of the scope of this work, however the recent trend to extend confinement models developed for concrete to the case of masonry elements imposes a careful re-evaluation of the theoretical bases.

Confinement modelling

First models proposed at the beginning of last century were based on solid mechanics, e.g. one of the pioneers was by Richart et al. [1] dating back to 1929 for concrete confinement. Its form was quite simple as it provided a linear formulation between masonry confined compressive strength, f_{mcd} , and lateral pressure, f_l , respectively, normalized with respect to masonry unconfined compressive strength, f_{md} and proportional to a coefficient k' :

$$\frac{f_{mcd}}{f_{md}} = 1 + k' \cdot \frac{f_l}{f_{md}} \quad (1)$$

Further models followed on an empirical base, assuming the following format:

$$\frac{f_{mcd}}{f_{md}} = 1 + a \left(\frac{f_l}{f_{md}} \right)^b \quad (2)$$

and the coefficients (or sometimes functions) a and b were repeatedly calibrated based on regression analyses, hence aiming at best fitting the experimental available data on masonry confinement. In these processes, two different uncertainties combine together: (i) the variability of masonry performance, hence a simple format involving only the compressive (unconfined) strength is weak; (ii) the former tests (on concrete only) involved passive confinement by means of steel jackets, hence the lateral pressure, at least close to the peak, was simply related to the (constant) yielding stress value of the steel material, however hundreds of further tests were based on fiber reinforced plastics (FRP) confinement. Such a confining material has a linear elastic behavior up to failure, so that the lateral pressure is continuously variable, and depends on the compatibility with lateral deformability of the

confined member. Assuming the ultimate stress of FRP as the relevant value to estimate the lateral pressure has been demonstrated to be non-conservative and many reasons for this were provided [2][3].

Furthermore many tests are conducted on non-axisymmetric elements, so that the confining pressure is not uniform and the correlation between lateral pressure and increase of compressive strength is even more complicated. Conventional approaches, like as parabolas and volumetric efficiency factors, have been provided (e.g. [4]), as long as more refined approaches taking into account the pointwise variability of lateral pressures not equal even in two orthogonal directions in the plane of the cross section [5][6].

To solve some of the issues remarked previously (i.e. on effective lateral pressure and cross sectional shape effect), in this work focus is made on circular masonry elements and for the validation, actively loaded cylindrical masonry specimens tested in triaxial compression device (Hoek cell) only were considered [7].

It is remarked that the extension of concrete models to masonry is weak because there are many differences between the two materials and the main intent is to avoid experimentally calibrated models, that do not reflect the intrinsic variability in masonry performance (apart providing a calibrated model for each masonry type), but to provide solid mechanic based models that can be implemented satisfactorily in Finite Element Models (FEMs), too, and allows to account also for non-uniform lateral pressures. Such models are necessarily multi-parameters. For instance CNR DT200R1 [8] suggests that the effects of lateral pressure on masonry confinement are proportional to the mass-density of the masonry, in the sense that heavier masonry has higher increases of compressive strength from the same amount of lateral pressure. This means that a in equation (2), according to [8], is proportional to the masonry mass density expressed in ton/m^3 . In this way the model is based on two parameters: unconfined compressive strength and mass density of masonry. Similarly, another two-parameters model was proposed [9], derived from the Mohr Coulomb failure criterion, including friction angle, Φ , and cohesion, c , to characterize different masonry materials (e.g. compared to [8], it is expected that two masonries having the same mass density, behaves differently according to other mechanical parameters). In that model [9], two independent parameters, out of the three (i.e. f_{mcd} , Φ and c), are used, e.g. k' in equation (1) is $(f_{md}/2c)^2$. However the definition of friction angle and cohesion for masonry is not always straightforward.

Recently some of the authors proposed [10] to extend the ultimate strength surface, based on five parameters, proposed by Argyris et al. [11] to masonry. This surface was previously adopted by Mander et al. [12] to calibrate their well-known solid mechanic based model for concrete confinement (and inserted also in international Codes, e.g. ACI440.2R-02 [13]). Despite its derivation is based on five parameters, the final form and coefficients, in particular, depend implicitly on the input parameters depicting the masonry behavior; the approach should be repeated for each masonry to provide the relevant confinement models. For instance the confinement equation proposed in ref [10] for clay brick masonry is as follows:

$$\frac{f_{mcd}}{f_{md}} = -1.07 + 2.07 \sqrt{1 + 7.56 \frac{f_l}{f_{md}}} - 2 \frac{f_l}{f_{md}} \quad (3)$$

To have a more flexible, solid mechanic based model, authors propose the following explicit two-parameter approach, based on the general failure surface developed by Stassi for hollow cylinders and hollow spheres [14].

Proposed two-parameters confinement model

Stassi [14] proposed a failure surface of general character that may be adopted for both soft and hard materials. The failure surface is expressed by the following equation:

$$J'_2 + \alpha \cdot I_1 = \beta \quad (4)$$

as a linear combination of first and second stress invariants, I_1 and J'_2 , respectively. The mean, or hydrostatic stress (i.e. I_1), is able to promote the difference between compressive, f_{md} , and tensile, f_{td} , strength and it is crucial to the failure in brittle and compacting porous materials. The parameters α and β have been related to f_{md} , and tensile, f_{td} , strengths (i.e. the failure surface passes through uniaxial strength points). Normalizing the principal stresses σ_1 , σ_2 , σ_3 (compression is positive) with respect to f_{md} and introducing the ratio $\rho_t = f_{td}/f_{md}$, the failure surface, $F=0$, becomes:

$$F\left(\frac{\sigma_1}{f_{md}}, \frac{\sigma_2}{f_{md}}, \frac{\sigma_3}{f_{md}}\right) = (\rho_t - 1) \left(\frac{\sigma_1}{f_{md}} + \frac{\sigma_2}{f_{md}} + \frac{\sigma_3}{f_{md}} \right) + \left[\left(\frac{\sigma_1}{f_{md}} \right)^2 + \left(\frac{\sigma_2}{f_{md}} \right)^2 + \left(\frac{\sigma_3}{f_{md}} \right)^2 - \frac{\sigma_1 \sigma_2}{f_{md}^2} - \frac{\sigma_2 \sigma_3}{f_{md}^2} - \frac{\sigma_1 \sigma_3}{f_{md}^2} \right] - \rho_t = 0 \quad (5)$$

The three-dimensional failure surface is plotted in Figure 1, assuming ρ_t changing from 0 to 1 with a step of 0.2. Equation (5) becomes a particular case, in fact it is the Von Mises failure criterion, when $\rho_t=1$.

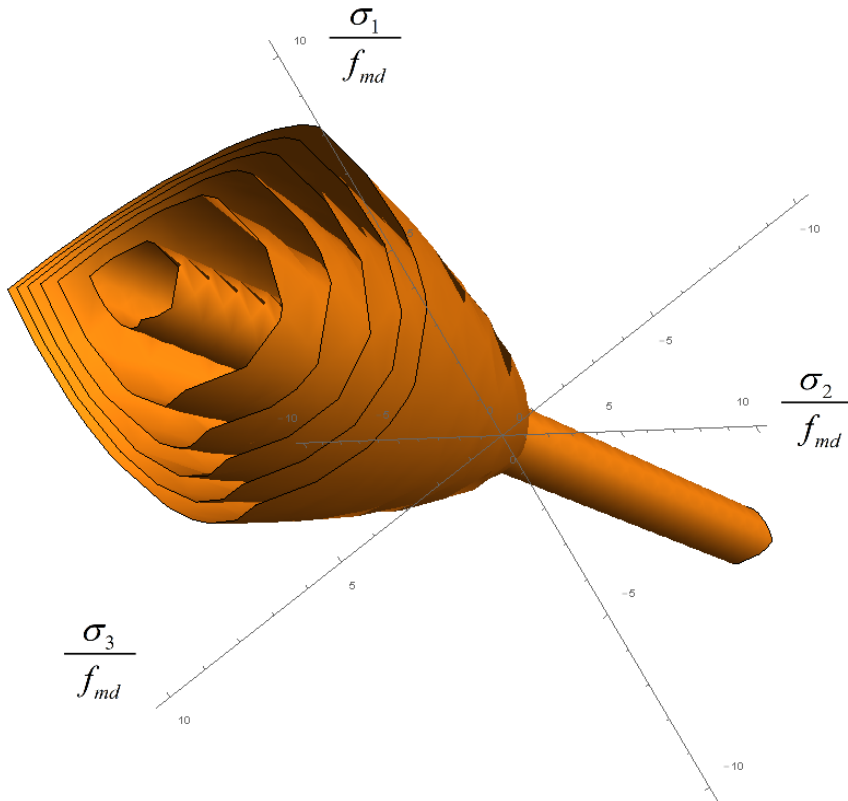


Figure 1. 3D failure surface assuming ρ_t changing from 0 to 1 with a step of 0.2

This failure surface was first used for finite element modelling of masonry structures by Sparacio and Russo Spena in 1980 [15], however it is still particularly suitable to derive a confinement model both accounting for uniform and non-uniform biaxial lateral confining pressure.

In the same format of previous equations for confinement modelling, the following positions can be assumed: axial stress $\sigma_1 = f_{mcd}$ and lateral stresses $\sigma_2 = \sigma_3 = f_l$. Equation (5) of the failure surface can be then solved with these assumptions, yielding to

$$\frac{f_{mcd}}{f_{md}} = \frac{1 - \rho_t}{2} + \sqrt{\left(\frac{1 + \rho_t}{2}\right)^2 + 3(1 - \rho_t)\frac{f_l}{f_{md}}} + \frac{f_l}{f_{md}} \quad (6)$$

hence an explicit two-parameter model, in terms of f_{md} and the ratio $\rho_t = f_{ld}/f_{md}$ is provided.

The proposed confinement model, based on equation (6) is plotted in Figure 2, assuming ρ_t changing from 0 to 1 with a step of 0.1. It is worth noting that, according to the proposed model, the lower is the ratio between tensile and compressive strength and the higher is the confinement effectiveness, i.e. given a lateral pressure, the increase of compressive strength is higher.

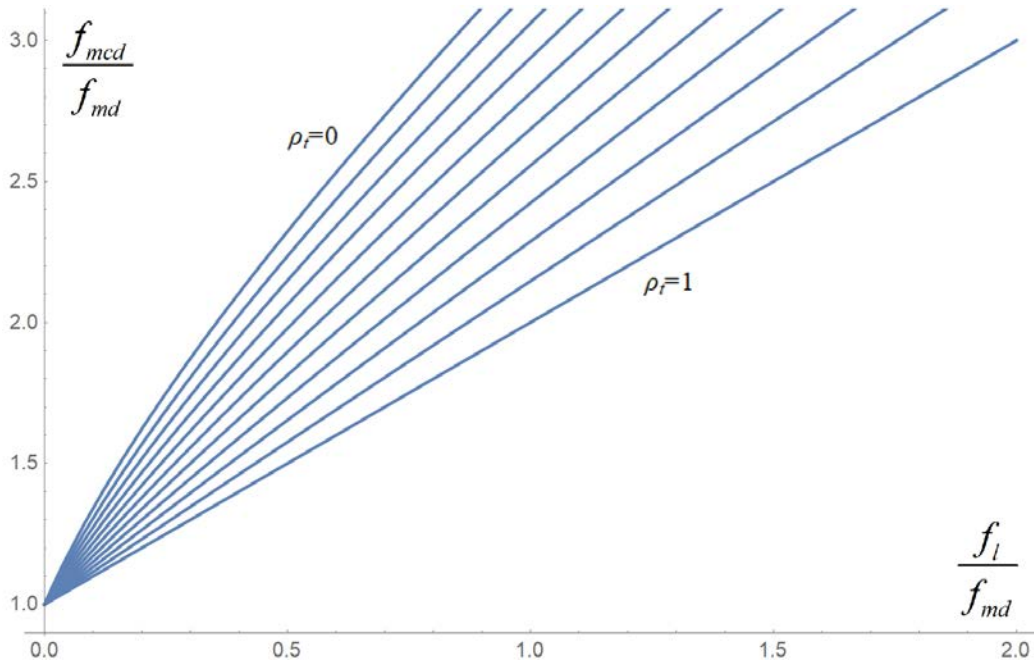


Figure 2. Proposed confinement model assuming ρ_t changing from 0 to 1 with 0.1 step

Experimental validation

The proposed confinement model provided by equation (6) is dependent on the two parameters f_{md} and ρ_t . The model is validated by means of comparison with experimental tests, where the two parameters are required, i.e. knowledge on compressive and tensile strengths. To reduce uncertainties related to the (lateral) confining pressure estimation, in particular those related with the linear elastic confining materials, like as FRP substituting the traditional steel hoops, an experimental program on cylindrical columns of 54 mm in diameter and 85 mm high, with 0.25 cm thick joints was considered [7].

The short dimensions of specimens are due to confine them actively by means of a triaxial compression device (Hoek cell) and reproduce a 1:4 scaled masonry column (however it

cannot be excluded that some size effect occurred). Three unconfined specimens were tested to assess the compressive strength and $f_{md}=13.58$ MPa with a CoV=6.22 %. Afterwards ten specimens were subjected to a uniform stress, ranging from 0.4 to 7 MPa, by a hydraulic pressure generator applied to the lateral surface of cylindrical specimens contained in a rubber tube.

Unfortunately nothing is said on the tensile strength of the masonry apart that the lime mortar was made of one portion of cement, one portion of hydrated lime, eight portions of sand, and two portions of water. In this case, on safe side, a $\rho_t=0.1$ has been assumed and the ten experimental results have been plotted as red squares to be compared with the confinement solid curve in figure 3. The comparison allows to satisfactorily validate the proposed confinement model, even if supplementary results are required to further validate the model, however usual experimental tests available in scientific literature, with confinement made by FRP, add the aforementioned uncertainties on the effective confining pressure f_l and eventually the cross sectional shape effects.

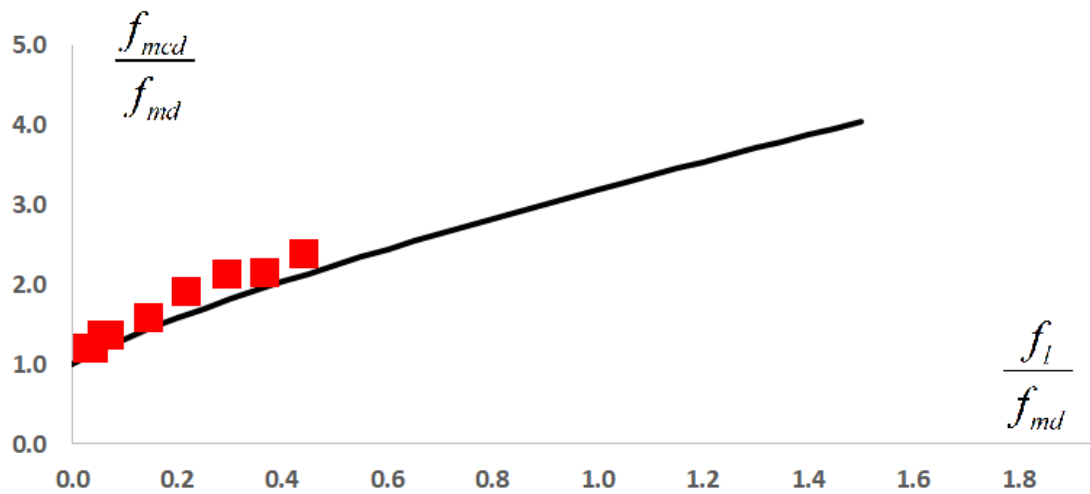


Figure 3. Experimental [7] validation of the proposed confinement model: Eq. 6 with $\rho_t=0.1$

Conclusions

Confinement of masonry is derived for similarity from concrete confinement, however many differences between the two materials can be outlined, as long as the parameters required to describe their behavior. Concrete confinement models are usually one-parameter models, as the knowledge of concrete behavior is usually comprehensive once given its cylindrical compressive strength; conversely masonry should be usually described by means of more parameters. Some attempts have been made to base the masonry confinement models on solid mechanics, hence to include as much information as possible on the masonry behavior.

In the present case, focus is made on two parameters, in particular tensile and compressive strengths, to characterize the masonry. A failure surface of general character is adopted as a linear combination of first and second stress invariants. The first invariant promotes the difference between compressive and tensile strength and is crucial to depict the failure in brittle and compacting porous materials. This confinement model adds up to few other models, mainly empirical in nature, relating the uniform lateral pressure to the increase of masonry compressive strength, however its nature allows to use its underlying failure criterion in FEM (hence including naturally the effects of confinement) and to evaluate confinement configurations where the confining stress is not uniform (hence in the case of

prismatic, non-circular elements). The proposed confinement model has been validated against few experimental tests, however there is lack of tests where the lateral pressure is known with reasonable accuracy without introducing further uncertainties on the effective confining pressure, commonly found in usual FRP confined tests.

References

- [1] Richart, F. E., Brandtzaeg A. and Brown R. L. (1929), *The failure of plain and spirally reinforced concrete in compression*, Bulletin No. 190, Engineering Experimental Station, Univ. of Illinois, Urbana, Ill., 3–72.
- [2] Zinno, A., Lignola, G. P., Prota, A., Manfredi, G. and Cosenza, E. (2010), Influence of free edge stress concentration on effectiveness of FRP confinement. *ELSEVIER Composites: Part B* **41**(7): 523-532.
- [3] Lignola, G. P., Nardone, F., Prota, A. and Manfredi, G. (2012), Analytical model for the effective strain in FRP-wrapped circular RC columns. *ELSEVIER Composites: Part B* **43**(8):3208-3218.
- [4] CNR-DT 200 (2004), *Guide for the Design and Construction of Externally Bonded FRP Systems for Strengthening Existing Structures*, Italian Council of Research (CNR), Rome (Italy) pages 1-144
- [5] Lignola, G. P., Prota, A., Manfredi, G. and Cosenza, E. (2008), Unified Theory For Confinement of RC Solid and Hollow Circular Columns. *ELSEVIER Composites: Part B* **39**(7-8):1151-1160
- [6] Lignola, G. P., Prota, A., and Manfredi, G. (2016) Simplified modeling of concrete confinement. *Proceedings of Fourth International Conference on Sustainable Construction Materials and Technologies (SCMT4)*, Las Vegas, USA, August 7-11, 2016. Paper ID S277: 1045-1054. Eds N. Ghafoori, P. Claisse, E. G., T.R. Naik, ISBN 9781535383943
- [7] Alecci, V., Briccoli Bati, S. and Ranocchiai, G. (2009) Study of Brick Masonry Columns Confined with CFRP Composite, *J. Compos. Constr.* **13**(1):179-187
- [8] CNR-DT 200 R1, (2013) *Guide for the Design and Construction of Externally Bonded FRP Systems for Strengthening Existing Structures*, Italian Council of Research (CNR), Rome (Italy) pages 1-167
- [9] Lignola, G. P., Angiuli, R., Prota, A. and Aiello, M. A. (2014) FRP Confinement of masonry: analytical modeling. *SPRINGER Materials and Structures*. **47**(12):2101-2115
- [10] Lignola, G. P., Prota, A., and Manfredi, G. (2014) Influence of masonry properties on confinement: a mechanical model. *Proceedings of MuRiCo4*. Ravenna, Italy, September 9-11 2014, 299-306.
- [11] Argyris, J. H., Faust, G., Szimmat, J., Warnke, E. P., William, K. J., (1974) Recent developments in the finite element analysis of prestressed concrete reactor vessels. *Nuclear Engineering and Design*, **28**(1):42–75.
- [12] Mander, J. B., Priestley, M. J. N., and Park, R. (1988) Theoretical stress–strain model for confined concrete. *Journal of the Structural Division ASCE* 114:1804–1826.
- [13] ACI Committee 440 (2002) *Guide for the Design and Construction of Externally Bonded FRP Systems for Strengthening Concrete Structures*, *ACI440.2R-02* American Concrete Institute, Farmington Hills, MI, pages 1-92.
- [14] Stassi, F., (1967) Flow and Fracture of Materials According to a New Limiting Condition of Yielding, *Meccanica*, **3**:178–195.
- [15] Sparacio, R., Russo Spena F. (1980) Verifica di un intervento consolidativo con il metodo degli elementi finiti. *Atti del III Convegno ASSIRCCO*, Palermo, ottobre 1980: pages 1-37

Frictional contact analysis of functionally graded materials using smoothed finite element methods

Y.F. Zhang¹, J.H. Yue¹, †R.P. Niu¹

¹Taiyuan University of Technology, Taiyuan, China

†Corresponding author: niu_ruiping2007@126.com

Abstract

In the paper, the smoothed finite element method (S-FEM) based on linear triangular elements is used to solve 2D solid contact problems of functionally graded materials. Both conforming and non-conforming contacts algorithms are developed using modified Coulomb friction contact models including tangential strength and normal adhesion. Based on smoothed Galerkin weak form, the system stiffness matrices are created using the formulation procedures of edge-based S-FEM (ES-FEM) and node-based S-FEM (NS-FEM), and the contact interface equations are discretized by contact point-pairs. Then these discretized system equations are converted into a form of linear complementarity problems (LCP), which can be further solved efficiently using the Lemke method. The singular value decomposition method is used to deal with the singularity of stiffness matrix in the procedure constructing the standard LCP, which can greatly improve the stability and accuracy of the numerical results. Numerical examples are presented to investigate the effects of functionally graded materials and comparisons have been made with analytical solutions and the standard FEM. The numerical results demonstrate that the strain energy solution of ES-FEM has higher convergence rate and accuracy compared with that of NS-FEM and FEM for functionally graded materials.

Key words: Smoothed finite element method; Contact problem; Linear complementarity problem; Strain energy; Functionally graded material

1. Introduction

Contact problems play an important role in many fields such as mechanical, civil engineering [1] and medicine. In fact, the use of functional graded materials (FGM) may become a critical issue for developing advanced lightweight structures, which meets the stringent requirements of high-tech fields. In many practical problems, the material of contact problems appears some particular physical properties changing with dimensions. For example, biological functionally gradient materials can be applied in medicine to achieve biological permanent repair and reconstruction of human hard tissue [4]. Therefore, it is necessary to study the contact problems of functionally graded materials.

Compared with solid mechanics, the geometric and material are discontinuous at the contact interface, which made it difficult to solve by the analytical method, so the numerical methods are needed. In this work, we study the contact problems with functionally graded materials using the S-FEM. Based on smoothed Galerkin weak form, the system stiffness matrices are

created using the formulation procedures of edge-based S-FEM (ES-FEM) and node-based S-FEM (NS-FEM), and the contact interface equations are discretized by contact point-pairs. However, the singularity of the system stiffness matrices makes the traditional solver for linear system of equation failed. We introduced the singular value decomposition method for solving linear system equations with singular system stiffness matrices. Through numerical simulations, it is demonstrated that the present algorithm is accurate and efficient for the contact problems of functionally graded materials.

2. Problem statement

2.1 Boundary value equations with contact

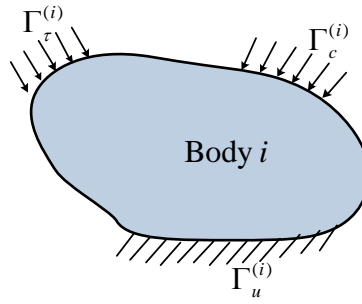


Figure 1. Configuration of the i th contact body

For solid body i as shown in Figure 1, the incremental forms from time t to $t + \Delta t$ of its static equilibrium equation, the constitutive equation and compatibility equation, the displacement increment, traction increment and contact traction increment boundary conditions are described, respectively, as follows

$$\begin{aligned}
 \mathbf{L}^T \boldsymbol{\sigma}^{(i)} + \mathbf{b}^{(i)} &= 0 \\
 \boldsymbol{\sigma}^{(i)} &= \bar{\mathbf{D}}^{(i)} \boldsymbol{\varepsilon}^{(i)} \\
 \boldsymbol{\varepsilon}^{(i)} &= \mathbf{L} \mathbf{u}^{(i)} \\
 \mathbf{u}^{(i)} &= \bar{\mathbf{u}}^{(i)}, \text{ on } \Gamma_u^{(i)} \\
 \mathbf{t}^{(i)} &= \mathbf{L}_n^T \boldsymbol{\sigma}^{(i)} = \bar{\mathbf{t}}^{(i)}, \text{ on } \Gamma_\tau^{(i)} \\
 \bar{\mathbf{t}}^{(i)} &= \mathbf{L}_n^T \boldsymbol{\sigma}^{(i)}, \text{ on } \Gamma_c^{(i)}
 \end{aligned} \tag{1}$$

2.2 Modeling of contact interfaces

By using the slack vectors of residual strength and contact gap from Kuhn-Tucker conditions, we get the conformability equations as following [3]:

$$\begin{aligned}
 -\hat{\mathbf{M}}_c \hat{\boldsymbol{\tau}} - \hat{\mathbf{k}} + \hat{\boldsymbol{\lambda}} &= 0 \\
 \hat{\mathbf{M}}_g \hat{\boldsymbol{\delta}} - \hat{\mathbf{g}} &= 0 \\
 \hat{\boldsymbol{\lambda}}^T \hat{\boldsymbol{\delta}} &= 0, \quad \hat{\boldsymbol{\lambda}} \geq 0, \quad \hat{\boldsymbol{\delta}} \geq 0
 \end{aligned} \tag{2}$$

3. Smoothed Galerkin weak form

Based on the smoothing operator, we have the following smoothed Galerkin weak form with contact boundary:

$$\begin{aligned} \sum_{c=1}^{N_c^{(i)}} \int_{\bar{\Omega}_c^{(i)}} \delta \bar{\boldsymbol{\varepsilon}}_c^{(i)T} \bar{\mathbf{D}}^{(i)} \bar{\boldsymbol{\varepsilon}}_c^{(i)} d\Omega - \int_{\Omega^{(i)}} \delta \mathbf{u}^{(i)T} \mathbf{b}^{(i)} d\Omega \\ - \int_{\Gamma^{(i)}} \delta \mathbf{u}^{(i)T} \bar{\mathbf{t}}^{(i)} d\Gamma - \int_{\Gamma_c^{(i)}} \delta \mathbf{u}^{(i)T} \boldsymbol{\Theta}^{(i)T} \boldsymbol{\tau}^{(i)} d\Gamma = 0, \end{aligned} \quad (3)$$

4. Discretized form for equation

The physical properties of functional graded materials (FGM) change with dimensions. For example, the Young's modulus is not constant, so that the material constant matrix is different for each smoothing domain. For each edge-based smoothing domain, using the material constant of the midpoint of the edge to represent the material constant of the smoothing domain. Similarly, for each node-based smoothing domain, using the material constant of the node to represent the material constant of the smoothing domain. For 2D contact problem of FGM, the material constant matrix $\bar{\mathbf{D}}$ can be written as:

$$\bar{\mathbf{D}} = [c_{ij}(x, y)], x, y \in \Omega \quad (4)$$

where c_{ij} is the material constant, which change with dimensions.

Based on smoothed Galerkin weak form, the system stiffness matrices are created using the formulation procedures of edge-based S-FEM (ES-FEM) and node-based S-FEM (NS-FEM), and the contact interface equations are discretized by contact point-pairs. Then we get the following equations:

$$\begin{aligned} \mathbf{K}\mathbf{U} - \mathbf{C}\boldsymbol{\tau} &= \mathbf{F} \\ -\mathbf{M}_c \boldsymbol{\tau} + \boldsymbol{\lambda} &= \mathbf{k} + \mathbf{M}_c \boldsymbol{\tau}_t \\ \mathbf{M}_g \boldsymbol{\delta} - \mathbf{G}\mathbf{U} &= \mathbf{G}_t \\ \boldsymbol{\lambda}^T \boldsymbol{\delta} &= 0, \boldsymbol{\lambda} \geq 0, \boldsymbol{\delta} \geq 0 \end{aligned} \quad (5)$$

From the Eq. (5) we have the following standard LCP, which can be very readily solved using the Lemke method.

$$\begin{cases} \boldsymbol{\lambda} = \mathbf{M}\boldsymbol{\delta} + \mathbf{q} \\ \tilde{\boldsymbol{\lambda}}^T \tilde{\boldsymbol{\delta}} = 0, \tilde{\boldsymbol{\lambda}} \geq 0, \tilde{\boldsymbol{\delta}} \geq 0 \end{cases} \quad (6)$$

The contact tractions and the displacement of the entire domain can be calculated using:

$$\boldsymbol{\tau} = \bar{\mathbf{K}}^{-1} (\mathbf{M}_g \boldsymbol{\delta} - \mathbf{G}_t - \mathbf{G}\mathbf{K}^{-1}\mathbf{F}) \quad (7)$$

$$\mathbf{U} = \mathbf{K}^{-1} (\mathbf{F} + \mathbf{C}\boldsymbol{\tau}) \quad (8)$$

where $\mathbf{M} = \mathbf{M}_c \bar{\mathbf{K}}^{-1} \mathbf{M}_g$, $\mathbf{q} = -\mathbf{M}_c \bar{\mathbf{K}}^{-1} (\mathbf{G}_t + \mathbf{G}\mathbf{K}^{-1}\mathbf{F}) + \mathbf{k} + \mathbf{M}_c \boldsymbol{\tau}_t$, $\bar{\mathbf{K}} = \mathbf{G}\mathbf{K}^{-1}\mathbf{C}$.

Note that the stiffness matrix \mathbf{K} in Eq. (5) may be singular, because some contact bodies may be suspended namely without displacement boundaries or constraints for rigid body movement. Assume that the i th body is suspended, and then its stiffness matrix $\mathbf{K}^{(i)}$ may be singular for static analysis. Based on singular value decomposition method, the $m \times m$ real matrix \mathbf{K} can be written as follows:

$$\mathbf{K} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (9)$$

where \mathbf{U} is an $m \times m$ orthogonal matrix, $\mathbf{\Sigma}$ is a $m \times m$ rectangular diagonal matrix with non-negative real numbers on the diagonal, \mathbf{V} is an $m \times m$ orthogonal matrix and \mathbf{V}^T is the transpose of \mathbf{V} . The diagonal entries σ_i of $\mathbf{\Sigma}$ are known as the singular values of \mathbf{K} . Then the inverse of the stiffness matrix \mathbf{K} can be written as

$$\mathbf{K}^{-1} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^T \quad (10)$$

5. Numerical example

In this section, a flat punch on an elastic foundation subjected to a uniform load as shown in Figure 2. The material parameters of flat punch and foundation are $p_0 = 1\text{Mpa}$, $w = 1\text{m}$, $h = 1.6w$, $H = 2w$, $W = 1.6w$.

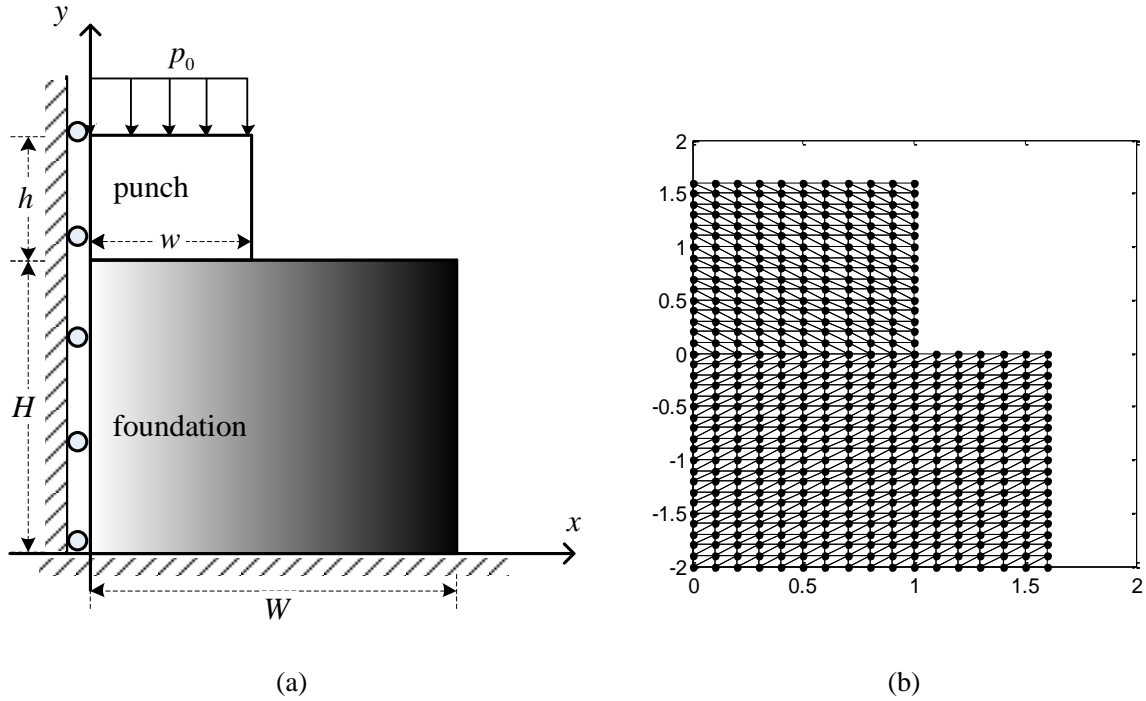


Figure 2. A flat punch on an elastic foundation subjected to a uniform load: (a) its half model with symmetric conditions imposed on the left; (b) discretized mesh model

5.1 Convergence of strain energy

The Young's modulus and Poisson ratio of the punch and foundation are $E_p = 1\text{Gpa}$, $E_f = 10\text{Mpa}$ and $\nu_p = \nu_f = 0.3$. We use six different mesh models to discretize the problem domain listed in Table 1, and the convergence of strain energy is obtained as shown in Figure 3.

Table 1 The number of nodes, T3 elements and the density of boundary nodes for flat punch

No.	M1	M2	M3	M4	M5	M6	Ref.
Nodes	153	264	544	2046	5508	12312	48622
T3 elements	240	440	960	3840	10600	24000	48000
Boundary nodes density	0.2	0.15	0.1	0.05	0.03	0.02	0.01

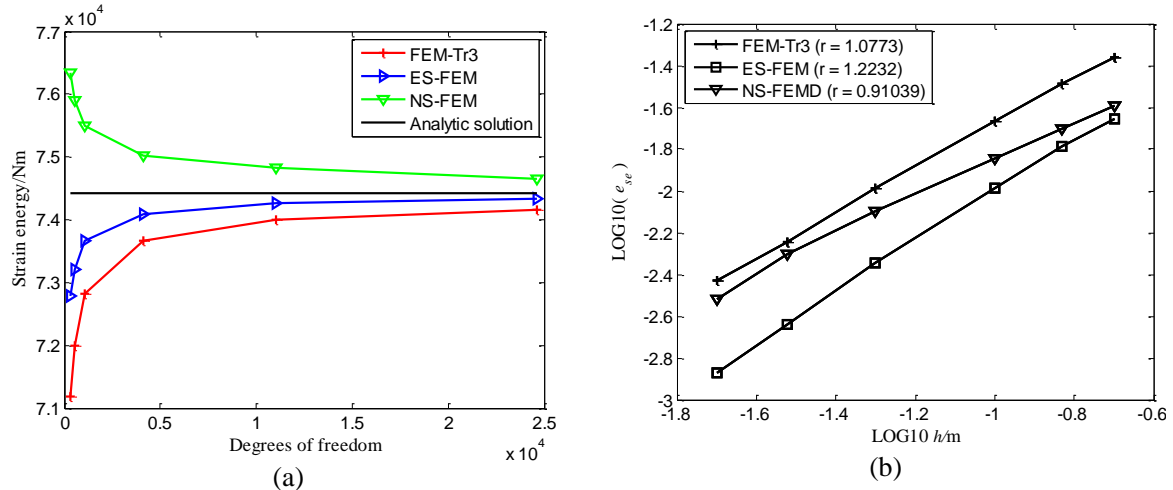


Figure 3. Using different methods for flat punch on an elastic foundation: (a) convergence of strain energy; (b) convergence ratios of relative error of strain energy

From Figure 3 (a), it is obviously observed that the solution of the NS-FEM model converges from above to the reference solution, while the ES-FEM model converges from below to the reference solution with the increase of nodal degree of freedom. Moreover, the strain energy solutions obtained by NS-FEM and ES-FEM models converge linearly with the characteristic length of the mesh, shown in Figure 3 (b). The convergence rate of ES-FEM is larger than that of NS-FEM and FEM-Tr3, which is about 20% higher than that of FEM-T3. The numerical results demonstrate that strain energy solutions of ES-FEM have higher convergence rate and accuracy compared with that of NS-FEM and FEM.

5.2 Influence of functionally graded materials

Let the Poisson ratio $\nu_p = \nu_f = 0.3$, E_0 and E_R represent the Young's modulus at the center of the foundation and at the right respectively and $r = E_R / E_0$, $E_0 = 10 \text{ Mpa}$. Let E_p represent the Young's modulus of the punch and $\kappa = E_p / E_0 = 100$. Young modulus of foundation is defined as follows: $E_f(x) = E_0 + (E_R - E_0)x/W$. In the work, let $r = 0.5, 1, 2, 5, 10$, we study the effect on the trend of normal contact traction with different r based on ES-FEM.

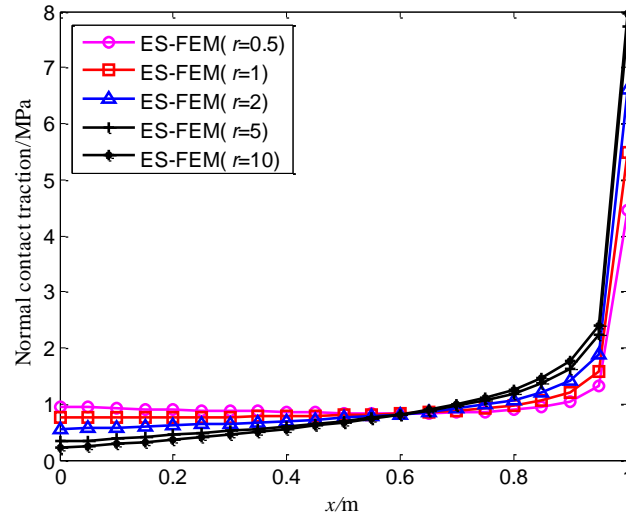


Figure 4. Normalized normal contact traction $\hat{\tau}_n$ in contact zone with different ratio r for flat punch on an elastic foundation

The Figure 4 shows that on the contact region from 0 to 0.6m, as r is increasing, the normal contact traction decreases. However, as the contact position is approaching to the right end of the punch, the contact traction increases with singularity. Furthermore, with the decreases of r the singularity of the right end decreases and the toughness of the elastic foundation can be maintained. The numerical example illustrates that the contact problems based on ES-FEM can be well solved for functional graded materials.

Conclusion

In the paper, we put forward conforming and non-conforming contacts algorithms using modified Coulomb friction contact models. Then according to S-FEM theory and contact point-pairs, discretized system equations are set up. Next, we use singular value decomposition method to solve the singularity of the stiffness matrix, through which the contact problems based on S-FEM can be well solved for functional graded materials. Through the intensive numerical simulations, we find that the presented algorithm is accurate and efficient for the contact problems of functionally graded materials.

References

- [1] Johnson KL. Contact mechanics. New York: Cambridge University Press; 1985
- [2] Miyamoto Y, Kaysser WA, Rabin BH, Kawasaki A, Ford RG. Functionally Graded Materials; Design, Processing and Applications. Kluwer Academic Publishers: Dordrecht, 1999.
- [3] Li, Y., Zhang, G.Y., Liu, G.R., Huang, Y.N., Zong, Z., 2013. A contact analysis approach based on linear complementarity formulation using smoothed finite element methods. Eng. Anal. Boundary Elem. 37 (10), 1244–1258.
- [4] Yang H, Li J, Fang H, et al. Synthesis, characterization and biological activity in vitro of FeCrAl (f)/HA asymmetrical biological functionally gradient materials[J]. Journal of Central South University, 2014, 21(2): 447-453.
- [5] Liu, G.R., Zhang, G.Y., 2013. Smoothed point interpolation methods. Smoothed Point Interpolation Methods: G Space Theory and Weakened Weak Forms.
- [6] Liu, G.R., Dai, K.Y., Nguyen, T.T., 2007. A smoothed finite element method for mechanics problems. Comput. Mech. 39 (6), 859–877

A sequential method in inverse estimation of the absorption coefficient for the spot laser welding process

Nguyen Nhut Phi Long¹, Nguyen Quan², Nguyen Hoai Son*¹, Le Trung Tin¹

¹GACES, University of Technical Education in Ho Chi Minh City, Viet Nam
e-mail: longnnp@hcmute.edu.vn,
e-mail: sonnh@hcmute.edu.vn,
e-mail: tinhktcn@gmail.com,

²Department of Engineering Technology, Pham Van Dong University, Viet Nam.
e-mail: nquan@pdu.edu.vn,
Corresponding Author (*): sonnh@hcmute.edu.vn

Abstract

In this paper, inverse determination of the absorption coefficient in spot laser welding by using a sequential method is presented. The advantages of this method are that the functional form for the unknown absorption coefficient is not necessary to preselect and nonlinear least-square do not need in the algorithm. Two examples have been fulfilled to demonstrate the proposed method. The obtained results can be concluded that the proposed method is an accurate and stable method to inversely determine the absorption coefficient in the spot laser welding.

Keywords: Inverse Problem, Absorption Coefficient, Spot Laser Welding.

Introduction

In recent years, the rapid development of the laser welding technology has gradually replaced the traditional welding techniques. Comparing the conventional welding process, laser welding was used widely because of its good behaviors such as high efficiency, narrow heat affected zone (HAZ), and high welding speed. Thus, the applications of laser welding have been broadened in modern industries including the aerospace and automotive industries, the microelectronics industry and the medical instrument industry nowadays [1, 2].

As shown in the literature, the absorption coefficient is an important factor in laser welding processing. The absorption coefficient depends on optical material properties, laser wavelength, surface temperature, and surface condition [3, 4]. Nevertheless, numerous studies in laser welding-related problems were assumed that the absorption coefficient is constant [5-7]. The effect of the absorption coefficient on the weld pool shape and on the temperature distribution was investigated by Bannour et al. [8]. The results showed that the peak temperature reaches 1250K for using the constant absorption coefficient and 1300k for the case of the temperature-dependent absorption coefficient. Furthermore, the results in Bannour et al. [8] also evidenced that the molten pool formation and temperature distribution significantly influenced by the absorption coefficient comparing with other parameters such as heat capacity, density and shielding gas. In other words, the use of appropriate absorption coefficients is significant in solving laser welding-related problems, especially the transient laser welding-related problems like spot laser welding and segment laser welding.

In fact, the direct measurement of the absorption coefficient during the laser welding process is not easy. As results, the inverse method is one of the good way to measure this coefficient. Sun et al. [9] used the direct sensitivity coefficient method to inversely estimate the absorptivity by implementing a two-dimensional quasi-static IHCP in laser hardening process. Chen et al. [10] proposed a hybrid technique of the Laplace transform and finite-difference methods to estimate the absorptivity in the laser surface heating process. Wang et al. [11] estimated the surface absorption coefficient in the laser surface hardening by using the

conjugate gradient method with the temperature-dependent thermal properties. However, these studies only deal with in the laser surface hardening in which the maximum temperature of the substrate is less than melting temperature. Thus, this result is no longer correct when the temperature field in the substrate reaches and exceeds the melting temperature in the welding process because the phase change of the laser welding process was not considered in the work. Furthermore, as our knowledge, only a few published papers implement and propose an effective method to determine the absorption coefficient in the spot laser welding until now.

In this paper, a robust and stable method is presented to determine the absorption coefficient in the spot laser welding process. In the proposed method, a modified Newton-Raphson method combined with the concept of the future time is used to solve the problem step by step [12-14]. The estimation of absorption coefficient in spot laser welding process at each time step consists of two phases: the process of direct analysis and the process of inverse analysis. In the process of direct analysis, the absorption coefficient and the boundary conditions are assumed as specified values and then the temperature field is solved by finite element method [15]. In finite element method, the effective heat capacity method [16, 17] are applied to take the latent heat into account due to the phase change in laser welding. Solution from this process are inputted to the sensitivity analysis and integrated with the measured temperature at the sensor's position. Thus, a set of nonlinear equations is formulated for the process of the inverse estimation. In the process of inverse analysis, an iterative method is used to guide the exploring points systematically to obtain the unknown variables. Then, the intermediate values are substituted for the unknown variables for the following analysis. That way, several iterations are performed to achieve the undetermined parameters. The advantage of this inverse method does not adopt the nonlinear least-squares error to formulate the inverse problem, but it is implemented a direct comparison between the measured temperature and the computed melting temperature.

Problem Statement

Considering the three-dimensional cylindrical workpiece, its top surface is heated by an incident laser beam with the laser beam radius of r_b . The rest of workpiece surface is cover by an adiabatic material to avoid the energy lost to the surroundings. The thermocouple is embedded inside the workpiece to capture the temperature history (as Figure 1).

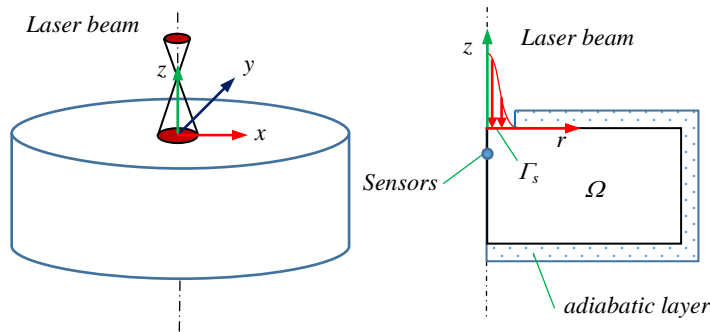


Figure 1: The model of spot laser welding.

The aim of this work is to propose the efficient method to estimate inversely the absorption coefficient in the spot laser welding. To simplify, the heat conduction-based method for this welding problem is thus considered. Due to the symmetry of cylindrical workpiece, the governing equation of transient heat conduction in two-dimensional cylindrical coordinates is given by:

$$\frac{1}{r} \frac{\partial}{\partial r} \left(k(T) r \frac{\partial T}{\partial r} \right) + \frac{\partial}{\partial z} \left(k(T) \frac{\partial T}{\partial z} \right) = \rho(T) C(T) \frac{\partial T}{\partial t} \text{ in } \Omega, \quad 0 \leq t \leq t_f \quad (1)$$

$$\begin{aligned} -k(T) \frac{\partial T}{\partial z} &= q(r) \text{ at } \Gamma_s, & 0 \leq t \leq t_f \\ -k(T) \frac{\partial T}{\partial n} &= 0 \text{ at otherwise surfaces} \end{aligned} \quad (2)$$

$$\frac{\partial T}{\partial r} = 0 \text{ at z-axis} \quad (3)$$

$$T(r, z, 0) = T_0 \quad (4)$$

where, $k(T)$, $C(T)$, $\rho(T)$ are respectively the thermal conductivity, heat capacity, and density; $T(r, z, t)$ is the temperature field; T_0 is the initial temperature; $q(r)$ is the laser heat flux; n is the normal vector. In this work, the heat source model proposed by Friedman [18] is considered and it can be expressed as following:

$$q(r) = \frac{3P\eta}{\pi r_b^2} \exp\left(\frac{-3r^2}{r_b^2}\right) \quad (5)$$

where, P is the laser power; η is the absorption coefficient, r_b is the effective radius of laser beam.

When the absorption coefficient, the boundary conditions and other input parameters are known, the temperature distribution in the domain can be solved numerically by the finite element method [15]. Furthermore, the effective heat capacity method is considered in the finite element method to take account of the latent heat of the phase change in laser welding [16, 17].

The inverse problem is to estimate the absorption coefficient in the process of spot laser welding when the temperature history is measured at $x = x_m$. Thus, a sequential method is proposed in the next section.

Methodology

The proposed method consists of the forward problem, the sensitivity problem, the operational algorithm, and the stopping criterion. The direct problem is implemented to obtain the temperature field, and the sensitivity problem is utilized to find out the search step in the inverse problem. Next, the operational algorithm is used to satisfy the process of the inverse analysis when the solution of both direct and sensitivity problems is available. Finally, the stopping criterion is shown to decide the termination of the iterative process.

Forward problem

The proposed method is based on a sequential algorithm in which the inverse solution is solved at each time step. Accordingly, Eqs. (1-4) are restricted to only one temporal step and the transient problem at $t = t_m$ is governed by the equations as follows:

$$\frac{1}{r} \frac{\partial}{\partial r} \left(k(T_m) r \frac{\partial T_m}{\partial r} \right) + \frac{1}{r} \frac{\partial}{\partial z} \left(k(T_m) r \frac{\partial T_m}{\partial z} \right) = \rho(T_m) \cdot C(T_m) \cdot \frac{\partial T_m}{\partial t} \text{ in } \Omega, t = t_m \quad (6)$$

$$\begin{aligned} -k(T_m) \frac{\partial T_m}{\partial z} &= q_m(r) \text{ at } \Gamma_s, & t = t_m \\ -k(T_m) \frac{\partial T_m}{\partial n} &= 0 \text{ at otherwise surfaces} \end{aligned} \quad (7)$$

$$\frac{\partial T_m}{\partial r} = 0 \text{ at z-axis} \quad (8)$$

$$T(r, z, t_{m-1}) = T_{m-1} \quad (9)$$

$$q_m(r) = \frac{3P \cdot \hat{\eta}_m}{\pi r_b^2} \exp\left(\frac{-3r^2}{r_b^2}\right) \quad (10)$$

where, $\hat{\eta}_m$ is the unknown absorption coefficient at $t = t_m$.

In the present work, the proposed method formulates the problem from the difference between the calculated temperature and the one measured directly. As well, instead of the optimization algorithm, the equation solver solves the inverse problem.

When the estimation is at $t = t_m$, the estimated condition from $t = t_1$ to $t = t_{m-1}$ has been evaluated, and the problem is to estimate the laser heat flux at $t = t_m$. In order to guarantee the stability of estimated results in the inverse algorithm, several future values of the estimation are temporally assumed to be constant or linear relation in the subsequent procedure [19]. Then, the unknown conditions are presented as follows:

$$\hat{\eta}_{m+\tau}^q = \hat{\eta}_m^q + \xi(\tau-1)(\hat{\eta}_m^q - \hat{\eta}_{m-1}^q) \quad (11)$$

where, τ is the number of the future time; $\xi=0$ is constant relation and $\xi=1$ is linear combination.

The forward problem, Eqs. (6-10) are solved in τ steps (from $t = t_m$ to $t = t_{m+\tau}$) and the undetermined absorption coefficient are set by Eq. (11).

Sensitivity problem

In the proposed method, a modified Newton-Raphson method is adapted to solve the inverse problem in which the sensitivity analysis is necessary to achieve the search step in each iteration. The derivative $\partial / \partial \hat{\eta}_m$ is taken at both sides of Eqs. (6-10). Furthermore, because of the small number of future time step and the small temporal increment, we can assume that the thermal properties at the estimating step t_m are constant. Then, we have:

$$\frac{1}{r} \frac{\partial}{\partial r} \left(k(T_m) r \frac{\partial X_m}{\partial r} \right) + \frac{1}{r} \frac{\partial}{\partial z} \left(k(T_m) r \frac{\partial X_m}{\partial z} \right) = \rho(T_m) C(T_m) \frac{\partial X_m}{\partial t} \quad (12)$$

$$-k(T_m) \frac{\partial X_m}{\partial z} = \frac{\partial q_m(r)}{\partial \hat{\eta}_m} \quad (13)$$

$$\frac{\partial X}{\partial r} = 0 \quad (14)$$

$$X(r, z, t_{m-1}) = X_{m-1} = 0 \quad (15)$$

Eqs. (12-15) describe the mathematical equations for sensitivity coefficient, X_m , which can be explicitly solved. These equations are the linear equations and the dependent variable, X_m , with respect to independent variables, x , y , z and t . Therefore, the sensitive solution can be obtained directly through by the finite element method.

Modified Newton-Raphson method

A modified Newton-Raphson method [14] is necessary in the proposed method to deal with the inverse problem with solving a set of nonlinear equations. This set of nonlinear equations is directly formulated the problem from the comparison between the computed temperature and the preselected temperature at the measurement locations. Therefore, the measured temperature Y_{meas}^j and the calculated temperature Y_c^j are evaluated first. Then estimation of the absorption coefficient, $\hat{\eta}_m$, at each time step can be recast as the solution of a set of nonlinear equations:

$$\mathbf{Y} = Y_c^j - Y_{meas}^j = 0 \quad (16)$$

where, $j = m, m+1, \dots, m+\tau$ is the number of equations which is equal to the number of the future times τ .

The derivative of \mathbf{Y} with respect to $\hat{\eta}_m$ is solved through Eqs. (12-15) and can be expressed as following:

$$\mathbf{X} = \frac{\partial \mathbf{Y}}{\partial \hat{\eta}_m} \quad (17)$$

where, \mathbf{X} is the sensitivity matrix

With the starting $\hat{\eta}_m^0$ and the above derivations from Eq. (17), we have the following equation:

$$\hat{\eta}_m^{k+1} = \hat{\eta}_m^k + \Delta^k \quad (18)$$

where, Δ^k is a linear least-squares solution for a set of over-determined linear equations and it can be derived as following:

$$\Delta^k = -[\mathbf{X}^T(\hat{\eta}_m^k)\mathbf{X}(\hat{\eta}_m^k)]^{-1} \mathbf{X}^T(\hat{\eta}_m^k)\mathbf{Y}(\hat{\eta}_m^k) \quad (19)$$

The preceding derivation is applied at each time step. This method can be carried out in the multi-sensor's measurement. Under this condition, the number of elements in Eq. (16) is based on the number of future time step and the number of measured positions.

The stopping criteria

The modified Newton-Raphson method (Eqs. (16-19)) is used to determine the unknown value of the absorption coefficient at the m -th time step, $\hat{\eta}_m$. The step size Δ^k goes from $\hat{\eta}_m^k$ to $\hat{\eta}_m^{k+1}$, and it is determined from Eq. (19). Once Δ^k has been calculated, the iterative to determine $\hat{\eta}_m^{k+1}$ is executed until the stopping criterion is satisfied.

The discrepancy principle [20] is widely used to evaluate the value of the stopping criterion in the inverse technique. Nevertheless, the convergence of the inverse solution is not guaranteed by the stopping criterion created by the discrepancy principle. Therefore, two criteria proposed by Frank and Wolfe [21] are chosen to assure the convergence and to stop iteration:

$$\|\hat{\eta}_m^{k+1} - \hat{\eta}_m^k\| / \|\hat{\eta}_m^{k+1}\| \leq \delta \quad (20)$$

$$\|\mathbf{S}(\hat{\eta}_m^{k+1}) - \mathbf{S}(\hat{\eta}_m^k)\| / \|\mathbf{S}(\hat{\eta}_m^{k+1})\| \leq \varepsilon \quad (21)$$

where

$$\|\mathbf{S}(\hat{\eta}_m^{k+1})\| = \sum_{i=1}^r [Y_c^i - Y_{meas}^i]^2 \quad (22)$$

where, ε and δ are small positive value known as the convergence tolerances.

Computational algorithm

We choose the number of the future time, r , the mesh configuration of the problem domain, and the temporal size, Δt first. Given overall convergence tolerance ε and δ , and the initial guess $\hat{\eta}_m^0$. The value of $\hat{\eta}_m^k$ is known at the k -th iteration.

Step 1: Let $j = m$ and $T(r, z, t_{j-1})$ is known.

Step 2: Collect the measured temperature, Y_{meas}^j .

Step 3: Calculate the sensitivity matrix, \mathbf{X} , by Eqs. (12-15).

Step 4: Solve the direct problem by Eqs. (6-10), and then obtain the calculated temperature Φ_c^j .

Step 5: Construct \mathbf{Y} by Y_{meas}^j and Y_c^j .

Step 6: Knowing \mathbf{Y} and \mathbf{X} , determine the step size Δ^k by Eq. (19).

Step 7: Knowing Δ^k and $\hat{\eta}_m^k$, calculate $\hat{\eta}_m^{k+1}$ through Eq. (18).

Step 9: Terminate the iteration if the stopping criterion (Eqs. (20-21)) is satisfied. Otherwise, return to step 5.

Step 10: Stop the process if the final time step is attached. Otherwise, let $j=m+1$ return to step 2.

Results and Discussion

Two simples are presented to demonstrate that the proposed method can estimate accurately the absorption coefficient in spot laser welding. In two examples, the cylindrical substrate has the height of $H = 5[\text{mm}]$ and the diameter of $d = 20[\text{mm}]$. The material used for these examples is commercial AISI304 which thermal properties are temperature-dependent and are taken from Sabarickanth [22]. The latent heat of fusion is $L = 272[\text{kJ/kg}]$, and the melting temperature range is from solidus temperature $T_s = 1673[\text{K}]$ to liquidus temperature $T_l = 1773[\text{K}]$. A thermocouple is located at $x_m(0, -1[\text{mm}])$. In addition, the measured temperature is generated from Eqs. (1-4) when the input parameters are preselected and it is presumed to have measurement errors. In other work, the random errors of measurement are added to the exact temperature. It can be achieved in the following equation:

$$T^{meas} = T^{exact} + \lambda \sigma \quad (23)$$

where, T^{exact} is the exact temperature, T^{meas} is the measured temperature, λ is random numbers calculated by the IMSL subroutine DRNNOR [23] and chosen over the range $-2.576 \leq \lambda \leq 2.756$, which presents the 99% confidence bond for the measured temperature.

The mesh in all cases is fine at the incident laser beam with $\Delta x \approx 2.10^{-5}[\text{mm}]$ and is coarse at away with $\Delta x \approx 1.5.10^{-3}[\text{mm}]$ (as Figure 2). As well, the time increment is $\Delta t = 0.02[\text{s}]$.

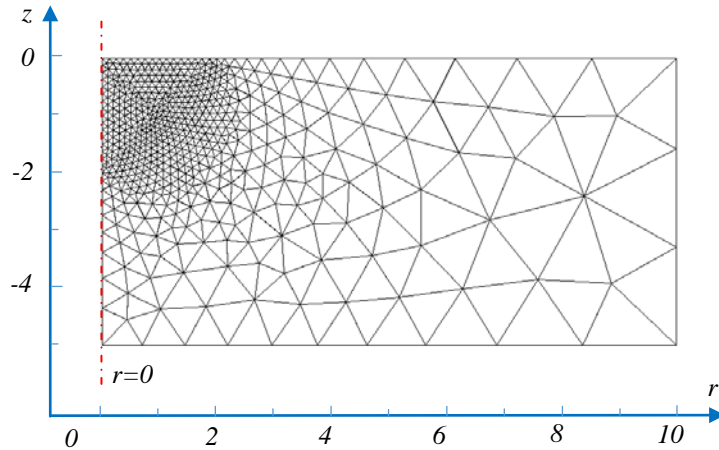


Figure 2: The mesh configuration

To investigate the deviation of the estimated results from the exact solution, the relative average error for the estimated solutions is defined as following:

$$\mu = \frac{1}{N_t} \sum_{i=1}^{N_t} \left| \frac{f - \hat{f}}{\hat{f}} \right| \quad (24)$$

where, f is the estimated results with measurement errors, \hat{f} is the exact results, and N_t is the number of the temporal step. It is declared that a smaller value of μ indicates a better estimation and vice versa.

Example 1: A constant absorption coefficient of 0.3 is assumed in this example. The workpiece is initially at a uniform temperature $T_0 = 27[^\circ\text{C}]$, and then is heated by a laser beam with its effective radius of $r_b = 0.63[\text{mm}]$ and power of $P = 400[\text{W}]$. In general, the laser heat source can be obtained as follows:

$$q_m(r) = \frac{3 \times 400 \times 0.3}{\pi (0.63 \cdot 10^{-3})^2} \exp\left(\frac{-3r^2}{(0.63 \cdot 10^{-3})^2}\right) \quad (25)$$

The estimated results of the absorption coefficient in the case of measurement error-free are shown in Figure 3 and Figure 4 shown the exact and computed temperature at the sensor position $(x_m(0, -l[\text{mm}]))$. As shown, when $\sigma = 0$, these estimated results are an excellent approximation of the exact solution for both constant and linear type of future time.

In the case of the measurement errors, the estimated results largely diverge from the exact solution (as Figure 4). Table 1 illustrates the relative average errors of the estimated results when the measurement errors are included. In general, the relative average errors are small in all cases. As shown in Table 1, even though the large measurement error of $\sigma = 1.5$, this value for the constant type of future time is 0.01. Furthermore, the relative average errors reduce with the increase of the number of future time step and the decrease of the measurement error. For example, the relative average error moderates from 0.0053 to 0.0022 about 58% as the number of future time steps increase from $\tau = 2$ to $\tau = 4$ for the linear type of future time and reduces from 0.01 to 0.0066 about 66% as the measurement error decrease from $\sigma = 1.5$ to $\sigma = 1$ for the constant assumption of future time.

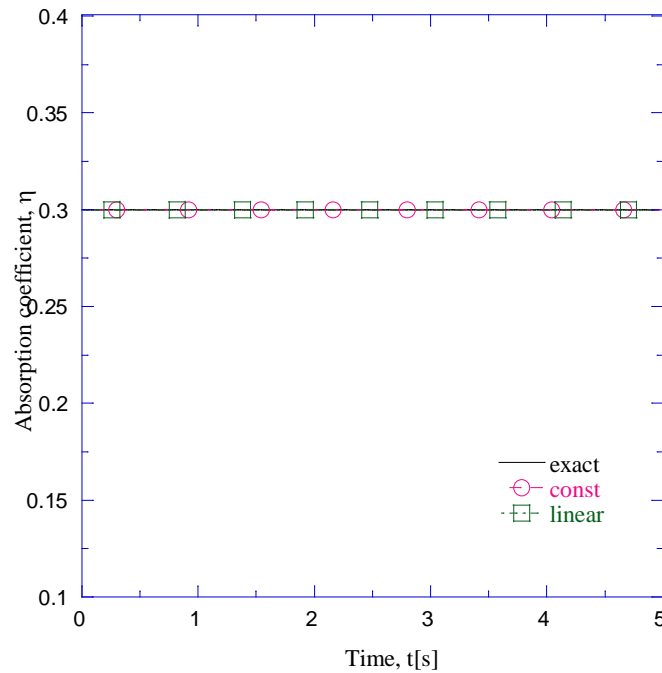


Figure 3: Estimated absorption coefficient in example 1 with $r = 2$ and $\sigma = 0$ with two function kinds of future time

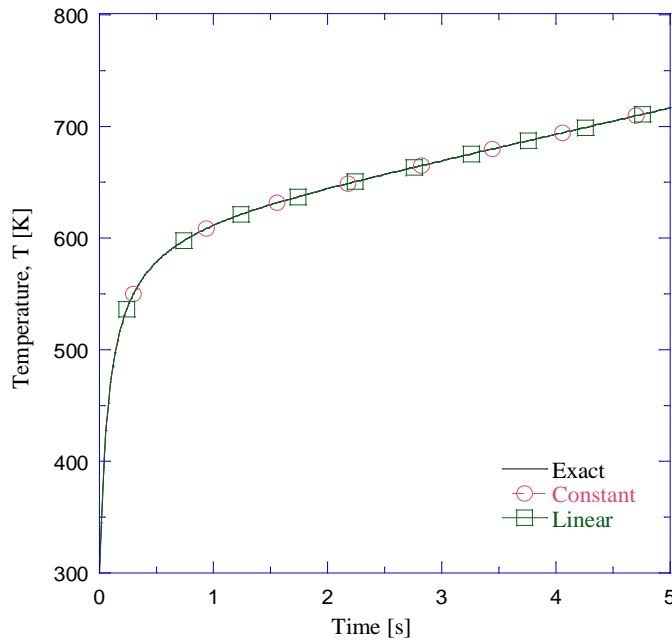


Figure 4: The temperature distribution at the sensor position with $r = 2$ and $\sigma = 0$ with two function kinds of future time

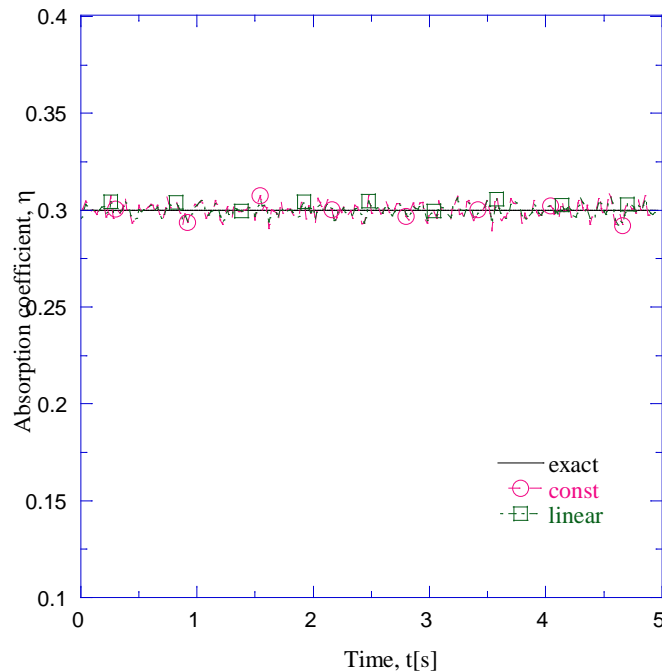


Figure 5: Estimated absorption coefficient in the example 1 with $\tau = 2$ and $\sigma = 1.5$ with two function kinds of future time

Additionally, the effect of the function type of future time on the estimated results is compared. As mentioned above, two kinds of future time description are considered in this work. One is constant type and the other is linear type. The results showed that the accuracy of estimated results for linear function of the future time is better than that for the constant function of the future time (Table 1). The estimated absorption coefficients with $\tau = 2$ and $\sigma = 1.5$ for both constant and linear functions of future time are shown in Figure 5. The results in this profile show that the estimated results for the linear combination of future time can close to the exact solution compared with that for the constant type of future time. In other words, the linear function of future time decreases the relative average error effectively (as Table 1).

Table 1: Relative average errors of example 1

Cases	Future time step	
	$\tau=2$	$\tau=4$
Linear		
$\sigma=1$	0.0053	0.0022
$\sigma=1.5$	0.008	0.0033
Constant		
$\sigma=1$	0.0066	0.0041
$\sigma=1.5$	0.01	0.0044

Example 2: In this example, the time variation of the absorption coefficient is assumed as follows:

$$\eta(t) = 0.3(1 - 6.5 \times 10^{-1} \exp(-(t-2)/0.75)) \quad (26)$$

The estimated results of the absorption coefficient in the example 2 are shown in Figure 6. Once again, Figure 6 shows that the estimated results have good approximation in the case of measurement free-error.

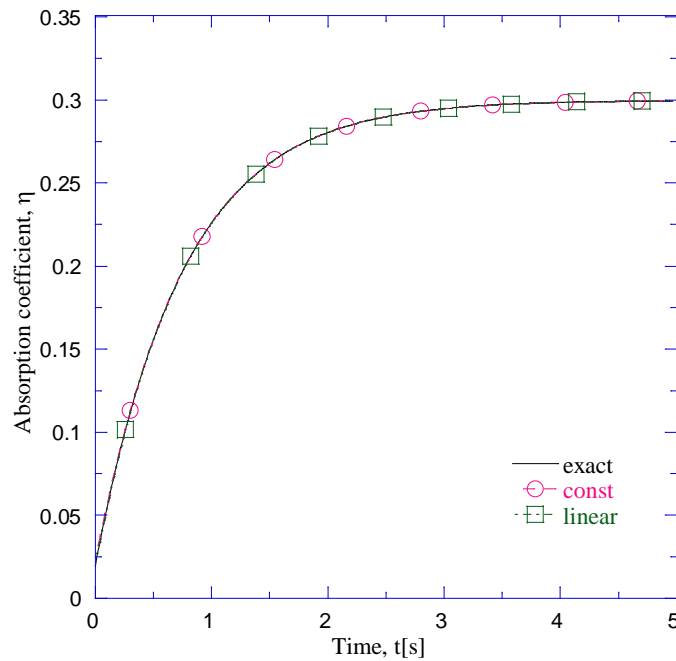


Figure 6: Estimation of laser heat flux in example 2 with $\tau = 2$ and $\sigma = 0$ with two function kinds of future time

Figure 7 and Figure 8 illustrate the estimated results with the measurement errors for the linear assumption of the future time in the cases of $\tau = 2$ and $\tau = 4$, respectively. From Figure 7 and Figure 8, in general, the estimated results have a good approximation to the exact solution with the measurement errors included.

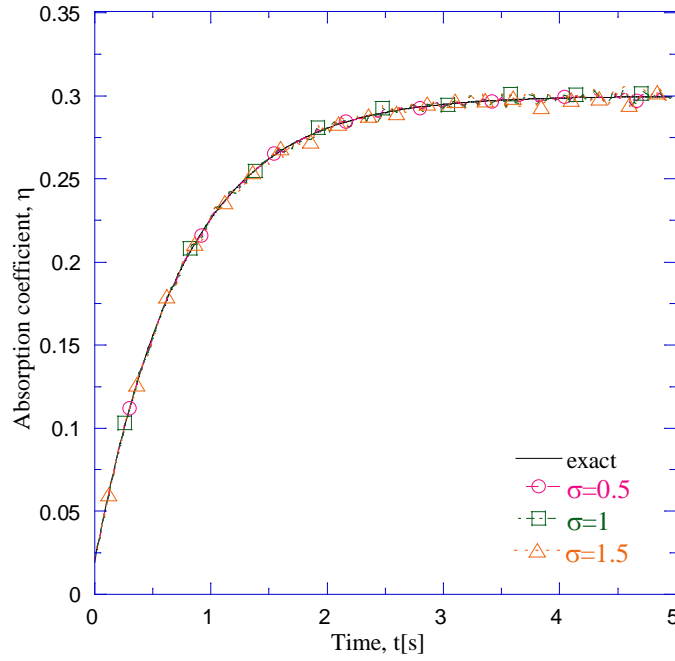


Figure 7: Estimated absorption coefficient in the example 2 with the measurement errors for the linear assumption of future time and $\tau=2$.

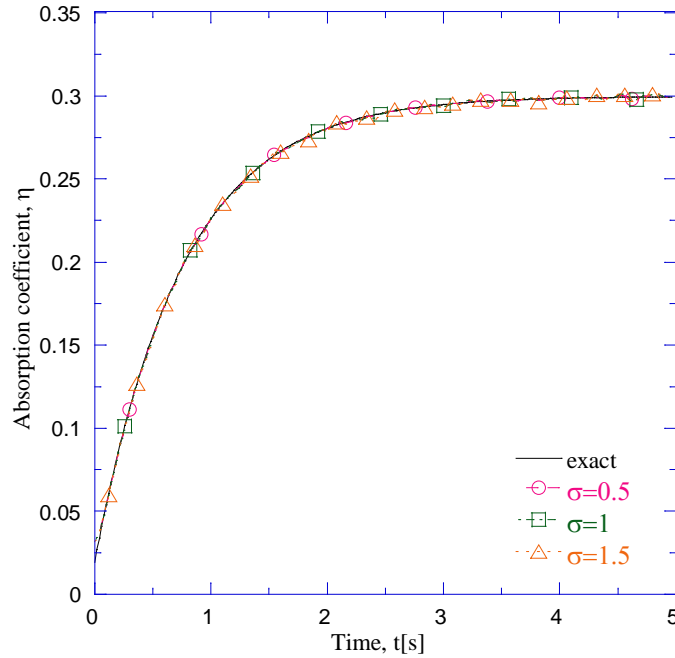


Figure 8: Estimated absorption coefficient in the example 2 with the measurement errors for the linear assumption of future time and $\tau=4$.

Table 2 presents the relative average errors with the different measurement errors and future time steps in the example 2. Table 2 shows that the relative average errors reduce as the measurement errors decrease. It can be noted that, in example 2, the relative average errors do not reduce when the number of the future time step increase. This phenomena is called the “leading error” as has been described by Lin [24, 25]. It appears in these results because of the temporary assumption in the constant and linear types of future time in Eq. (11), as these assumptions might not exactly match the form of unknown absorption coefficient. Furthermore, with the form of undetermined absorption coefficient in the example 2, the

linear combination of future time has a better approximation than the constant assumption of future time. Thus, the relative averages errors for the linear combination of future time is less than that for the constant assumption of future time (as Table 2). In general, the relative averages errors in all cases are small. This implies that the proposed method estimates accurately the absorption coefficient in the spot laser welding.

Table 2: Relative average errors of example 2

Cases	Future time step	
	$\tau = 2$	$\tau = 4$
Linear		
$\sigma = 0.5$	0.0042	0.0061
$\sigma = 1$	0.0065	0.0072
$\sigma = 1.5$	0.0089	0.0094
Constant		
$\sigma = 0.5$	0.0071	0.0121
$\sigma = 1$	0.0102	0.0134
$\sigma = 1.5$	0.0133	0.0148

From the results and discussion above, it can be declared that the proposed method is an effective and stable method to estimate the absorption coefficient in the spot laser welding.

Conclusion

In this paper, the estimation of the absorption coefficient in the spot laser welding was present by using a sequential method. As well, the inverse solution at each time step is solved by a modified Newton-Raphson method. The advantage of this proposed method is that the nonlinear least-squares error is not adopted to formulate the inverse problem, but it is implemented a direct comparison of the measured and calculated temperature. In addition, the special characteristics of this method are that preselected functional form for the unknown absorption coefficient is not necessary. Two examples have been fulfilled to demonstrate the proposed method. The accuracy of the estimated results with the different measurement errors and number of future time steps is investigated. The results show that the accuracy of the estimated results increases when the measurement error decreases and the number of future time step increases. Additionally, two kinds of function of future time are also discussed. In two examples, the results showed that the estimated results with the linear relation of future time is more accurate than that with constant type of future time. In conclusion, from the results in the examples, it can be concluded that the proposed method is an accurate and stable method to determine the absorption coefficient in the spot laser welding.

Reference

- [1] Yilbas B.S., Akhtar S., Shuja S.Z., 2013, *Laser Forming and Welding Processes*, Springer International Publishing.
- [2] Steen W.M., Mazumder J., (2010) *Laser Material Processing*, Springer.
- [3] Wieting T., DeRosa J., (1979) *Effects of surface condition on the infrared absorptivity of 304 stainless steel*, J Appl Phys, 50(2) 1071-1078.
- [4] Wieting T., Schriempf J., (1976) *Infrared absorptances of partially ordered alloys at elevated temperatures*, J Appl Phys, 47(9) 4009-4011.
- [5] Akbari M., Saedodin S., Toghraie D., Shoja-Razavi R., Kowsari F., (2014) *Experimental and numerical investigation of temperature distribution and melt pool geometry during pulsed laser welding of Ti6Al4V alloy*, Optics & Laser Technology, 59(0) 52-59.
- [6] Shanmugam N.S., Buvanashakaran G., Sankaranarayanan K., (2013) *Some Studies on Temperature Distribution Modeling of Laser Butt Welding of AISI 304 Stainless Steel Sheets*, World Academy of Science, Engineering and Technology, 7.

- [7] Trivedi A., Bag S., De A., (2007) *Three-dimensional transient heat conduction and thermomechanical analysis for laser spot welding using adaptive heat source*, Sci Technol Weld Joi, 12(1) 24-31.
- [8] Bannour S., Abderrazak K., Mhiri H., Le Palec G., (2012) *Effects of temperature-dependent material properties and shielding gas on molten pool formation during continuous laser welding of AZ91 magnesium alloy*, Optics & Laser Technology, 44(8) 2459-2468.
- [9] Sun Y.S., Weng C.I., Chen T.-C., Li W.-L., (1996) *Estimation of Surface Absorptivity and Surface Temperature in Laser Surface Hardening Process*, Japanese Journal of Applied Physics, 35(6R) 3658.
- [10] Chen H.T., Wu X.Y., (2006) *Estimation of surface absorptivity in laser surface heating process with experimental data*, Journal of Physics D: Applied Physics, 39(6) 1141.
- [11] Wang J.-T., Weng C.I., Chang J.G., Hwang C.C., (2000) *The influence of temperature and surface conditions on surface absorptivity in laser surface treatment*, J Appl Phys, 87(7) 3245-3253.
- [12] Yang C.-Y., (2003) *Estimation of boundary conditions in nonlinear inverse heat conduction problems*, Journal of Thermophysics and Heat Transfer, 17(3) 389-395.
- [13] Lin D.T.W., Yang C.-y., (2007) *The estimation of the strength of the heat source in the heat conduction problems*, Applied Mathematical Modelling, 31(12) 2696-2710.
- [14] Nguyen Q., Yang C.-Y., (2015) *Inverse determination of laser power on laser welding with a given width penetration by a modified Newton–Raphson method*, International Communications in Heat and Mass Transfer, 65 15-21.
- [15] Carnahan B., Luther H.A., Wilkes J.O., 1969, *Applied Numerical Methods*, Wiley, New York.
- [16] Huang H.-C., Usmani A.S., 1994, *Finite Element Analysis for Heat Transfer: Theory and Software*, 1 ed., Springer-Verlag London.
- [17] Lewis R.W., Ravindran K., (2000) *Finite element simulation of metal casting*, International Journal for Numerical Methods in Engineering, 47(1-3) 29-59.
- [18] Friedman E., (1975) *Thermomechanical Analysis of the Welding Process Using the Finite Element Method*, Journal of Pressure Vessel Technology, 97(3) 206-213.
- [19] Beck J.V., Blackwell B., St.Clair C.R., (1985) *Inverse heat conduction ill-posed problems*.
- [20] Tikhonov A.N., Arsenin V.I.A.k., 1977, *Solutions of ill-posed problems*, Vh Winston.
- [21] Frank M., Wolfe P., (1956) *An algorithm for quadratic programming*, Naval research logistics quarterly, 3(1-2) 95 -110.
- [22] Sabarikanth R., Sankaranarayanamsamy K., Siva Shanmugam N., Buvanashakaran G., (2009) *A study of laser welding modes with varying beam energy levels*, Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 223(5) 1141-1156.
- [23] IMSL, (1987) *Library Edition 10.0, User's Manual: Math Library Version 1.0*, Houston, Tex.
- [24] Lin S.-M., Chen C.o.-K., Yang Y.-T., (2004) *A modified sequential approach for solving inverse heat conduction problems*, International Journal of Heat and Mass Transfer, 47(12-13) 2669-2680.
- [25] Lin S.-M., (2011) *A sequential algorithm and error sensitivity analysis for the inverse heat conduction problems with multiple heat sources*, Applied Mathematical Modelling, 35(6) 2607-2617.

Limit analysis of masonry structures based on fictitious associative-type contact interface laws

Francesco Trentadue^{1*}, Giuseppe Quaranta^{2†}

¹Department of Civil Engineering and Architecture, Technical University of Bari, Italy

²Department of Structural and Geotechnical Engineering, Sapienza University of Rome, Italy

*Presenting author: francesco.trentadue@poliba.it

†Corresponding author: giuseppe.quaranta@uniroma1.it

Abstract

We illustrate an original method for the limit analysis of masonry structures modeled as assemblies of dry rigid blocks with Coulomb-type (non-associative) contact interface laws. The method resorts to a fictitious system characterized by cohesive-type contact interface laws that depend on the axial forces of the real block system. Two theorems establish the connection between the collapse state of the real (frictional) block assembly and that of the fictitious one. Hence, an alternative problem of mathematical programming is presented to evaluate the minimum collapse load multiplier. According to the proposed formulation, the complementarity condition is not introduced as constraint but is obtained as Karush-Kuhn-Tucker condition. Several numerical results concerning with masonry arches, portals and panels are provided to illustrate the application of the proposed approach, which is also validated through the comparison with some existing methods.

Keywords: **Masonry structure; Friction; Limit analysis; Mathematical programming.**

Introduction

Limit analysis provides an effective framework to study collapse load and failure mechanisms of the structures. Among the potential applications, those pertaining to block assemblies in presence of friction at the contact surfaces have received several attentions in the last decades because of the relevant practical implications. For instance, the collapse load estimation of rigid block systems interacting through frictional interfaces is of particular importance for the assessment of masonry structures. In this field, Baggio and Trovalusci [1] studied the limit analysis of no-tension and frictional three-dimensional discrete systems. In their study, the solution of the nonlinear programming problem is obtained by solving a preliminary linear programming problem that corresponds to a linearized limit analysis with dilatancy at the interfaces. Ferris and Tin-Loi [2] calculated the collapse loads of discrete rigid block systems with frictional contact interfaces by formulating a special constrained optimization problem and proposed an algorithm based on the relaxation of the complementarity constraint for its solution. The relaxation parameter is progressively reduced to zero through a succession of nonlinear sub-problems. Orduña and Lourenço [3] presented a model for the limit analysis of three-dimensional block assemblages interacting through frictional interfaces and included a proposal to take into account torsional failure modes. The model also accounted for limited compressive stresses at the interfaces. Gilbert and co-workers [4] illustrated an iterative procedure based on the successive solution of linear programming sub-problems. The method presented by these authors assumes fictitious values of cohesion and negative angles of friction, which are progressively relaxed toward zero. A finite-element-based approach has

been described by Mihai [5] for the limit analysis of planar systems formed by linear elastic bodies in non-penetrative contact with Coulomb friction.

In the present contribution, we illustrate a new method for the limit analysis of discrete systems formed by dry rigid blocks characterized by Coulomb-type (non-associative) contact interface laws [6]. The proposed method resorts to a discrete system with fictitious cohesive-type contact interface laws depending on the axial forces of the real block system. Once the connection between the collapse state of the fictitious block assembly and that of the real one is demonstrated, a new formulation of the mathematical programming problem intended to estimate the collapse load is proposed. In particular, the minimum collapse multiplier is here obtained by solving a nonlinear mathematical programming problem where the constraints include: (i) equilibrium conditions, (ii) kinematic conditions, and (iii) a further condition imposing that the collapse multiplier is kinematically admissible for the fictitious system with cohesive-type contact laws. In doing so, the classical complementarity condition is not introduced as constraint but is obtained as Karush-Khun-Tucker condition.

Proposed method for the limit analysis of masonry structures

An assembly of dry n_b blocks is considered. The constituent blocks are rigid and are allowed to slide over each other. Moreover, a Coulomb model is assumed to represent the frictional contact at the interfaces of the blocks. Contact forces and moments for the j th constituent block are defined as shown in Fig. 1.

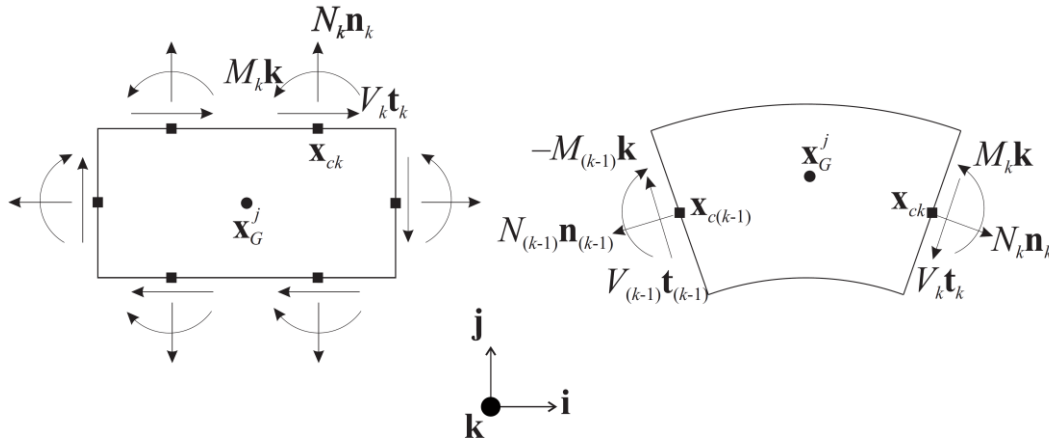


Figure 1. Contact forces and moments

The well-known equilibrium equations for the blocks are the following:

$$\begin{cases} \mathbf{f}_0^j + \alpha \bar{\mathbf{f}}^j + \sum_{\mathbf{x}_{ck} \in \partial B^j} (V_k \mathbf{t}_k + N_k \mathbf{n}_k) = \mathbf{0} \\ \mathbf{x}_G^j \wedge (\mathbf{f}_0^j + \alpha \bar{\mathbf{f}}^j) + \sum_{\mathbf{x}_{ck} \in \partial B^j} [(\mathbf{x}_{ck} - \mathbf{x}_G^j) \wedge (V_k \mathbf{t}_k + N_k \mathbf{n}_k) \pm M_k \mathbf{k}] = \mathbf{0} \end{cases}, \quad (1)$$

where $j=1, \dots, n_b$, \mathbf{f}_0^j is the j th constant force vector acting on the j th block, $\bar{\mathbf{f}}^j$ is the j th base external force vector amplified by the load multiplier α . Moreover, N_k are the normal contact forces, V_k are the shear contact forces, M_k are the contact bending moments (\mathbf{n}_k , \mathbf{t}_k and \mathbf{k} are unit vectors). The k th resultant internal force or moment is applied on \mathbf{x}_{ck} , which is the k th contact interface point (with $k=1, \dots, n_c$) belonging to the boundary ∂B^j of the j th block whose

center of mass is denoted as \mathbf{x}_G^j . Based on Eq. (1), the equilibrium of the structure can be expressed as follows [2]:

$$\mathbf{A}_f \mathbf{s}_f + \alpha \bar{\mathbf{f}} + \mathbf{f}_0 = \mathbf{A} \mathbf{s} + \mathbf{f}_0 = \mathbf{0}, \quad (2)$$

where $\alpha \bar{\mathbf{f}}$ and \mathbf{f}_0 are the vectors collecting the j th forces $\alpha \bar{\mathbf{f}}^j$ and \mathbf{f}_0^j , respectively, whereas \mathbf{s}_f is the vector of the contact forces (internal forces and reactions). It is also introduced a vector \mathbf{s} that includes the contact forces vector \mathbf{s}_f and the load multiplier α , namely

$$\begin{aligned} \mathbf{s}^T &= \left\{ \mathbf{s}_f^T \quad \alpha \right\} = \left\{ \mathbf{s}_N^T \quad \mathbf{s}_V^T \quad \mathbf{s}_M^T \quad \alpha \right\} \\ &= \left\{ N_1 \quad \dots \quad N_{n_c} \quad V_1 \quad \dots \quad V_{n_c} \quad M_1 \quad \dots \quad M_{n_c} \quad \alpha \right\}. \end{aligned} \quad (3)$$

The following conditions must be fulfilled for such system at each contact interface:

$$\begin{cases} \mu N_k + V_k \leq 0 \\ \mu N_k - V_k \leq 0 \\ d_k N_k - M_k \leq 0 \\ d_k N_k + M_k \leq 0 \end{cases}, \quad (4)$$

where μ is the static friction coefficient and $d_k > 0$ is the maximum eccentricity of the resultant contact force at the k th contact surface. The conditions in Eq. (4) imply that the axial contact forces N_k must be negative or null.

It is now considered a fictitious (conjugate) block assembly, identical to the real one presented before but characterized by cohesive-type contact interface laws. The cohesive strengths in such fictitious system are taken equal to $-\mu \mathbf{s}_N$, thus depending on the axial contact forces N_k of the real system. Henceforth, S will denote the set that collects the statically admissible equilibrium states of the real block assembly whereas K will identify the set of the kinematically admissible displacement fields. The following theorems proved in Ref. [6] establish the connection between the collapse state of the real block system and that of the fictitious one:

- given any collapse state of the frictional block assembly, the collapse load multiplier α_c is always equal to the collapse load multiplier $\alpha_{as}(\mathbf{s}_N)$ of the block assembly with fictitious associative-type contact interface laws;
- given any statically admissible equilibrium state \mathbf{s} of the frictional block assembly, if the load multiplier α is equal to any kinematically admissible load multiplier $\tilde{\alpha}_K = (\mathbf{u}, \xi, \mathbf{s}_N)$ of the fictitious system, then \mathbf{s} is a collapse state and $(\mathbf{u}, \xi) \in K$ is a collapse displacement field of the real system (\mathbf{u} collects displacements and rotations of the blocks with respect to their centroids \mathbf{x}_G^i whereas ξ collects relative displacements and rotations between the blocks).

Therefore, it is concluded that $\alpha = \alpha_{as}(\mathbf{s})$ if and only if \mathbf{s} identifies a collapse state. Hence, the minimum collapse load multiplier α_c for the frictional (real) block assembly can be determined by solving the following mathematical programming problem [6]:

$$\begin{aligned} \alpha_c &= \min \{ \alpha \} \\ &\left\{ \begin{array}{l} \mathbf{s} \in S \\ (\mathbf{u}, \xi) \in K \\ \alpha = -\mathbf{u}^T \mathbf{f}_0 - \xi^T \mathbf{N}_\mu \mathbf{s}_N \end{array} \right., \end{aligned} \quad (5)$$

where $N_\mu = [\mu \mathbf{I} \ \mu \mathbf{I} \ \mathbf{0} \ \mathbf{0}]^T$. Equation (5) provides an original approach for the limit analysis of frictional block assemblies. By solving the mathematical programming problem in Eq. (5), the minimum collapse load multiplier and the corresponding failure mode can be estimated. As far as the resolution technique is concerned, it is important to highlight that the admissible domain is not convex because of the last condition in Eq. (5). As a consequence, multiple local optima might exist and numerical resolution technique with global search capability is needed. In this work, the mathematical programming problem in Eq. (5) is solved by means of a genetic algorithm.

Numerical applications

An arch structure and two portals are first examined in order to illustrate the application of the proposed method. The j th base external force vector and the j th constant force vector are

$$\bar{\mathbf{f}}^j = \begin{Bmatrix} -W^j & 0 & 0 \end{Bmatrix}^T, \quad \bar{\mathbf{f}}_0^j = \begin{Bmatrix} 0 & -W^j & 0 \end{Bmatrix}^T, \quad (6)$$

respectively, where W^j is the weight of the j th block. The weight per unit of volume of the blocks is 1.0 whereas the friction coefficient is 0.5. A unit-width slice of the structures is analyzed. Geometry and collapse mechanisms are shown in Figs. 2-4.

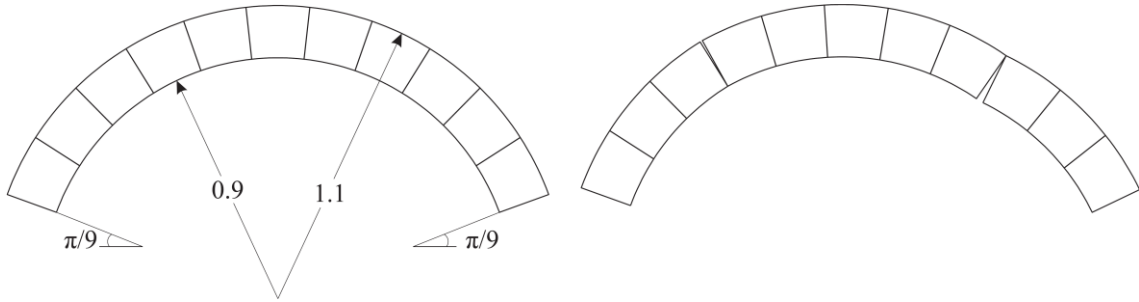


Figure 2. Geometry of the arch (left) and its collapse mechanism (right)

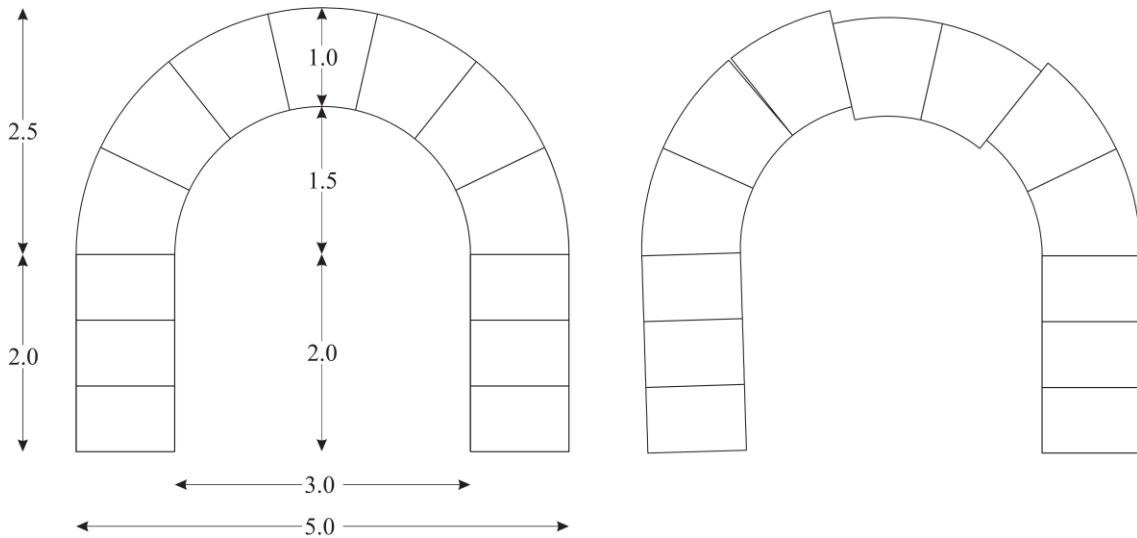


Figure 3. Geometry of the portal with constant thickness (left) and its collapse mechanism (right)

As shown in Fig. 2, the collapse of the masonry arch is based on a hinging-dominated mechanism (the corresponding collapse load multiplier is 0.63425). By counting the blocks from the left, the first hinge occurs at the intrados of the arch, between the third and the fourth

voussoirs. The second hinge takes place at the extrados of the arch, between the eighth and the ninth blocks. Finally, a third hinge occurs at the right impost of the arch, and it is placed on the intrados. The left impost also slides along its base and moves toward the outside. Combined sliding and hinging collapse modes occur for the two masonry portals, as shown in Fig. 3 and Fig. 4 (the corresponding collapse load multipliers are 0.20628 and 0.30148, respectively). In both failure mechanisms, a hinge occurs at the base of the left pier, thus causing its counterclockwise rotation. Another hinge takes place on the intrados of the arches. The sliding collapse mode involves the keystone of the arches and some voussoirs adjacent to it.

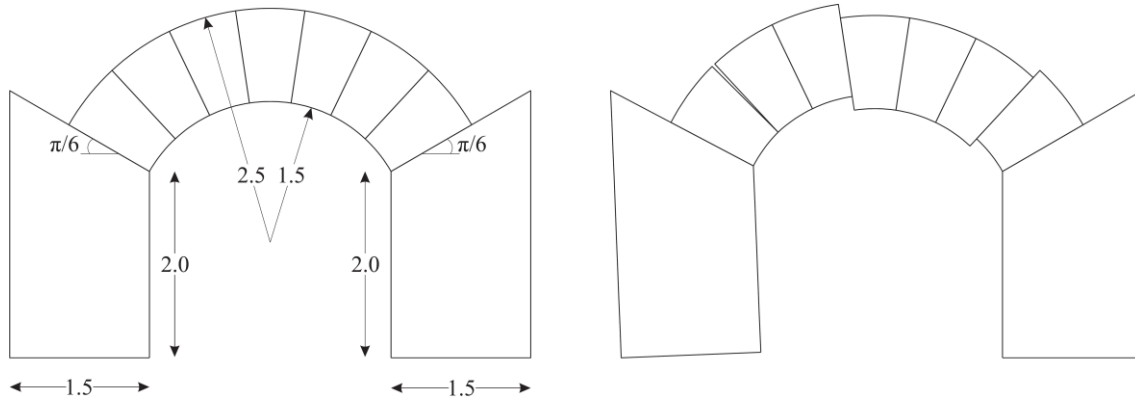


Figure 4. Geometry of the portal with large columns formed by a single block (left) and its collapse mechanism (right)

Two masonry panels are finally considered in order to demonstrate the correctness of the proposed method by comparing the corresponding collapse load multipliers with those estimated by Ferris and Tin-Loi [2], Gilbert et al. [4] and Mihai [5]. These examples are concerned with free standing walls supported on a rigid horizontal plane and subjected to in-plane forces applied to the centroid of each block. The full block size is 4×1.75 whereas the half block size is 2×1.75 . The friction coefficient is 0.65. Each full block is subjected to a vertical body force (oriented downwards), which is calculated by assuming a weight equal to 1.0. Moreover, each full block is subjected to a unit horizontal live load (directed from left to right). One panel is formed by $n_b=33$ blocks whereas the second panel is formed by $n_b=55$ blocks. The collapse mechanisms of the examined panels are shown in Fig. 5. The corresponding collapse load multipliers are listed in Tab. 1, together with reference solutions reported in some existing studies. This comparison substantiates the correctness of the proposed approach.

Table 1. Collapse load multipliers of the considered walls

Wall	Ref. [2]	Ref. [4]	Ref. [5]	Proposed approach
$n_b=33$	0.63898	0.63982	0.63945	0.63911
$n_b=55$	0.55742	0.56262	0.55751	0.55749

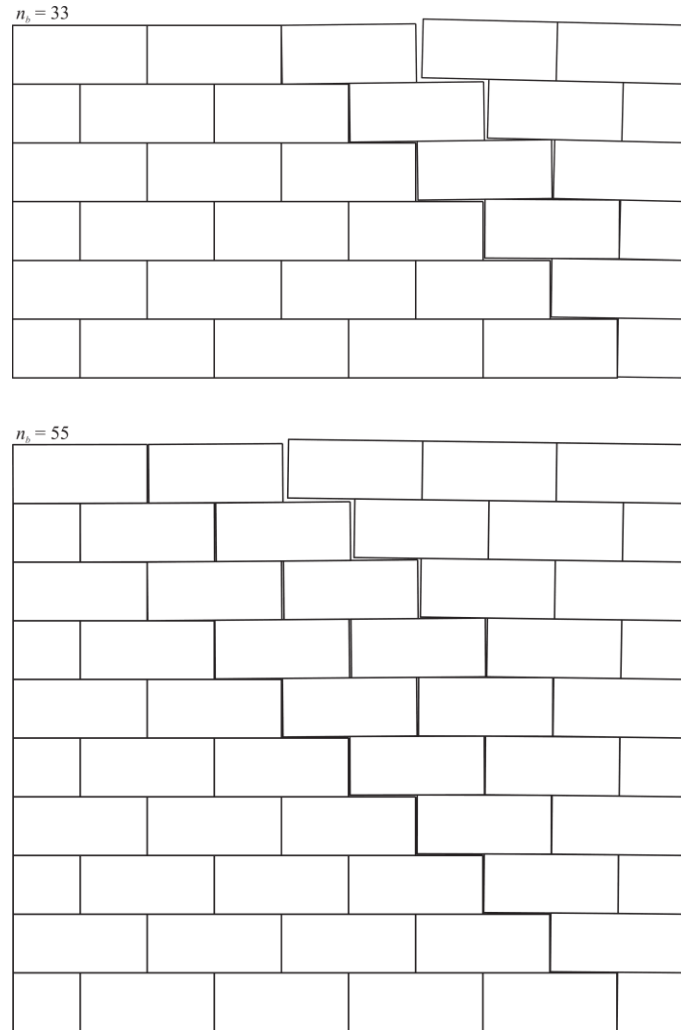


Figure 5. Collapse mechanisms of the considered walls

Conclusions

In the present work, we have illustrated an original strategy to address the limit analysis of frictional block assemblies by means of fictitious associative-type interfacial laws. Once the connection between the collapse state of the fictitious system and that of the real one has been highlighted, an original mathematical programming problem has been presented to estimate collapse load multiplier and failure mechanism. Herein, the introduction of the complementarity condition as constraint is not required because it is obtained as Karush-Khun-Tucker condition. Several numerical applications concerning with the limit analysis of masonry structures have been also included in order to demonstrate the application of the proposed approach and its correctness.

References

- [1] Baggio, C. and Trovalusci, P. (1998) Limit analysis for no-tension and frictional three-dimensional discrete systems, *Mech Struct Mach* **26**(3), 287–304.
- [2] Ferris, M. C. and Tin-Loi, F. (2001) Limit analysis of frictional block assemblies as a mathematical program with complementarity constraints, *Int J Mech Sci* **43**(1), 209–224.
- [3] Orduña, A. and Lourenço P. B. (2005) Three-dimensional limit analysis of rigid blocks assemblages. PartI: torsion failure on friction interfaces and limit analysis formulation, *Int J Solids Struct* **42**(18-19), 5161–5180.

- [4] Gilbert, M., Casapulla, C. and Ahmed, H. M. (2006) Limit analysis of masonry block structures with non-associative frictional joints using linear programming, *Comput Struct* **84**(13-14), 873–887.
- [5] Mihai, L. A. (2010) A fixed-point approach to the limit load analysis of multibody structures with Coulomb friction, *Comput Struct* **88**(13-14), 859–869.
- [6] Trentadue, F. and Quaranta, G. (2013) Limit analysis of frictional block assemblies by means of fictitious associative-type contact interface laws, *Int J Mech Sci* **70**, 140–145.

Molecular dynamics simulation of the initiation of plastic deformation in nanocrystalline material

Takuya Uehara

Department of Mechanical Systems Engineering, Yamagata University, Japan
Corresponding author: uehara@yz.yamagata-u.ac.jp

Abstract

Molecular dynamics simulations were carried out to investigate the initiation mechanism of the plastic deformation in nanocrystalline material. A polycrystalline model consisting of four grains each of which had nano-meter order size was prepared. All atoms were arranged on the face-centered cubic lattice and rotated around the [001] axis keeping the (001) face on the x - y plane. The rotation angles were set different in every grain so that the grain boundaries were formed. Then an external compressive load was imposed at a constant rate until a plastic deformation was observed apparently. Several models were prepared by varying the combination of the crystal orientations of the grains, and the change in the configuration of atoms and variation of the stress as well as their dependency on the model were investigated. Common to all models, the stress increased monotonously in the early stage, when the atomic configuration sustains the initial state. Then an abrupt drop in stress was observed. At this moment, linear or planar defects were generated, and the initiation sites were inferred on the grain boundaries. It also revealed that a plateau was observed around the stress peak when a linear defect was generated, while the peak was sharp when a planar defect was generated.

Keywords: Molecular dynamics, Plastic deformation, Dislocation, Yielding, Computer simulation

Introduction

Microstructure of a metallic material strongly affects the macroscopic properties of the material, and, for instance, the strength of polycrystalline material increases as the grain size becomes smaller. Therefore, great efforts have been devoted to refining the crystal grains, and several types of methods, such as equal-channel angular pressing (ECAP), accumulative roll bonding (ARB), and asymmetric rolling methods, have been successfully developed [1-3]. As a result of this kind of severe plastic deformation (SPD) processes, the grain size of the material falls in the submicron order, and such materials are generally termed nanocrystals. The mechanical behaviors of these materials show a characteristic tendency due to relatively higher occupancy of the grain-boundary area, and especially the plastic behavior is significantly affected. To clarify the deformation mechanism in nanocrystalline materials, an atomistic investigation is necessary, and the molecular dynamics (MD) simulation is a quite effective tool for this purpose. To date, various kinds of simulations have been demonstrated, including generation of dislocation, restructuring and migration of grain boundaries, interaction between dislocation and grain boundaries, as well as grain refining or coarsening due to mechanical loading or thermal activation [4-6]. The author also has demonstrated various simulations on the plasticity including grain-boundary migration [7], deformation of polycrystalline shape-memory alloy [8], and transformation-induced plasticity [9]. However, comprehensive understanding and construction of unified theory on the nano-scale deformation are still difficult due to the complexity of the nanocrystalline material. For

example, in relatively large scale, the grain boundary energy is often represented as a function of the misorientation angle, but the grain-boundary energy is actually different if the atomistic structure is different even though the misorientation angle is the same. Therefore, a systematic investigation using a simple model, instead of realistic large model, is considered effective to achieve the goal. The author has demonstrated MD simulations on a severe tensile or compressive deformation process of a polycrystalline material. In our previous report [10-12], change in microstructure was focused, and the grain refinement was successfully observed. In this study, the focus is concentrated on the initiation of the plastic deformation based on a similar polycrystalline model. Defects in a grain as well as their initiation and motion were made visible using the potential energy so that plastic behavior is clearly shown.

Fundamental Equations

Classical molecular dynamics method is applied in this study. The fundamental equation is the following.

$$\ddot{\mathbf{r}}_i = \frac{1}{m_i} \mathbf{F}_i = \frac{1}{m_i} \sum_j \mathbf{f}_{ij}, \quad \mathbf{f}_{ij} = -\frac{d\phi}{dr} \frac{\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|} \quad (1)$$

Here, \mathbf{r}_i and m_i are the position vector and the mass of the i -th atom, and \mathbf{F}_i is the force acting on the i -th atom. In addition, \mathbf{F}_i is assumed to be represented by the summation of the two-body interatomic force \mathbf{f}_{ij} and the force is represented by a two-body interatomic potential energy ϕ . As a representative form for a face-centered-cubic (fcc) crystal, the following Lennard-Jones-type function is applied.

$$\phi = 4D \left((R/r)^{12} - (R/r)^6 \right) \quad (2)$$

Here, D and R are the material parameters in the dimension of energy and length, respectively, while they are diminished in a non-dimensional form. In this paper, all physical parameters and variables are represented in the dimensionless values.

Simulation Model

Figure 1 represents the simulation model. Two types of nanocrystalline models consisting of four grains are prepared. A pair of grains is arranged perpendicular to the external load in Type A, while they are arranged parallel to the load in Type B. Periodic boundary conditions are imposed in the all directions; accordingly, Grain 3 separated in Fig. 1 is actually a single grain domain. The thickness of the model is five unit cells without any grain boundary. The atoms are set on the lattice points of the fcc structure so that the (001) plane is on the x - y

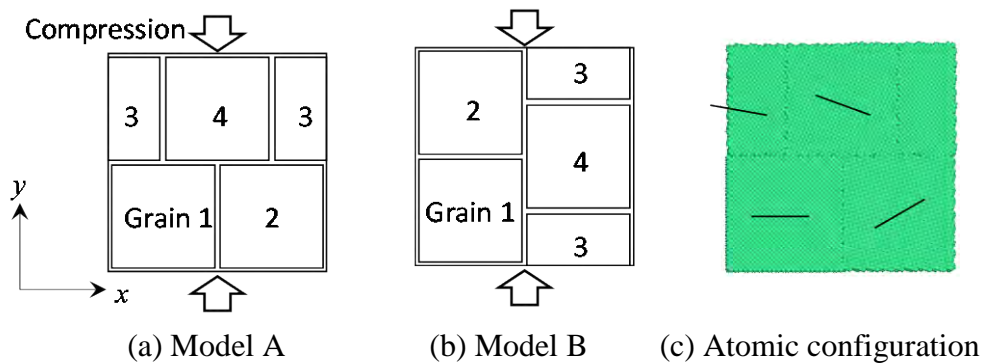


Fig. 1. Illustration of the simulation model.

plane and the [100] direction to be the x axis. Then the crystal is rotated around the z -axis, where the rotation angles θ_k of the k -th grain are set distinctively.

Simulation Conditions

The combination of the rotation angles of the four grains is set variously, and the three cases shown in Table 1 are presented in this paper. Each simulation model is identified by the symbol combined with the type of grain arrangement, A or B; e.g. Model Az1. Fig. 1 (c) represents the initial configuration of atoms for Model Az1, where short lines drawn in every grain show the [100] orientation. A significant irregularity is induced at the grain-boundary area by this operation, and the 15000 time-steps are devoted for relaxation of the arrangement. Then an external compression is imposed by reducing the length in the y -direction L_y of the model at a constant rate. Then the average normal components of the stress in the lateral directions, x and z , are kept constant at zero by adjusting the edge length L_x and L_z . The compression rate is set as $\Delta L_y = 1.0$ per 8000 time steps (L_y is approximately 73.0 for all models).

Table 1. Rotation angles in each grain.

Model ID	Grain No			
	1	2	3	4
z1	0	30	-10	-20
z2	0	-20	30	-10
z3	0	-10	-20	30

Simulation Results --- Model Az1

Simulation result for Model Az1 is shown in Fig. 2, which represents the variation of configuration of atoms during compressive deformation process. The color indicates the potential energy of each atom, where red and blue represent the maximum and minimum values, respectively, and the intermediate colors are continuously allocated. Accordingly, grain boundaries are depicted in green or yellow.

Figure 2(a) exhibits the configuration of atoms at the beginning of loading after relaxation. Every grain maintains the original shape of square, though the top-right grain boundary between Grains 3 and 4 is rather curved. Overall, no specific change in grain arrangement and

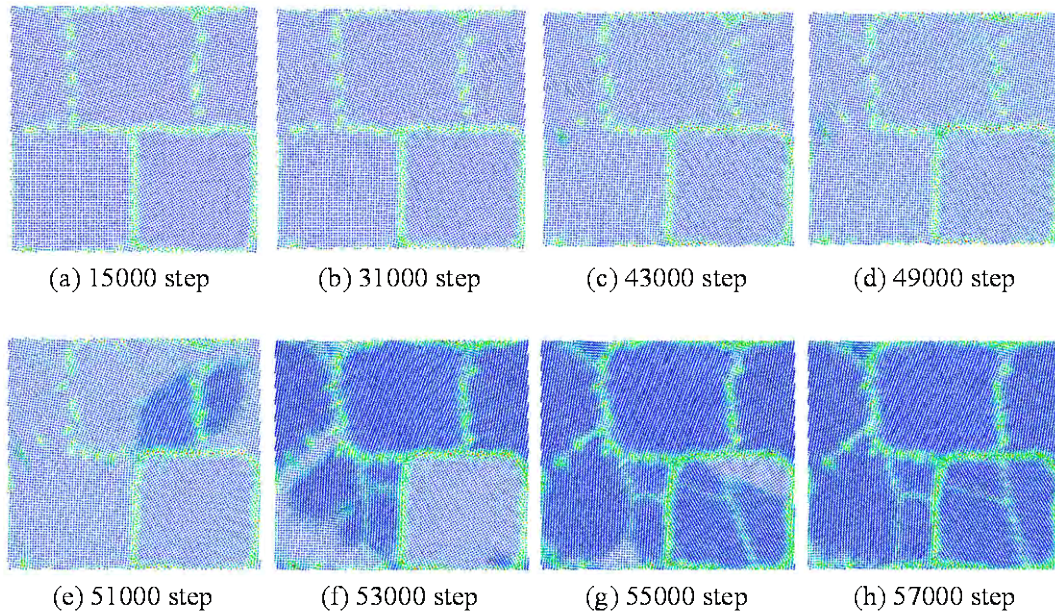


Fig. 2. Variation of the configuration of atoms for Model Az1. The color indicates the potential energy of each atom.

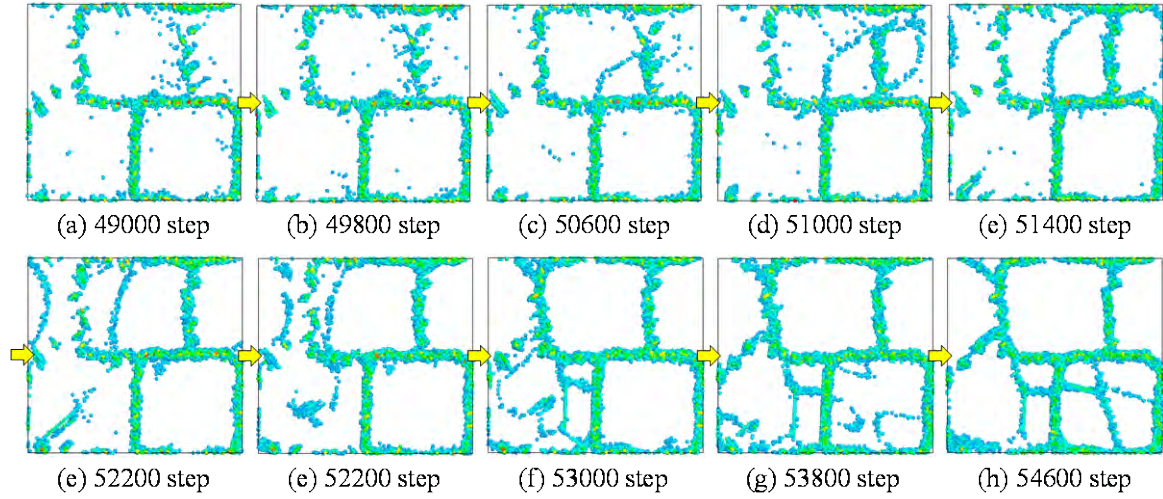


Fig. 3. Variation of atomic configuration in the duration of initiation of the plastic deformation for Model Az1. Only the atoms having potential energy valued higher than a certain threshold, here $\phi = -7.8$.

crystal structure is observed until the 49000th time step, while slight variations are observed in the grain boundaries. Subsequently, an apparent change is observed at the 51000th time step, as shown in Fig. 2(e): an area depicted in dense blue color appears around the top-right area. This is caused by the change in the crystal orientation in the depth direction; the [001] direction completely accorded with the z -axis originally, but the orientation shifted slightly and the atoms in the depth direction become visible. It is also notable that this area is centered at a grain boundary, and spread across two grains. This area extends quickly over Grains 3 and 4. A new dense-blue area is also generated in Grain 1 by the 53000th time step. Subsequently, similar change appeared in Grain 2, and finally a series of the change complete by the 57000th time step.

To see the phenomena occurring more clearly, only the atoms at unstable state are made visible and displayed at short intervals between the 49000 and 55000th time steps in Fig. 3. Values in the potential energy are taken as the indicator of the instability, and the threshold value is chosen as -7.8 . Most unstable atoms are located on the grain boundaries except for

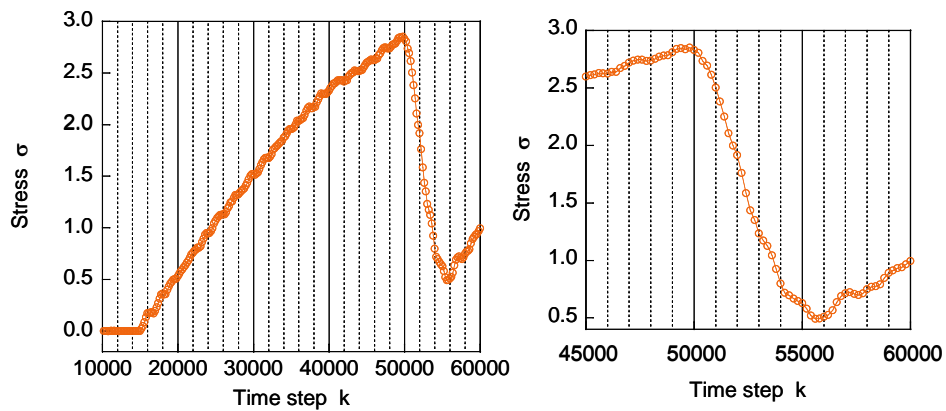


Fig. 4. Variation of stress in the y direction for Model Az1. (a) Overall view from the beginning of loading and until the end of yielding. (b) Magnified view around the yielding.

several atoms scattered inner grains at the 49000th and 49800th time steps. Then a linear defect appears in the top-center grain (Grain 3) at the 50600th time step, and the defect spread in Grain 3. Compared with Fig. 2 (e) and Figs. 3 (c) and (d), the apparent linear defect in Fig. 3 corresponds to the boundary of the dense-blue area in Fig. 2, and actually the defect is planer one. In addition, the origin of the defect is the right-bottom corner of the grain, or the triple junction of the grains, and the transformed area spread in Grain 3. Then the transformed area crosses over the grain boundary and extended in Grain 4.

The variation in the compressive stress is shown in Fig. 4. The stress increases monotonously as compressive load is imposed from the 15000th time step. Then the stress decreases abruptly at the 50000th time step. The peak point is very sharp, but the decrease is not momentary; it takes about 5000 time steps until the stress reaches the bottom. The initiation of the defects was determined from Fig. 3 to be the 50000th time step, which clearly corresponds to the peak in Fig. 4. Additionally, it is also confirmed that the duration of the expansion of the deformed area corresponds well to the duration of the stress decrease, from the 50000th to 55000th time step. Therefore, it can be considered that yielding occurred in this duration.

Simulation Results --- Other Cases

Simulation results for all cases are shown in Fig. 5. Variations in the stress show similar tendency; the stress increases monotonously and decreases drastically at around the 50000th time step. The peak values in stress are different for every case; minimum $\sigma = 2.5$ for Model Bz1, and maximum $\sigma = 3.1$ for Model Az3. The reduction depth and duration spent for the stress drop are approximately the same. An overall tendency that Models A exhibits higher peak than Models B can be seen, though more data is necessary for draw a conclusion. Another interesting feature is that some cases, for example Models Az2 and Bz1, show some plateau duration around the peak. Figure 5 (b) shows the configuration of unstable atoms just after the stress peak for some typical cases. Common to Models Az2 and Bz1 both of which exhibit a plateau around the stress peak, a straight line defect is generated, as shown in Figs 5(b)(i) and (iii). In other cases for which a sharp peak is observed, in contrast, a

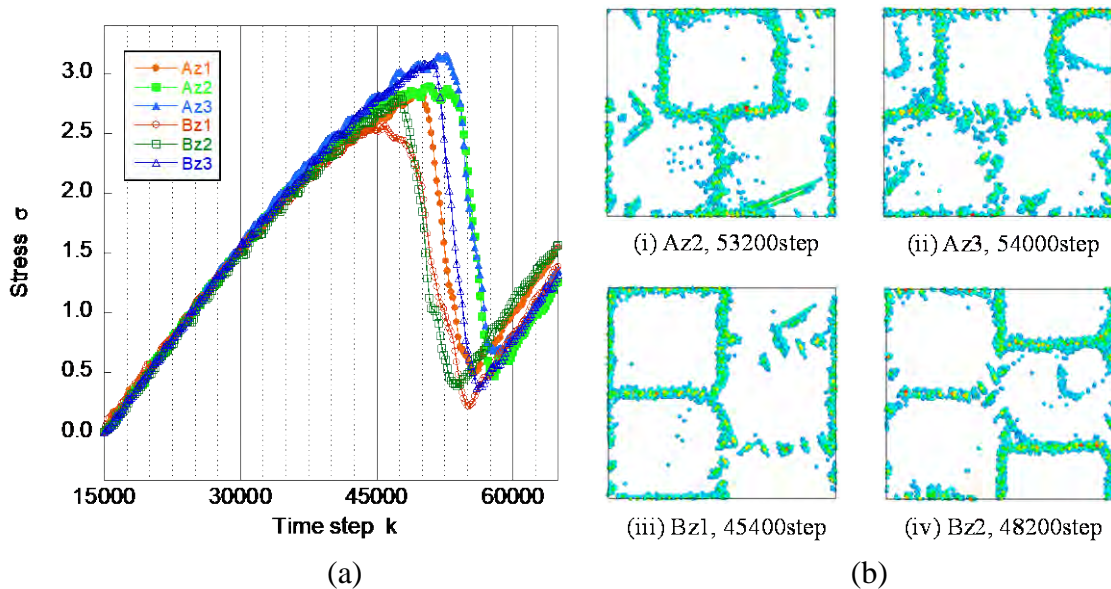


Fig. 5. Variation in stress for all models (a), and the configuration of unstable atoms around the initiation of the plastic deformation for some typical cases (b).

transformation domain surrounded by curved defects are generated, as shown in Figs 5(b)(ii) and (iv) for Az3 and Bz2 as well as Fig. 3(c) for Az1. These defects correspond to planar defects as can be seen in Fig. 2(e). The cause of generation of different type of defects, linear or planar ones, is inferred to be the structure of grain boundary, since the origins of both types of defects are considered to be on the grain boundaries. Additionally, the crystal orientation of the grains in which linear defects are generated, i.e. Grain 2 in Model Az2 and Grain 4 in Model Bz1, is commonly -20 deg., as listed in Table 1. The orientation with respect to the external load is also one of the key factors dominating the behavior, which will be discussed in our future work in detail.

Conclusion

Molecular dynamics simulations were carried out to investigate the initiation mechanism of plastic deformation in a nanocrystalline material. Two types of simple models having four grains with different arrangements were prepared, and a uniaxial compression was imposed. Simulations were carried out varying the combination of the crystal orientations of every grain. As a result, the following results were obtained. Common to all models, the stress increased monotonously, when the grain sustains the initial configuration. Then, an abrupt drop in stress was observed. At this moment, a linear or planar defect was generated, and this is considered the initiation of the plastic deformation. The defects spread in a certain time duration, which is considered as yielding. Further investigation is still required for clarifying the mechanism in more detail. Full three-dimensional model is necessary to simulate slip phenomena on the (111) plane. Application of more precise inter-atomic potential is required not only for quantitative evaluation but also for distinguishing some more delicate mechanism. However, it revealed that fundamental insights can be derived using the present model, and further investigation is now in progress.

References

- [1] Segal, V.M. (1999) Equal channel angular extrusion: from macromechanics to structure formation. *Materials Science*, **54**, 427-510.
- [2] Saito, Y., Utsunomiya, H., Tsuji, N. and Sakai, T. (1999) Novel ultra-high straining process for bulk materials --- Development of the accumulative roll-bonding (ARB) process. *Acta Materialia*, **47**, 579-583.
- [3] Sidor, J., Miroux, A., Petrov, R. and Kestens, L. (2008) Microstructural and crystallographic aspects of conventional and asymmetric rolling processes. *Acta Materialia*, **56**, 2495-2507.
- [4] Van Swygenhoven, H., Caro, A. and Farkas, D. (2001) A molecular dynamics study of polycrystalline fcc metals at the nanoscale: grain boundary structure and its influence on plastic deformation. *Materials Science and Engineering A*, **309-310**, 440-444.
- [5] Trautt, Z.T. and Mishin, Y. (2012) Grain boundary migration and grain rotation studied by molecular dynamics. *Acta Materialia*, **60**, 2407-2424.
- [6] Spearot, D.E. and Sangid, M.D. (2014) Insights on slip transmission at grain boundaries from atomistic simulations. *Current Opinion in Solid State and Materials Science*, **18**, 188-195.
- [7] Uehara, T., Wakabayashi, N., Hirabayashi, Y. and Ohno, N. (2008) An atomistic study of grain boundary stability and crystal rearrangement using molecular dynamics techniques. *International Journal of Mechanical Sciences*, **50**, 956-965.
- [8] Uehara, T., Asai, C. and Ohno, N. (2009) Molecular dynamics simulation of shape-memory behaviour using a multi-grain model. *Modelling and Simulation in Materials Science and Engineering*, **17**, 035011.
- [9] Uehara, T. (2015) Molecular dynamics simulation on transformation-induced plastic deformation using a Lennard-Jones model. *Key Engineering Materials*, **626**, 414-419.
- [10] Uehara, T. (2017) Molecular dynamics simulation of the variation in the microstructure of a polycrystalline material under tensile load. *Key Engineering Materials*, **748**, 375-380.
- [11] Uehara, T. (2017) Molecular dynamics simulation of microstructural change in a poly-crystalline FCC metal under compression. *Proceedings of XIV Int. Conf. on Computational Plasticity*, 106-113.
- [12] Uehara, T. (2017) Molecular dynamics simulation of grain refinement in a polycrystalline material under severe compressive deformation, *Materials Sciences and Applications*, **8**, 918-932.

2-D inverse scattering analysis using pure SH wave for delamination in carbon fiber reinforced plastic

†Takahiro Saitoh¹, Takashi Onodera¹, Akira Furukawa² and *Sohichi Hirose²

¹Department of Civil and Environmental Engineering, Gunma University, Japan

²Department of Civil and Environmental Engineering, Tokyo Institute of Technology, Japan

*Presenting author: hirose.s.aa@m.titech.ac.jp

†Corresponding author: t-saitoh@gunma-u.ac.jp

Abstract

In this paper, a linearized inverse scattering technique with the aid of the convolution quadrature time-domain boundary element method (CQBEM) has been developed for the reconstruction of a delamination in carbon fiber reinforced plastic (CFRP) with anisotropic property. The CQBEM is utilized to obtain scattered wave data from a delamination in CFRP. The wave forms obtained by the CQBEM are adequately treated to implement the shape reconstruction of a delamination in CFRP. The Kirchhoff approximation is applied to the unknown delamination opening displacement. A far-field approximation of the 2-D fundamental solution in frequency-domain for general anisotropic elastodynamics is used for the proposed inverse scattering formulation. Numerical examples for a delamination in various types of CFRPs are shown to verify the proposed method.

Keywords: Time-domain BEM, inverse scattering analysis, anisotropic elastodynamics, carbon FRP

Introduction

Some anisotropic materials have attracted lots of interest in the fields of the mechanical and civil engineering in recent years. The carbon fiber reinforced plastic (CFRP) is known as one of the typical anisotropic materials, and is generally used as a material of construction for bridges and aircrafts, because the FRP has the characteristics of high tension strength, corrosive resistance and light weight. The ultrasonic non-destructive testing is most widely used in order to provide evidence of safety for structural materials. The exact identification of position, size, and shape of a defect in materials is an important factor for structural monitoring and health diagnostics. Defect shape reconstruction methods for materials have been developed by several researchers since several years ago [1][2]. The inverse scattering is an effective defect shape reconstruction method, and has been applied to many engineering problems [3][4][5]. However, no numerical example using the inverse scattering method can be seen for the reconstruction of a defect in CFRP with the anisotropic property. The acoustic anisotropic property makes it difficult for the nondestructive engineers to evaluate a defect. Therefore, in this study, an inverse scattering technique is developed for a defect shape reconstruction for CFRP with anisotropic property. The pure SH wave mode is only considered in this study for simplicity. The convolution quadrature time-domain boundary element method (CQBEM) [6][7][8][9] is utilized to obtain the scattered wave data from a delamination in a CFRP, which is required for the implementation of the inverse scattering formulation. The proposed inverse scattering formulation is achieved in the frequency-domain. Therefore, the scattered wave data in frequency-domain are calculated by using the Fourier transform of those in time-domain

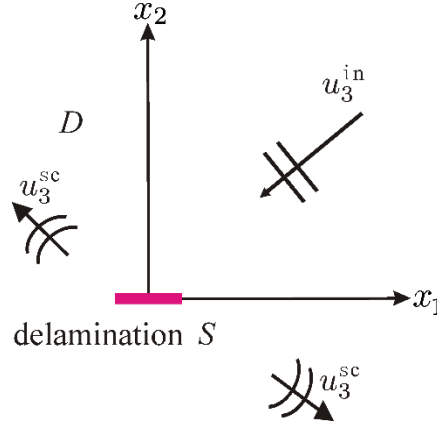


Figure 1 Analysis model.

obtained by CQBEM. In the following sections, the problem statement and proposed inverse scattering formulation are discussed. Some numerical results for the shape reconstruction of a delamination in various types of CFRPs are presented. Finally, some comments and our future research plans are remarked.

Problem statement

The proposed 2-D inverse scattering formulation using pure SH wave is based on the Kirchhoff approximation [10]. Some important equations for the study on this inverse scattering for a delamination in CFRP are shown in this section, because of the page limitation. In this research, we assume that the pure SH wave is generated by the interaction between the incident wave $u_3^{\text{in}}(\mathbf{x}, t)$ and a delamination S in CFRP, as shown in Fig.1, namely, the elastic waves generated in CFRP can be decomposed into the pure SH wave and in-plane wave modes. The equation of motion and constitutive equation at the position \mathbf{x} and time t for the anisotropic elastodynamics are defined as follows:

$$\rho \ddot{u}_i(\mathbf{x}, t) = \sigma_{ij,j}(\mathbf{x}, t) \quad (1)$$

$$\sigma_{ij}(\mathbf{x}, t) = C_{ijkl} u_{k,l}(\mathbf{x}, t) \quad (2)$$

where σ_{ij} is the stress, ρ is the density of CFRP, $u_i(\mathbf{x}, t)$ is the displacement, $(\cdot)_{,i}$ is the partial derivative with respect to $\partial/\partial x_i$, and $(\cdot)^{\cdot}$ is the time derivative. In addition, C_{ijkl} is the elastic constant. The fourth order elastic constant C_{ijkl} is related to the Voigt notation elastic constant C_{IJ} ($i, j = 1, \dots, 6$) [10] expressed by

$$I = \begin{cases} i & : i = j \\ 9 - (i + j) & : i \neq j \end{cases}, \quad J = \begin{cases} k & : k = l \\ 9 - (k + l) & : k \neq l \end{cases} \quad (3)$$

2-D inverse scattering formulation using pure SH wave

The delamination must be carefully taken into consideration during CFRPs in-service period. In the frequency domain, the boundary integral equation for the scattered wave $u_3^{\text{sc}}(\mathbf{x}, \omega)$ with the time-harmonic frequency ω in infinite domain D , as shown in Fig.1, can be written as follows:

$$u_3^{sc}(\mathbf{x}, \omega) = - \int_S C_{3\alpha 3\beta} e_\alpha(\mathbf{y}) \frac{\partial U_{33}(\mathbf{x}, \mathbf{y}, \omega)}{\partial y_\beta} [u_3(\mathbf{y}, \omega)] dS_y \quad (4)$$

where $[u]_3$ and e_α show the delamination opening displacement for anti-plane direction and the unit normal vector with respect to the outer normal direction on \mathbf{y} , respectively. In addition, $U_{33}(\mathbf{x}, \mathbf{y}, \omega)$ denotes the traction fundamental solution for 2-D anti-plane anisotropic elastodynamics in frequency-domain. The fundamental solution $U_{33}(\mathbf{x}, \mathbf{y}, \omega)$, derived by Wang and Achenbach [12], is given as follows:

$$U_{33}(\mathbf{x}, \mathbf{y}, \omega) = \frac{1}{8\pi^2} \int_{|\mathbf{n}|=1} \frac{1}{\rho c^2(\mathbf{n})} \phi(k(\mathbf{n})|\mathbf{n} \cdot (\mathbf{x} - \mathbf{y})|) d\mathbf{n} \quad (5)$$

where $c(\mathbf{n})$ is the phase velocity with respect to the direction \mathbf{n} over the unit sphere and k is the wave number defined by $k(\mathbf{n}) = \omega/c(\mathbf{n})$. The fundamental solution $U_{33}(\mathbf{x}, \mathbf{y}, \omega)$ involves the numerical integration over the unit circle with respect to $|\mathbf{n}| = 1$. The function $\phi(\xi)$ is defined by

$$\phi(\xi) = i\pi e^{i\xi} - 2\{\cos(\xi) \text{ci}(\xi) + \sin(\xi) \text{si}(\xi)\}. \quad (6)$$

In eq. (6), the functions $\text{si}(\xi)$ and $\text{ci}(\xi)$ are Sine and Cosine integrals, respectively, which are defined as follows:

$$\text{si}(\xi) = - \int_\xi^\infty \frac{\sin(s)}{s} ds, \quad \text{ci}(\xi) = - \int_\xi^\infty \frac{\cos(s)}{s} ds. \quad (7)$$

The numerical evaluation of the integration over the unit sphere in eq. (5) is very time-consuming. Therefore, a far-field approximation is introduced to decrease the required computational time. In addition, the use of a far-field approximation allows us to achieve the inverse scattering formulation. If the observation point \mathbf{x} is far enough from the source point \mathbf{y} , the fundamental solution $U_{33}(\mathbf{x}, \mathbf{y}, \omega)$ can be approximated by using the stationary phase method as follows:

$$U_{33}(\mathbf{x}, \mathbf{y}, \omega) = \frac{i}{C_{44}} \sqrt{\frac{1}{8\pi k_0 |\mathbf{x}| |f''(\varphi^s)|}} S^2(\varphi^s) \cdot \exp[ik_0(|\mathbf{x}| - \hat{\mathbf{x}} \cdot \mathbf{y})f(\varphi^s) + i\frac{\pi}{4} \text{sgn}\{f''(\varphi^s)\}] \quad (8)$$

where $\hat{\mathbf{x}}$ is the unit vector of \mathbf{x} and k_0 is given by $k_0 = \omega/c_0$. c_0 is given by $c_0 = \sqrt{C_{44}/\rho}$. φ^s and ψ satisfy $f'(\varphi^s) = 0$ and $(\cos\psi, \sin\psi) = (\mathbf{x} - \mathbf{y})/|\mathbf{x} - \mathbf{y}|$, respectively. In addition, $S(\varphi) = c_0/c(\varphi)$ and, $f(\varphi) = S(\varphi)\cos(\varphi - \psi)$. The symbol “sgn” shows the sign function. Equation (8) is the far-field approximation of the fundamental solution $U_{33}(\mathbf{x}, \mathbf{y}, \omega)$. Substituting eq. (8) into eq. (4), we can obtain

$$\begin{aligned}
 u_3^{sc}(\mathbf{x}, \omega) = & -\frac{f(\varphi^s)}{C_{44}} \sqrt{\frac{k_0}{8\pi|\mathbf{x}||f''(\varphi^s)|}} S^2(\varphi^s) \\
 & \cdot \exp \left[ik_0|\mathbf{x}|f(\varphi^s) + i\frac{\pi}{4} \text{sgn}\{f''(\varphi^s)\} \right] \\
 & \cdot \int_S e_\alpha(\mathbf{y}) \exp\{-ik_0f(\varphi^s)\hat{\mathbf{x}} \cdot \mathbf{y}\} [u_3(\mathbf{y}, \omega)] dS_y.
 \end{aligned} \tag{9}$$

In eq. (9), the delamination opening displacement $[u_3(\mathbf{y}, \omega)]$ is unknown. Therefore, $[u_3(\mathbf{y}, \omega)]$ can be approximated by using the Kirchhoff approximation, which approximates the unknown delamination opening displacement as the sum of the incident wave $u_3^{in}(\mathbf{x}, \omega)$ and the reflected wave from the delamination. In addition, the singular function $\gamma(\mathbf{y})$, which has the characteristic of

$$\int_D \gamma(\mathbf{y}) dV_y = \int_S dS_y, \tag{10}$$

is considered. Applying the Kirchhoff approximation to eq. (9), and using both Gauss's divergence theorem and the singular function $\gamma(\mathbf{y})$ defined in eq. (10) yield the following equation:

$$\begin{aligned}
 u_3^{sc}(\mathbf{x}, \omega) = & \frac{if(\varphi^s)F(\omega)}{C_{44}} \sqrt{\frac{k_0}{2\pi|\mathbf{x}||f''(\varphi^s)|}} C_{3\alpha 3\beta} \hat{x}_\beta S^2(\varphi^s) (k_0f(\varphi^s)\hat{x}_\alpha - k\hat{d}_\alpha^{in}) \\
 & \cdot \exp \left[ik_0|\mathbf{x}|f(\varphi^s) + i\frac{\pi}{4} \text{sgn}\{f''(\varphi^s)\} \right] \\
 & \cdot \int_D \gamma(\mathbf{y}) \exp[-i\{k_0f(\varphi^s)\hat{\mathbf{x}} - k\hat{\mathbf{d}}^{in}\} \cdot \mathbf{y}] dV_y
 \end{aligned} \tag{11}$$

where $\hat{\mathbf{d}}^{in}$ denotes the propagation vector of the incident wave. In eq. (11), the Ricker wavelet [13] is considered as the incident wave $u_3^{in}(\mathbf{x}, \omega)$. The Ricker wave in frequency-domain, $F(\omega)$, is defined by

$$F(\omega) = -\frac{\sqrt{2\pi}\omega^2 \exp(i\omega t_s)}{2\exp(\omega^2/\omega_p^2)\omega_p^3} \tag{12}$$

where ω_p and t_s show the peak frequency and peak location of the Ricker wavelet, respectively. In eq. (11), the singular function $\gamma(\mathbf{y})$ is the Fourier transform with respect to $K = k_0f(\varphi^s)\hat{\mathbf{x}} - k\hat{\mathbf{d}}^{in}$. Therefore, the singular function $\gamma(\mathbf{y})$, which shows the delamination surface, can be obtained by the inverse Fourier transform as follows:

$$\begin{aligned}
 \gamma(\mathbf{y}) = & -iC_{44} \int_0^{2\pi} \int_0^\infty \left[\frac{f(\varphi^s)}{c_0} - \frac{1}{c} \cos(\psi - \psi^{in}) \right] \\
 & \cdot \frac{u_3^{sc}(\mathbf{x}, \omega)}{F(\omega)C_{3\alpha 3\beta} \hat{x}_\beta S^2(\varphi^s) (k_0f(\varphi^s)\hat{x}_\alpha - k\hat{d}_\alpha^{in})}
 \end{aligned} \tag{13}$$

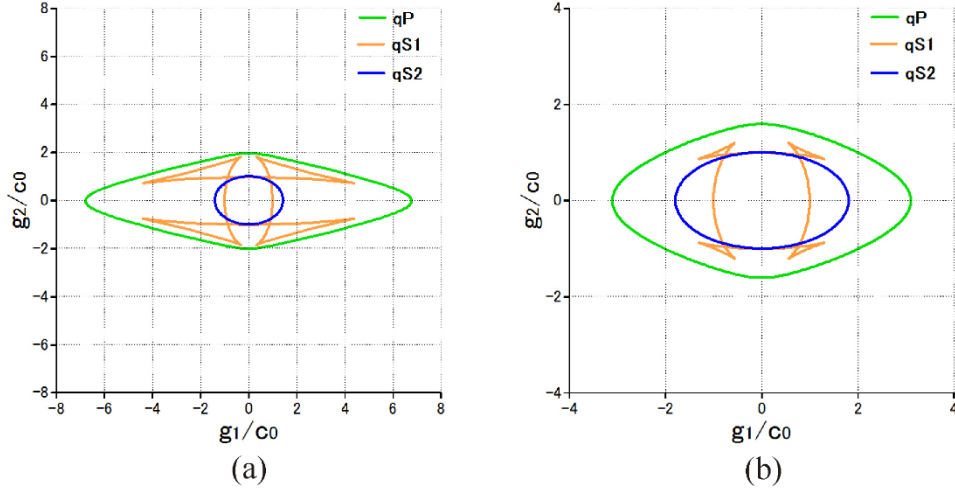


Figure 2 Group velocity curves for (a) unidirectional CFRP (b) quasi-isotropic CFRP.

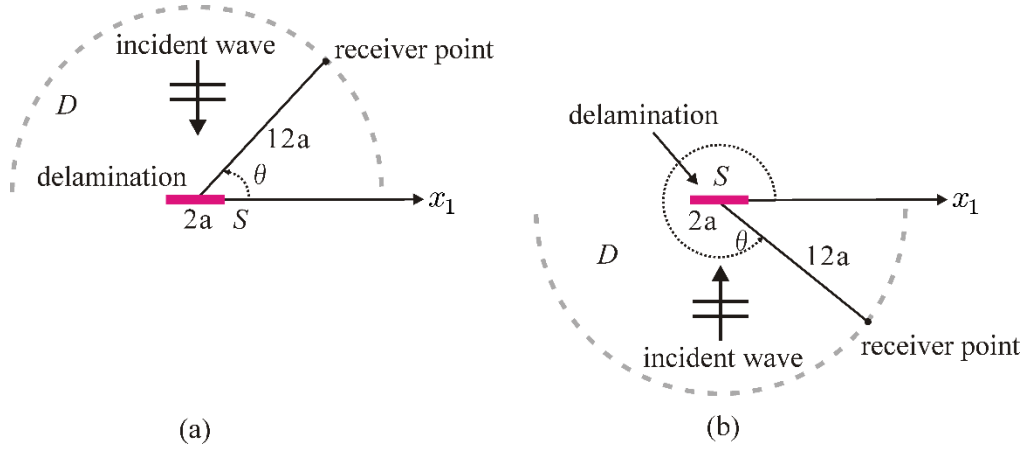


Figure 3 Forward and inverse scattering analysis models (a) downward and (b) upward incidences.

$$\begin{aligned}
 & \cdot \sqrt{\frac{k_0 |\mathbf{x}| |f''(\varphi^s)|}{8\pi^3}} \exp \left[-ik_0 |\mathbf{x}| f(\varphi^s) - i\frac{\pi}{4} \text{sgn}\{f''(\varphi^s)\} \right] \\
 & \cdot \exp[i\{k_0 f(\varphi^s) \hat{\mathbf{x}} \cdot \mathbf{y} - k \hat{\mathbf{d}}^{\text{in}} \cdot \mathbf{y}\}] d\omega d\psi
 \end{aligned}$$

where ψ^{in} is the incident wave angle. The shape reconstruction of the delamination is implemented by the calculation of the right-hand side of eq. (13).

Numerical results

In this section, some numerical results for the shape reconstruction of a delamination in two types of CFRPs are demonstrated by using the proposed method. The two types of CFRPs are unidirectional and quasi-isotropic CFRPs. The elastic constants of them are given in the normalized form by C_{44} as follows:

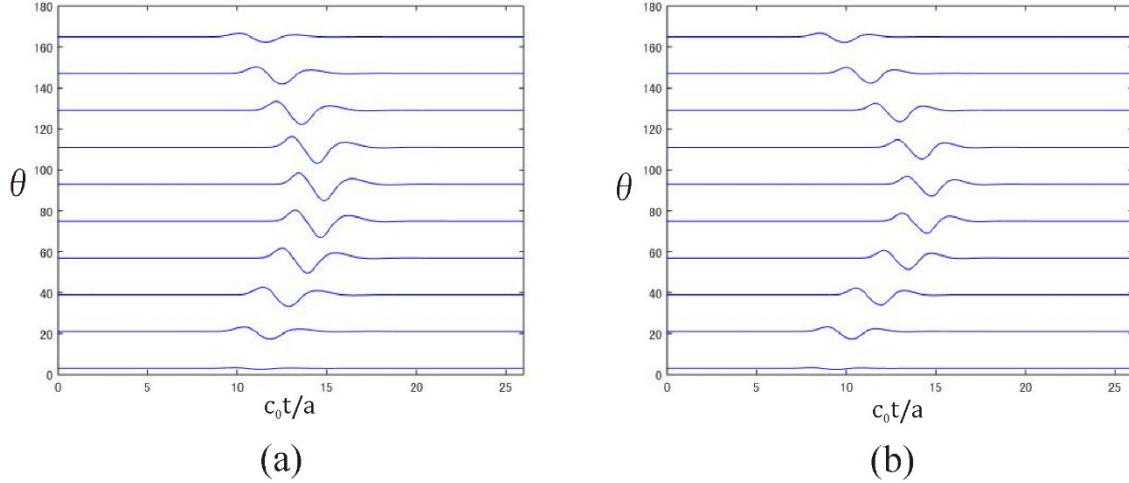


Figure 4 Scattered wave forms $u_3^{sc}(x, t)$ obtained by CQBEM for the case of (a) unidirectional CFRP (b) quasi-isotropic CFRP.

$$\frac{c_{\alpha\beta}}{c_{44}} = \begin{pmatrix} 45.914 & 1.829 & 41.874 & 0 & 0 & 0 \\ & 3.977 & 1.829 & 0 & 0 & 0 \\ & & 45.914 & 0 & 0 & 0 \\ & & & 1.0 & 0 & 0 \\ \text{sym.} & & & & 2.02 & 0 \\ & & & & & 1.0 \end{pmatrix} \text{ (unidirectional CFRP)} \quad (14)$$

$$\frac{c_{\alpha\beta}}{c_{44}} = \begin{pmatrix} 9.63 & 0.77 & 4.0 & 0 & 0 & 0 \\ & 2.54 & 0.77 & 0 & 0 & 0 \\ & & 9.63 & 0 & 0 & 0 \\ & & & 1.0 & 0 & 0 \\ \text{sym.} & & & & 3.24 & 0 \\ & & & & & 1.0 \end{pmatrix} \text{ (quasi-isotropic CFRP)} \quad (15)$$

Figure 2(a) and (b) show the group velocity curves for unidirectional CFRP and quasi-isotropic CFRP, respectively. As shown in Fig.2, three distinct waves, the qP wave (longitudinal wave), and qS1 and qS2 waves (shear waves), exist in each CFRP. In addition, the qP wave, which is faster than qS1 and qS2 waves, are observed. The velocity of the qP wave for the horizontal direction is faster than that for the vertical direction, due to the anisotropic property. In this analysis, the qS2 wave, which is called pure SH wave, is used to reconstruct a delamination in CFRPs. The scattered wave data $u_3^{sc}(x, \omega)$ of eq. (13) can be calculated by using the Fourier transform of $u_3^{sc}(x, t)$ obtained by the CQBEM.

Forward analysis results obtained by CQBEM

The results for 2-D elastic wave scattering by a delamination with the length $2a$ in CFRPs are demonstrated in this section. Figure 3 shows the forward analysis model and the scattered waves $u_3^{sc}(x, \omega)$ at several receiver points, which are away from the center of the delamination by $12a$, are calculated with the aid of the CQBEM. In this analysis, two cases which are downward and upward incidences for the delamination, as shown in Fig.3(a) and (b), respectively, are considered. The delamination is discretized by the piecewise constant boundary elements and the number of boundary elements M is given by $M = 20$. The time increment $c_0 t/a$, the number of total time steps N , and the central frequency of the Ricker wavelet ω_p are given by

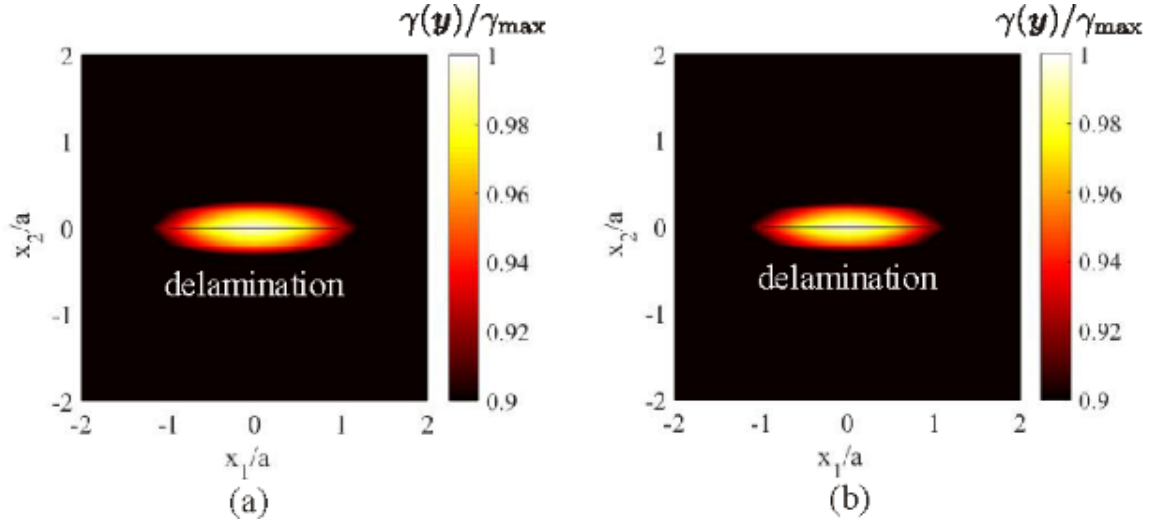


Figure 5 Shape reconstruction results using the proposed inverse scattering technique for the delamination in (a) unidirectional CFRP (b) quasi-isotropic CFRP.

$c_0 t/a = 0.02$, $N = 2048$, and $\omega_p = \pi$, respectively. Figure 4(a) and (b) show time variation of scattered wave forms at the receiver points, $(r, \theta) = (12a, \theta = 3^\circ + 18^\circ n \ (n = 0, \dots, 9))$ in Fig.3(a). The unidirectional and quasi-isotropic CFRPs whose elastic constants are given in eq.(14) and (15) are considered for Fig.4(a) and (b), respectively. The time-domain transformed wave for the Ricker wave defined in eq. (12) is considered for this analysis. We can see that the scattered waves $u_3^{sc}(\mathbf{x}, t)$ arrive at different times for each receiver point, as shown in Fig.4, due to the anisotropic property of CFRPs. The shapes of the group velocity curves for both CFRPs are elliptical, which are shown by blue lines in Fig.2, and the group velocity of qS2 (pure SH wave) for horizontal direction is faster than that for vertical one. These scattered wave forms $u_3^{sc}(\mathbf{x}, t)$ can be used for the following inverse scattering analysis.

Inverse scattering analysis results

The shape reconstruction results by the proposed method are demonstrated in this section. As mentioned before, the scattered wave forms $u_3^{sc}(\mathbf{x}, t)$ in time-domain can be obtained by using the CQBEM. However, scattered wave forms $u_3^{sc}(\mathbf{x}, \omega)$ in frequency-domain are required for the computation of right-hand side of eq. (13). The scattered wave forms $u_3^{sc}(\mathbf{x}, \omega)$ in frequency-domain are calculated by using the Fourier transform of those $u_3^{sc}(\mathbf{x}, t)$ in time-domain in this research. Figure 5(a) and (b) show the results for the shape reconstruction of the delamination in unidirectional and quasi-isotropic CFRPs, respectively. The singular function $\gamma(\mathbf{y})/\gamma_{max}$, γ_{max} is the maximum value of γ , is plotted around the delamination. The central straight black line in Fig.5 denotes the actual delamination shape and position. The scattered wave forms $u_3^{sc}(\mathbf{x}, t)$ at the receiver points $(r, \theta) = (12a, \theta = 3^\circ + 18^\circ n \ (n = 0, \dots, 9))$ for the downward incidence and $(r, \theta) = (12a, \theta = 183^\circ + 18^\circ n \ (n = 0, \dots, 9))$ for upward incidence, as shown in Fig.3 (a) and (b), respectively, are used for this inverse scattering analysis for the delamination. We can see that the singular function $\gamma(\mathbf{y})/\gamma_{max}$ shows large values around the delamination in Fig.5. Therefore, our proposed inverse scattering technique has the potential to realize the identification of an unknown delamination in various types of CFRP with anisotropic property.

Conclusions

In this study, the inverse scattering technique for the reconstruction of a delamination in CFRP was proposed. The mathematical formulation for the proposed technique was derived, and tested numerically to verify the proposed method by solving the fundamental inverse scattering problem for the delamination in various types of CFRPs. In this study, only the pure SH wave (qS2 wave) was used for the reconstruction of the delamination in CFRPs. Therefore, in the future, we will try to implement the shape reconstruction using the qP wave. In addition, the extension to 3-D problem is also our next challenge.

Acknowledgements

This work is supported by JSPS KAKENHI (17H0329400) and "Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures"(jh180049) in Japan.

References

- [1] Doctor, S. R., Hall, T. E. and Reid, L.D. (1986) Saft - the evolution of a signal processing technology for ultrasound testing, *NDT Int.*, **19**(3), 163-167.
- [2] Colton, D., Coyle, J. and Monk, P. (2000) Recent developments in inverse acoustic scattering theory, *SIAM Review*, **42**(3), 369-414.
- [3] Kitahara, M. and Hirose, S. (1998) Elastodynamic inversion of 3D cavity from backscattering data, *Inverse Problems in Engineering Mechanics*, 163-170.
- [4] Kitahara, M., Nakahata, K. and Hirose, S. (2002) Elastodynamic inversion for shape reconstruction and type classification of flaws, *Wave Motion*, **36-4**, 443-455.
- [5] Nakahata, K., Saitoh, T. and Hirose, S. (2006) 3-D flaw imaging by inverse scattering analysis using ultrasonic array transducer, *Review of Progress in Quantitative Nondestructive Evaluation*, **34**, 717-724.
- [6] Saitoh, T., Hirose, S., Fukui, T. and Ishida, T. (2007) Development of a time-domain fast multipole BEM based on the operational quadrature method in a wave propagation problem, *Advances in Boundary Element Techniques VIII*, V. Minutolo and MH Aliabadi, eds., 355-360.
- [7] Saitoh, T. and Hirose, S. (2010) Parallelized fast multipole BEM based on the convolution quadrature method for 3-D wave propagation problems in time-domain, IOP Conf. Series: *Materials Science and Engineering* **10**, 012242.
- [8] Furukawa, A., Saitoh, T. and Hirose, S. (2014) Convolution quadrature time-domain boundary element method for 2-D and 3-D elastodynamic analyses in general anisotropic elastic solids, *Engineering Analysis with Boundary Elements*, **39**, 64-74.
- [9] Maruyama, T., Saitoh, T., Bui, T. Q. and Hirose, S. (2016) Transient elastic wave analysis of 3-D large-scale cavities by fast multipole BEM using implicit Runge-Kutta convolution quadrature, *Computer Methods in Applied Mechanics and Engineering*, **303**, 231-259.
- [10] Sedow, A. and Schmerr, L.W. (1986) The time domain elastodynamic Kirchhoff approximation for cracks, the inverse problem, *Wave Motion*, **8**, 15-26.
- [11] Auld, B. A. (1990) Acoustic fields and waves in solids, Vol.1,2, R. E. Krieger.
- [12] Wang, C.-Y. and Achenbach, J. D. (1994) Elastodynamic fundamental solutions for anisotropic solids, *Geophys. J. Int.*, **118**, 384-392.
- [13] Ricker, N. (1953) The form and laws of propagation of seismic wavelets, *Geophys.*, **18**, 10-40.

Tsunami Run-Up Simulation Using Particle Method and its Visualization with Unity

†Takahiro Saitoh¹, Gohki Noguchi¹, and Takumi Inoue¹

¹Department of Civil and Environmental Engineering, Gunma University, Japan

*Presenting author: t-saitoh@gunma-u.ac.jp

†Corresponding author: t-saitoh@gunma-u.ac.jp

Abstract

This paper presents a mobile-processed virtual reality (VR) tsunami simulator, which could help people evacuate in the event of a powerful tsunami in the future. The explicit moving particle simulation (E-MPS) method is developed for a tsunami simulation, and the results are demonstrated on the proposed VR with the help of Unity, an open-source software product for VR visualization using a head-mounted display. The experience of the developed VR system might be also helpful for training related to tsunami disaster preparation.

Keywords: Tsunami simulation, Virtual reality, Unity, Particle method

Introduction

The Great East Japan Earthquake, which occurred on March 11, 2011, inflicted serious damage on civil structures in northeastern Japan. In particular, the tsunami generated by the earthquake caused heavy damage not only to civil structures but also to the Japanese economy. Since this heavy disaster, several measures have been taken to prevent tsunami disasters, such as the one that followed the Nankai Trough Earthquake, which may occur in the future. Evacuation training and drawing-up hazard maps for tsunamis are important measures. However, residents are likely to forget what they have learned about tsunami disaster prevention. Therefore, it is necessary to take additional measures to raise citizens' awareness of tsunami disaster prevention.

Recently, virtual reality (VR) technology has been developed and is attracting attention in many engineering fields. Virtual reality is a computer-generated scenario that simulates an experience in a VR space. The advantage of VR is its ability to take us to places we have never been before and experience things that we could not otherwise experience with any sense of realism. In general, we cannot experience a tsunami many times in our life. Consequently, we sometimes fail to remember past catastrophic natural disasters, such as earthquakes and tsunamis. Therefore, the application of VR to tsunami disaster preparation may help residents create memories of catastrophic tsunami disasters, such as the incident that occurred on March 11, 2011. The key to increasing the effect of a VR system for tsunami disaster preparation is to reproduce realistic tsunami behaviors in a VR space. Innovative numerical simulation tools are helpful in reproducing such realistic tsunami disasters in a VR space. Motohashi et al. [1] used an open-source software program to estimate the hydraulic force exerted on a bridge by a tsunami. Flouri et al. [2] implemented a simulation of earthquake-generated tsunamis using a finite difference computational model. Wei et al. [3] calculated the hydraulic force of a tsunami using smoothed particle hydrodynamics (SPH), which is one of the particle methods. That particle method is known as a powerful numerical technique and does not require computational meshes, such as the ones used in the finite element method (FEM) [4] and boundary element method (BEM) [5][6][7]. Moreover, a particle method can easily handle the large deformation

of continuum bodies. As mentioned here, the particle method [8] has many advantages for fluid analysis.

Therefore, in this research, the explicit-moving particle simulation (E-MPS) [9][10], which is one of the particle methods, is developed and integrated into a VR system to create a tsunami disaster experience. A smartphone is used for high portability and usability of the developed VR system. Unity, which is an open-source software product for VR experience development, is utilized to visualize the numerical results obtained by the E-MPS and construct the VR space of the developed system. The following text explains the E-MPS formulation. Then, the VR visualization obtained with Unity is discussed. Finally, the developed VR system for tsunami disaster preparation is demonstrated, and some comments on our future research works are provided.

E-MPS formulation

In this research, the continuity and Navier Stokes equations at time t are solved using the E-MPS. The continuity and Navier Stokes equations are defined as

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{u} = 0 \quad (1)$$

$$\frac{D\mathbf{u}}{Dt} = -\frac{1}{\rho} \nabla P + \nu \nabla^2 \mathbf{u} + \mathbf{g} \quad (2)$$

where ρ is the density, \mathbf{u} is the fluid velocity, and P is the fluid pressure. Moreover, ν is the kinematic viscosity and \mathbf{g} is the acceleration of gravity. The continuity and Navier Stokes equations defined in Eqs. (1) and (2), respectively, are calculated as

$$\langle \nabla P \rangle_i = \frac{d}{n_{grad}^0} \sum_{j \neq i} \left[\frac{(P_j + P_i)(\mathbf{r}_j - \mathbf{r}_i)}{|\mathbf{r}_j - \mathbf{r}_i|^2} \omega_{grad}(|\mathbf{r}_j - \mathbf{r}_i|) \right] \quad (3)$$

$$\langle \nabla^2 \mathbf{u} \rangle_i = \frac{2d}{\lambda^0 n_0} \sum_{j \neq i} [(\mathbf{u}_j - \mathbf{u}_i) \omega(|\mathbf{r}_j - \mathbf{r}_i|)] \quad (4)$$

where d and λ^0 represent the number of space dimensions and correction parameters used in the particle method, respectively. In addition, \mathbf{r} is the position vector of the particle. Note that the subscript i (or j) indicates the parameter of the i (or j)-th particle. Meanwhile, $\omega(|\mathbf{r}|)$ and n_0 indicate the weight function and initial value of the particle number density, respectively. The subscript “*grad*” represents the calculation term for gradient. The weight functions $\omega_{grad}(r)$ and $\omega(r)$ are defined as follows:

$$\omega_{grad}(r) = \begin{cases} \frac{r_e}{r} - \frac{r}{r_e} & (r < r_e) \\ 0 & (r \geq r_e) \end{cases}, \quad (5)$$

$$\omega(r) = \begin{cases} \frac{r_e}{r} + \frac{r}{r_e} - 2 & (r < r_e) \\ 0 & (r \geq r_e) \end{cases} \quad (6)$$

In Eqs. (5) and (6), r is the distance between particles and r_e is the influence radius of a particle. The parameters n_{grad}^0 and n^0 are defined as follows:

$$n_{grad}^0 = \sum_{j \neq i} \omega_{grad}(|\mathbf{r}_j - \mathbf{r}_i|) \quad (7)$$

$$n^0 = \sum_{j \neq i} \omega(|\mathbf{r}_j - \mathbf{r}_i|). \quad (8)$$

The Navier Stokes equation (2) can be calculated using Eqs. (3) and (4) [8]. However, in this study, the left-hand side of Eq. (2) can be calculated by the explicit Euler method in the E-MPS algorithm using the intermediate velocity \mathbf{u}_i^* as

$$\frac{D\mathbf{u}}{Dt} = \frac{\mathbf{u}_i^{k+1} - \mathbf{u}_i^k}{\Delta t} = \frac{\mathbf{u}_i^* - \mathbf{u}_i^k}{\Delta t} + \frac{\mathbf{u}_i^{k+1} - \mathbf{u}_i^*}{\Delta t} \quad (9)$$

where \mathbf{u}_i^k denotes the particle velocity of particle i at the k -th time step. In addition, Δt represents the time increment. The superscript $*$ denotes the physical quantity at the intermediate time step. The intermediate particle velocity \mathbf{u}_i^* can be calculated as follows:

$$\mathbf{u}_i^* = \mathbf{u}_i^k + (\nu \langle \nabla^2 \mathbf{u} \rangle_i^k + \mathbf{g}) \Delta t. \quad (10)$$

Moreover, the particle position \mathbf{r}_i^* at the intermediate step can be obtained as follows:

$$\mathbf{r}_i^* = \mathbf{r}_i^k + \mathbf{u}_i^* \Delta t. \quad (11)$$

The particle velocity \mathbf{u}_i^{k+1} and position \mathbf{r}_i^{k+1} can be calculated using Eq. (9) as

$$\mathbf{u}_i^{k+1} = \mathbf{u}_i^* - \frac{\Delta t}{\rho_i^0} \langle \nabla P \rangle_i^{k+1} \quad (12)$$

$$\mathbf{r}_i^{k+1} = \mathbf{r}_i^* + (\mathbf{u}_i^{k+1} - \mathbf{u}_i^*) \Delta t \quad (13)$$

where ρ^0 is the initial density of the fluid. In the conventional MPS, Poisson's equation for the pressure obtained by Eq. (12) can be solved implicitly to obtain the pressure P^{k+1} . However, in the E-MPS, the pressure P^{k+1} can be evaluated as a function of density, as follows:

$$P^{k+1} = \begin{cases} c^2(\rho^* - \rho^0) & (\rho^* > \rho^0) \\ 0 & (\rho^* \leq \rho^0) \end{cases} \quad (14)$$

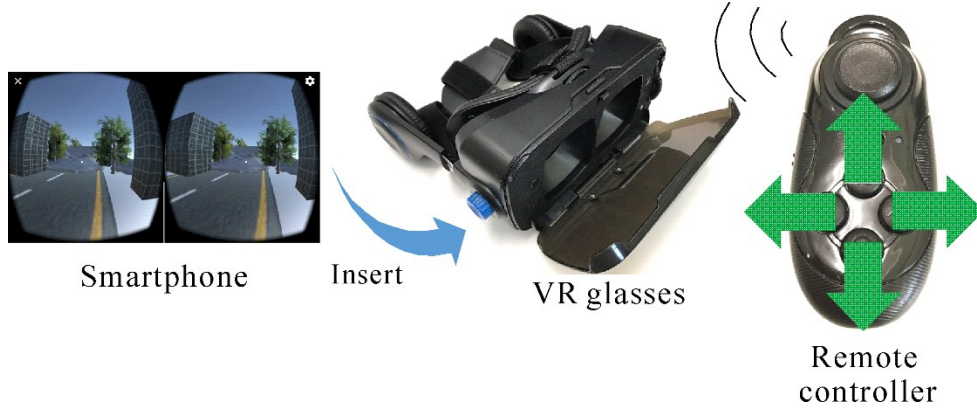


Figure 1. VR system using smartphone and remote controller.

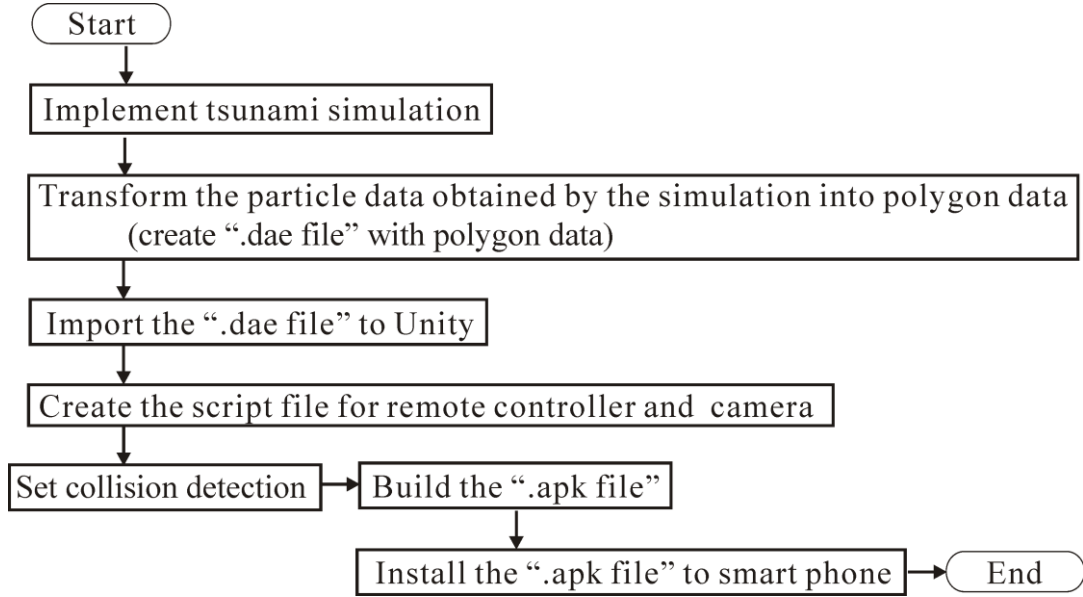


Figure 2. Flowchart of creating the ".apk file" required for the proposed VR system.

where c is the speed of sound. The density ρ^* in the intermediate step can be calculated by assuming that ρ^* is proportional to the sum of the weight function ω , as follows:

$$\rho^* = \frac{\rho^0}{n^0} \sum_{j \neq i} w(|\mathbf{r}_j^* - \mathbf{r}_i^*|). \quad (15)$$

In Eq. (14), the speed of sound c is given by $c = u_{max}/0.2$, where u_{max} is a predicted value of the maximum fluid velocity.

VR visualization using Unity

The tsunami run-up behavior can be calculated using the E-MPS, as explained in the previous section. Therefore, in this section, we briefly describe how to integrate the numerical results obtained by the E-MPS into a VR space. Although there are several kinds of VR systems, a mobile VR head-mounted display (HMD), which uses a smartphone for the display, is considered in this research, as shown in Fig.1. A VR system with a large-scale screen for

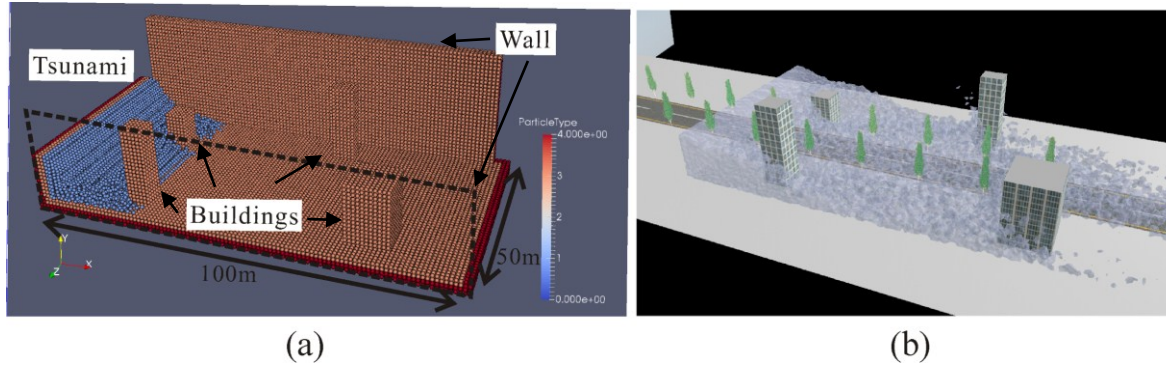


Figure 3. Visualization results for tsunami simulation. An example of (a) particle data used in this analysis and (b) visualization of the numerical result with the aid of Maya.

tsunami simulation was developed by Kawabe et al. [11]. In addition, Tanaka et al. [12] proposed a tsunami virtual reality system using a head-mounted display connected to a high-performance notebook PC with a graphics-processing unit (GPU). However, their proposed system requires expensive high-performance equipment. Such systems are not suitable for allowing many people to experience a VR system in different places at the same time. For that reason, a VR system using a smartphone, which can easily and inexpensively provide a virtual tsunami experience, was constructed in this research.

Hereafter, a brief description of the VR system construction procedure is provided. The procedure used Maya [13] and Unity [14], which are the 3D computer software product with powerful modeling and widespread use as free VR development platform, respectively. In this research, first, a tsunami simulation using the E-MPS was implemented as the scenario that the VR user experiences. In general, the numerical results are output as a binary or text file. However, these file extensions are not suitable for Unity. Therefore, the output files are transformed by Maya into files with the extension “.dae”, which are compatible with Unity. In this process, the particle data, which show the tsunami behavior, are obtained by the E-MPS and are transformed into polygon data. At that time, some models used in the E-MPS analysis, such as buildings, grounds, and timbers are rendered for a more realistic visualization. The setting of the camera, movement operation using the remote controller, and collision detection for walls, buildings, and trees, are also performed using Unity. This setting is implemented through the C#-like script programming in Unity. Finally, a file with the Android application extension, “.apk” is built and copied to a smartphone that is inserted into the VR glasses, as shown in Fig.1. The flowchart of these operation procedures can be seen in Fig.2. Obviously, this developed VR system is economical and does not need a high-performance PC and large screen, except for the workstation used in the particle simulation. Therefore, this system is suitable for disaster preparation training, such as in a situation in which the instructor has a large number of trainees for tsunami evacuation drills.

VR demonstration results

The VR demonstration results can be seen in this section. The E-MPS was applied to the tsunami simulation, whose model is illustrated in Fig.3(a). The scenario in this simulation is that of a virtual urban area with some buildings, which is hit by a tsunami. In general, the numerical models for particle methods are represented as an assembly of particles. Therefore, the fluid, riverbed, and buildings are represented by particles, as shown in the figure. The virtual walls are set with reflected boundary conditions for front and back, as shown in the same figure,

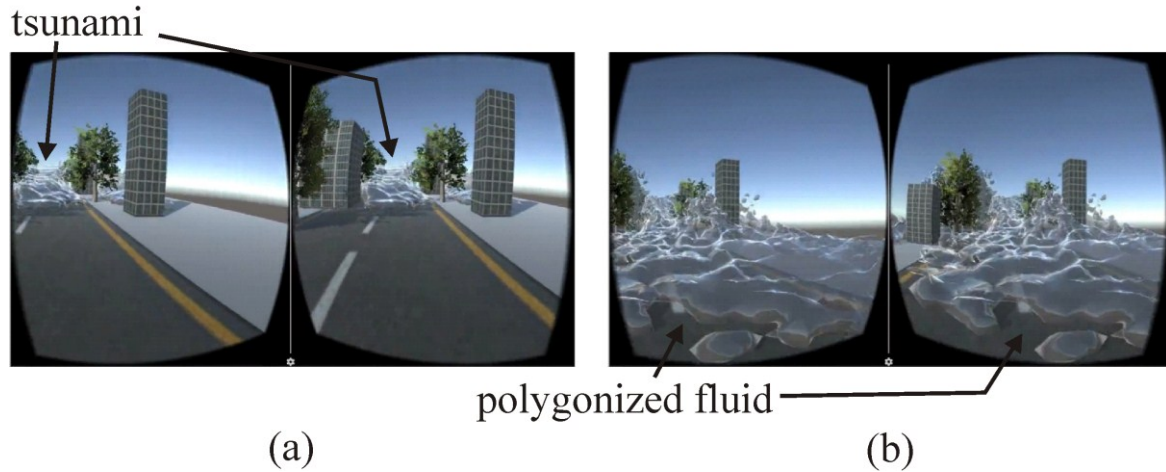


Figure 4. VR demonstration results at selected time steps:(a) a VR user can see the tsunami behind the building; (b) a VR user is swallowed by the tsunami.

to prevent the fluid from spreading widely in the VR space. The walls and buildings are treated as wall particles. In this analysis, the density of the fluid ρ (or ρ^0), speed of sound c , gravity \mathbf{g} , and time step size Δt are set as ρ (or ρ^0) = 1000kg/m³, c = 100.0m/s, \mathbf{g} = (0.0, -9.8m/s², 0.0), and Δt = 1.0×10^{-2} s, respectively. In addition, the kinematic viscosity ν and number of the space dimension d are given by ν = 1.301×10^{-2} m²/s and d = 3, respectively. The particle data obtained by the E-MPS are transformed into polygon data and visualized by Maya, as explained in the previous section. Figure 3(b) shows an example of the visualized E-MPS numerical result obtained using Maya. Note that the trees in that figure are not considered in this numerical analysis model. In fact, the trees are simply located in the VR space to enhance the realism of the simulation in the steps of the Maya-visualization. The realistic visualization of this simulation was achieved with the aid of Maya. All of the static images for each time step, as shown in Fig.3(b), are prepared for the VR demonstration with a smartphone. These static images are gathered in Unity. Then, the Android application with the extension “.apk” is built and the resulting file is copied to the smartphone, which is inserted into the VR glasses.

Figure 4 shows an example of the VR demonstration results, which were obtained using Unity, according to the flowchart in Fig.2. Figure 4(a) and (b) show the stereo rendering images for the left and right eyes, respectively, through the VR glasses. In Fig.4(a), a VR user sees the tsunami approaching from behind the building. In Fig.4(b), the VR user is swallowed by the tsunami. From Fig.4, someone using the developed VR system using a smartphone can feel the velocity and height of the tsunami approaching, something that cannot be readily experienced in the real world. For example, a big tsunami such as the one observed after the Great East Japan Earthquake of March 11, 2011 occurs only once every several hundred years. Therefore, it was concluded that the developed VR system for experiencing a tsunami may be helpful as a method for tsunami disaster preparation drills.

Conclusion

In this study, a tsunami simulation was implemented using the E-MPS, which is one of the particle methods. The particle data obtained from the E-MPS were transformed into polygon data using Maya, and realistic static images were created. These static images were gathered, and the VR application was built using Unity. The VR demonstration results with a smartphone device installed (containing the developed VR application) were shown to prove that it worked

properly. In the future, this developed VR system will be used for tsunami disaster preparation drills.

Acknowledgements

This work was supported by grants from the Takahashi Industrial and Economic Research Foundation, and the foundation for the Fusion Of Science and Technology.

References

- [1] Motohashi, H., Sugatsuke, K., Nonaka, T., Kawasaki, K. and Harada, T. (2013) Tsunami damage simulation of KOIZUMI bridge, *Journal of JSCE*, B2 **69**, 831-835 (in Japanese).
- [2] Flouri, E. T., Kalligeris, N., Alexandrakis, G., Kampanis, N. A., and Synolakis, C. E. (2013) Application of a finite difference computational model to the simulation of earthquake generated tsunamis, *Applied Numerical Mathematics* **67**, 111-125.
- [3] Wei, Z. and Dalrymple, R. A. (2016) Numerical study on mitigating tsunami force on bridges by an SPH model, *Journal of ocean engineering and marine energy* **2**(3), 365-380.
- [4] Zienkiewicz, O. C., Taylor, R. L., and Nithiarasu, P. (2013) The finite element method for fluid dynamics, Seventh Edition, Butterworth-Heinemann.
- [5] Saitoh, T., Hirose, S., Fukui, T. and Ishida, T. (2007) Development of a time-domain fast multipole BEM based on the operational quadrature method in a wave propagation problem, *Advances in Boundary element technique*, VIII, 355-360.
- [6] Saitoh, T. and Hirose, S. (2010) Parallelized fast multipole BEM based on the convolution quadrature method for 3-D wave propagation problems in time-domain, IOP Conf. Series: *Materials Science and Engineering* **10**(1), 012242.
- [7] Brebbia, C. A., Telles, J. C. F. and Wrobel, L. C. (2012) Boundary element techniques: Theory and applications in engineering, Springer.
- [8] Koshizuka, S. and Oka, Y. (1996) Moving-particle semi-implicit method for fragmentation of incompressible fluid, *Nucl. Sci. Wng* **123**, 421-434.
- [9] Shakibaeinia, A. and Jin, Y. C. (2010) A weakly compressible MPS method for modeling of open-boundary free-surface flow, *Int. J. Numer. Methods Fluids* **63** (10), 1208-1232.
- [10] Oochi, M., Koshizuka, S. and Sakai, M. (2010) Explicit MPS algorithm for free surface flow analysis, *Proc. conf. Comput. Eng. Sci.* **15**(2), 589-590.
- [11] Kawabe, T., Kashiya, K., Miyachi, H., Iwatsuka, Y., Furumaki, D. and Nishihata, T. (2014) Development of a simulation based experience system for tsunami using sonification technology, *Journal of JSCE*, F3 **70**(2), 235-242.
- [12] Tanaka, N., Saito, F. and Kitajima, M. (2015) Development of tsunami virtual reality system using head-mounted display, *CFD Symposium of Japan society of Fluid Mechanics*, A05-3.
- [13] Riddell, D. and Britt, A. (2001) Maya: visual quickstart guide, Pearson Education.
- [14] Linowes, J. (2015) Unity virtual reality projects: explore the world of virtual reality by building immersive and fun VR projects using Unity 3D, Packt Publishing.

A class of novel tetrahedron elements with curved surfaces for three-dimensional solid mechanics problems with curved boundaries

C.Q. Wang¹, †J.H. Yue¹, Ming Li²

¹Taiyuan University of Technology, Taiyuan, China

²Taiyuan University of Technology, Taiyuan, China, liming04@gmail.com

†Corresponding author: woyuejunhong@163.com

Abstract

Linear tetrahedral elements with four nodes (Te4) are currently the simplest and most widely used in finite element (FE) developed for solving 3D mechanics problems. However, the standard Te4 elements cannot be used to simulate accurately the 3D problems with curved boundaries, because of the flat surfaces of the standard Te4 elements. In this paper, we develop a set of new elements having curved surfaces to simulate the curved boundaries, by adding nodes to the standard Te4 elements. These novel elements include five-noded, six-noded, and seven-noded tetrahedron elements (Te5, Te6, and Te7). Based on the Te4 FE mesh, a hybrid mesh can be conveniently built for 3D problems with curved boundaries, in which the standard Te4 elements are used for the interior, and Te5, Te6, and Te7 elements are used for the curved boundaries. Compared with the standard FEM with Te4 elements, our mixing mesh can significantly improve the accuracy of the solution at the curved boundaries. Several solid mechanics problems are studied using hybrid meshes to validate the effectiveness of the present new elements.

Keywords: finite element method; curved boundaries; five-noded, six-noded, seven-noded tetrahedron element

Introduction

Common three-dimensional element of FEM is linear tetrahedral element with four nodes (Te4), which can automatically generate for complex geometries [1]. Due to its high efficiency, robustness and adaptability for complex geometries, the Te4 element is the most commonly used for general solid mechanics problems. However, the accuracy of FEMs using Te4 elements is poor in terms of stress, especially at the curved boundaries. Tetrahedral elements with ten nodes (Te10) [2], wedge elements with six nodes (W6) and hexahedron elements with eight nodes (H8) are proposed for this problem to obtain higher accuracy, but the computational cost is too large. Considering the above characteristics of the Te4, the Te10 and the higher order element, we propose a hybrid class of multi-node tetrahedral elements.

For the 3D problem domain with curved surfaces, some of the edges of tetrahedron boundary elements locate on the curved boundaries. If the edge is on the curved boundaries, we use a curved edge instead of the straight edge used in the standard Te4 element. Then we add a new node in the middle point of the curved edge to accurately simulate the changing trend. We analyze the type of the boundary elements, and put forward three kinds of new tetrahedron elements which are five-noded, six-noded and seven-noded tetrahedron (Te5, Te6, and Te7) elements.

The shape functions of the Te5, Te6 and Te7 elements for FEM

Then we construct the shape functions for the Te5, Te6 and Te7 elements. Figure 1(a) is a Te5 element that one additional node was added on the middle node of the curved edge. By using shape functions of the four-noded triangular (Tr4) element which can be found in [3], we can construct the shape function of a standard five-noded tetrahedron element in the natural system, which can be seen in Figure 1(b).

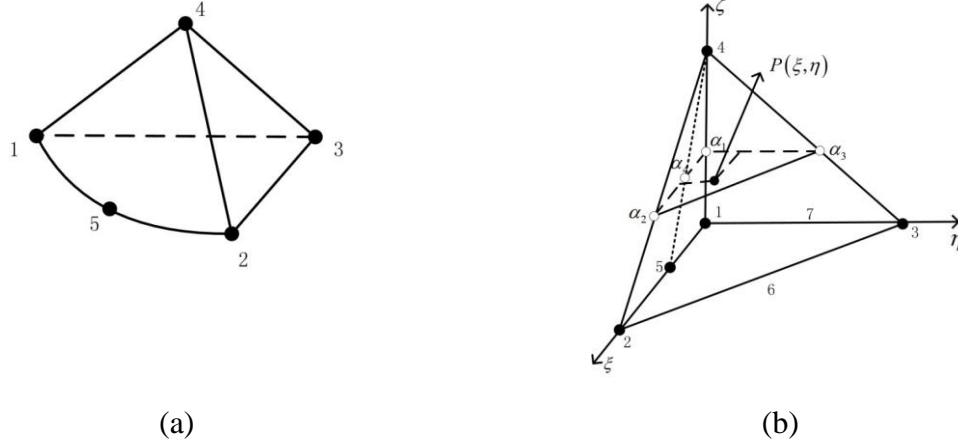


Figure 1. (a) The five-noded tetrahedron element; (b) the five-noded tetrahedron element in natural system.

For any triangular $\Delta_{\alpha_1\alpha_2\alpha_3}$ paralleled to the triangular Δ_{123} , the displacement can be approximated using

$$u = \sum_{i=1}^4 \phi_{\alpha_i}^{Tr4} u_{\alpha_i} \quad (1)$$

where the shape function of the Tr4, $\Delta_{\alpha_1\alpha_2\alpha_3}$, can be written as

$$\left\{ \begin{aligned} \phi_{\alpha_1}^{Tr4} &= \frac{2\left(\frac{\xi}{1-\zeta}\right)^2}{1-\left(\frac{\eta}{1-\zeta}\right)} - 3\frac{\xi}{1-\zeta} + \left(1 - \frac{\eta}{1-\zeta}\right) = \frac{2\xi^2}{(1-\zeta-\eta)(1-\zeta)} + \frac{1-3\xi-\zeta-\eta}{1-\zeta} \\ \phi_{\alpha_2}^{Tr4} &= \frac{2\left(\frac{\xi}{1-\zeta}\right)^2}{1-\left(\frac{\eta}{1-\zeta}\right)} - \frac{\xi}{1-\zeta} = \frac{2\xi^2}{(1-\zeta-\eta)(1-\zeta)} - \frac{\xi}{1-\zeta} \\ \phi_{\alpha_3}^{Tr4} &= \frac{\eta}{1-\zeta} \\ \phi_{\alpha_4}^{Tr4} &= 4\frac{\xi}{1-\zeta} - \frac{4\left(\frac{\xi}{1-\zeta}\right)^2}{1-\frac{\eta}{1-\zeta}} = \frac{4\xi}{1-\zeta} - \frac{4\xi^2}{(1-\zeta-\eta)(1-\zeta)} \end{aligned} \right. \quad (2)$$

Invoking the simple fact that

$$\frac{l_{4-\alpha_1}}{l_{4-1}} = \frac{l_{4-\alpha_2}}{l_{4-2}} = \frac{l_{4-\alpha_3}}{l_{4-3}} = \frac{l_{4-\alpha_5}}{l_{4-5}} = 1 - \zeta, \quad (3)$$

where l_{i-j} is the distance between two points i and j . So we have the relationships as listed:

$$\begin{aligned}
 u_{\alpha_i} &= \zeta u_4 + (1 - \zeta) u_i; \quad i = 1, 2, 3 \\
 u_{\alpha_i} &= \zeta u_4 + (1 - \zeta) u_{i+1}; \quad i = 4
 \end{aligned} \tag{4}$$

Therefore, the displacement in the standard five-noded tetrahedron element can be evaluated using the following formulation

$$\begin{aligned}
 u &= \sum_{i=1}^4 \phi_{\alpha_i}^{Tr4} u_{\alpha_i} \\
 &= (\zeta u_4 + (1 - \zeta) u_1) \phi_{\alpha_1}^{Tr4} + (\zeta u_4 + (1 - \zeta) u_2) \phi_{\alpha_2}^{Tr4} \\
 &\quad + (\zeta u_4 + (1 - \zeta) u_3) \phi_{\alpha_3}^{Tr4} + (\zeta u_4 + (1 - \zeta) u_5) \phi_{\alpha_4}^{Tr4} \\
 &= (1 - \zeta) \phi_{\alpha_1}^{Tr4} u_1 + (1 - \zeta) \phi_{\alpha_2}^{Tr4} u_2 + (1 - \zeta) \phi_{\alpha_3}^{Tr4} u_3 + \zeta u_4 + (1 - \zeta) \phi_{\alpha_4}^{Tr4} u_5 \\
 &= N_1 u_1 + N_2 u_2 + N_3 u_3 + N_4 u_4 + N_5 u_5
 \end{aligned} \tag{5}$$

Substitute Eq.(2) into the above equation, the shape functions N_i ($i = 1, 2, 3, 4, 5$) of the standard Te5 element can be formulated as

$$\begin{cases}
 N_1 = \frac{2\xi^2}{(1 - \zeta - \eta)} + 1 - 3\xi - \zeta - \eta \\
 N_2 = \frac{2\xi^2}{(1 - \zeta - \eta)} - \xi \\
 N_3 = \eta \\
 N_4 = \zeta \\
 N_5 = 4\xi - \frac{4\xi^2}{(1 - \zeta - \eta)}
 \end{cases} \tag{6}$$

where the parameters $\xi \in [0, 1]$, $\eta \in [0, 1]$, $\zeta \in [0, 1]$.

Figure 2(a) is a Te6 element that two additional nodes were added on the middle node of each curved edge.

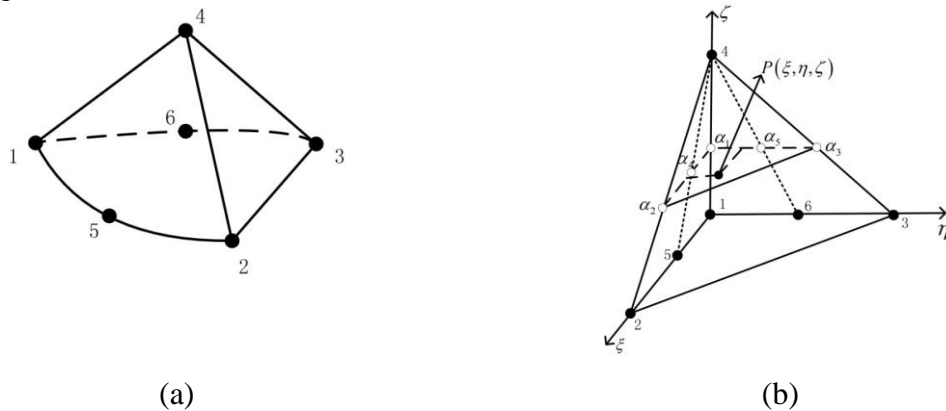


Figure 2. (a) The six-node tetrahedron element; (b) the standard six-node tetrahedron element in the natural system.

Similarly, the shape functions N_i ($i = 1, 2, 3, 4, 5, 6$) of the standard Te6 element, which can be seen in Figure 2(b), can be formulated as

$$\left\{ \begin{array}{l} N_1 = (1-\zeta)\phi_{\alpha_1}^{Tr5} = \frac{(\xi + \eta + \zeta - 1)(2\xi + 2\eta + \zeta - 1)}{(1-\zeta)} \\ N_2 = (1-\zeta)\phi_{\alpha_2}^{Tr5} = \frac{\xi(2\xi + 2\eta + \zeta - 1)}{(1-\zeta)} \\ N_3 = (1-\zeta)\phi_{\alpha_3}^{Tr5} = \frac{\eta(2\xi + 2\eta + \zeta - 1)}{(1-\zeta)} \\ N_4 = \zeta \\ N_5 = (1-\zeta)\phi_{\alpha_4}^{Tr5} = \frac{4\xi(1-\xi-\eta-\zeta)}{(1-\zeta)} \\ N_6 = (1-\zeta)\phi_{\alpha_5}^{Tr5} = \frac{4\eta(1-\xi-\eta-\zeta)}{(1-\zeta)} \end{array} \right. \quad (7)$$

where the parameters $\xi \in [0,1], \eta \in [0,1], \zeta \in [0,1]$.

Figure 3(a) is a Te7 element that three additional nodes were added on the middle node of each curved edge.

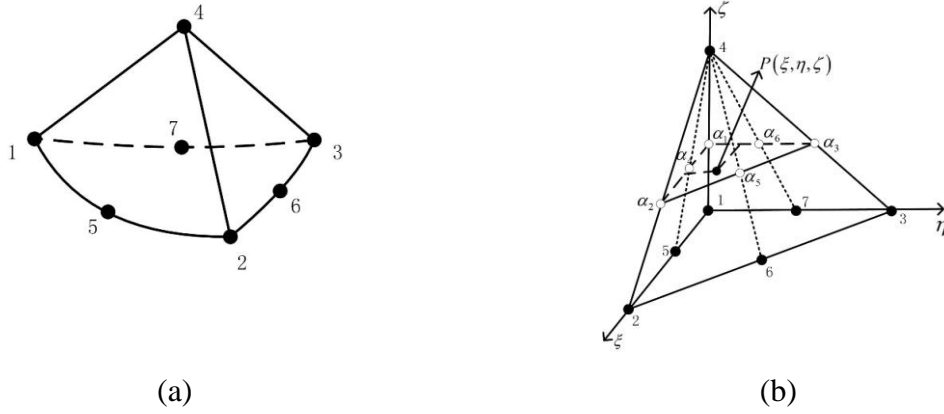


Figure 3. (a) The seven-node tetrahedron element; (b) the standard seven-node tetrahedron element in the natural system.

Similarly, the shape functions N_i ($i = 1, 2, 3, 4, 5, 6, 7$) of the Te7 element, which can be seen in Figure 3(b), can be formulated as

$$\left\{ \begin{array}{l} N_1 = (1-\zeta)(2L_1-1)L_1 = (1-\xi-\eta-\zeta)\left(\frac{2(1-\xi-\eta-\zeta)}{1-\zeta}-1\right) \\ N_2 = (1-\zeta)(2L_2-1)L_2 = \xi\left(\frac{2\xi}{1-\zeta}-1\right) \\ N_3 = (1-\zeta)(2L_3-1)L_3 = \eta\left(\frac{2\eta}{1-\zeta}-1\right) \\ N_4 = \zeta \\ N_5 = 4(1-\zeta)L_1L_2 = \frac{4(1-\xi-\eta-\zeta)\xi}{1-\zeta} \\ N_6 = 4(1-\zeta)L_2L_3 = \frac{4\xi\eta}{1-\zeta} \\ N_7 = 4(1-\zeta)L_3L_1 = \frac{4(1-\xi-\eta-\zeta)\eta}{1-\zeta} \end{array} \right. \quad (8)$$

where the parameters $\xi \in [0,1], \eta \in [0,1], \zeta \in [0,1]$.

Numerical simulation

The domain of the hollow sphere is defined as $\Omega = B(0, 2)/B(0, 1.0)$, which the origin $O(0, 0, 0)$, inner radius $a = 1.0m$, and outer radius $b = 2.0m$. The hollow sphere is subjected to an internal pressure $P=1 \text{ N/m}^2$ on the inner spherical surface. Because of the symmetric characteristics of the problem, only one-eighth of hollow sphere needs to be modeled as shown in Figure 1, and symmetric conditions are imposed on the symmetric planes.

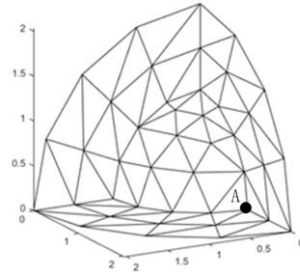


Figure 1. one-eighth of hollow sphere discretized using Te4 elements

Table1. Relative errors in displacement component u of the added nodes in the curved edges for the inner surface

Mesh	62 nodes	371 nodes	770 nodes	1482 nodes
FEM-Te4	0.2429	0.0532	0.0331	0.0140
FEM-HM	0.1796	0.0361	0.0246	0.0086

We use FEM-Te4 to represent the finite element method using Te4 elements and FEM-HM to represent the finite element method using a hybrid mesh with Te4, Te5, Te6 and Te7 elements. Table1 gives the relative errors in displacement component u of the added nodes in the curved edges for the inner surfaces using different elements and mesh sizes. The results show that the hybrid mesh with Te4, Te5, Te6 and Te7 elements can improve the accuracy of the displacement result on the curved boundaries, compared to the mesh with Te4 elements.

Figure 2 shows relative errors in radial stress σ_r of Point A (marked in Figure 1) against mesh sizes using different elements, which obtains the maximum radial stress easily observed in the analytical solution. It is clearly seen that the hybrid mesh with Te4, Te5, Te6 and Te7 elements stands out in the radial stress, compared the mesh with Te4 elements.

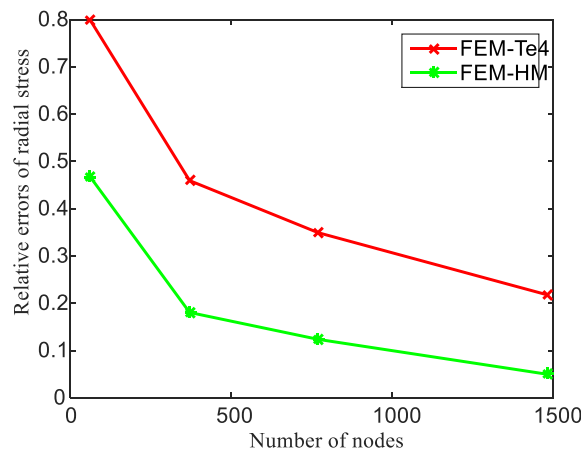


Figure 2. Relative errors of the radial stress of Point A against mesh sizes using different elements for 3D Lamé problem.

Conclusions

In this paper, we present a novel hybrid mesh using Te5, Te6, and Te7 elements to accurately approximate the curved boundaries of problem domains. The hybrid mesh not only remains the advantages of the linear tetrahedral element, but also greatly improves the accuracy of the stress solution. Based on the shape functions of the standard Tr4 element, the standard Tr5 element and the standard Tr7 element, we formulate the shape functions for Te5, Te6, and Te7 elements separately. Through intensive numerical examples, it is concluded that our novel hybrid mesh with the multi-node tetrahedral element can simulate the curved boundaries efficiently and accurately.

References

- [1] Borouchaki H, George PL (1996) Optimal Delaunay point insertion. *Int J Numer Methods Eng* 39:3407–3437.
- [2] M. Jabareen, E. Hanukah, and M. Rubin, "A ten node tetrahedral Cosserat Point Element (CPE) for nonlinear isotropic elastic materials," *Computational Mechanics*, vol. 52, pp. 257-285, 2013.
- [3] J.H. Yue, G.R. Liu, R.P. Niu, M. Li, A Four-Noded Triangular (Tr4) Element for Solid Mechanics Problem with Curved Boundaries. *International Journal of Computational Methods*, Vol.15, No. 1(2018) 1844003.

Multiscale Damage Modelling of Sustainable Composites

Shyam M. Panamoottil¹, †* Raj Das², and Krishnan Jayaraman¹

¹Department of Mechanical Engineering, University of Auckland, New Zealand

²School of Engineering, RMIT University, Australia

*Presenting author: raj.das@rmit.edu.au

†Corresponding author: raj.das@rmit.edu.au

Abstract

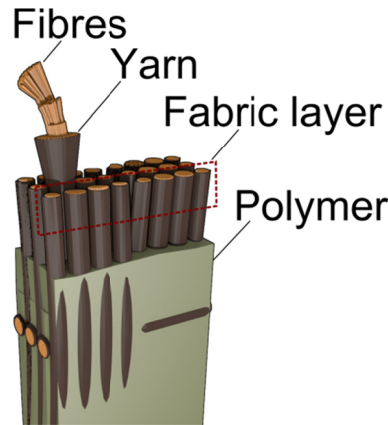
A computational homogenization technique has been implemented using the ABAQUS finite element solver to analyse the behaviour of flax fibre composites subjected to three-point bending. Macroscale models of bending specimens were coupled with numerically estimated damage rules for two systems each of flax/polypropylene and flax/epoxy. The results obtained for the failure strength were between 7.5-11.2% lower than the test average values. Validation studies were also performed, using geometry and material parameters for a glass/epoxy composite, and the predicted failure strength was only 12.6% lower than the experimental average for the glass/epoxy three-point bending specimens.

Keywords: Multiscale, Homogenization, Damage, Sustainability, Flax, Composites.

Introduction

Composites can be manufactured in a range of different configurations, using short or long fibres, which are present in either random or oriented manner in the matrix material. For higher performance applications, it is common to utilise long fibres, which can be obtained in fabric form. Shown in Figure 1 are the various stages of producing a composite from fabric. Yarns are composed of several fibres twisted together (Figure 1), which are then woven to produce fabrics (Figure 1). Layers of fabric are then stacked together and infused with matrix/resin to form a composite. As shown in the illustration (Figure 1), the multiscale structure of composites materials results in the response to loading being contributed to by components at different length scales.





(c)

Figure 1. Illustration of a composite made from yarn fabric - (a) yarn composed of twisted fibres, (b) fabric from yarn, (c) stack of fabric layers and (d) fabric stack infused with resin

The heterogeneous and multiscale nature of composite materials causes them to demonstrate failure behaviour that is quite different from that of metals or other materials which might have been traditionally used in certain applications. The various failure processes observed could include fibre fracture and interface failure at the fibre-interface scale [1] (microscale), matrix cracking and delamination at the yarn-lamina scale (mesoscale) [2]. Local damages/cracks interact with each other, sometimes synergistically, and finally may lead to catastrophic failure [3] at the composite level. Macroscale damage may present as longitudinal splitting, shear crippling or complete rupture, among other modes. Local phenomena thus indirectly dictate the final failure properties of composites.

Due to the complexity involved in composite failure, coupled multiscale modelling techniques are necessary to capture the failure mechanisms at the different geometric scales [4, 5]. Such techniques typically couple local analysis of RVE models with global or macro analysis [6-8]. Most of these techniques can be classified as superposition techniques [9-15] mathematical homogenization techniques [16-21], domain decomposition techniques [22-26] and multiscale computational techniques [8, 27-30]. Among the various techniques explored, the approach taken by Smit et al. [8, 29, 31] is to be noted for being *non-intrusive*, meaning that it can be easily implemented with commercial finite element (FE) codes.

Natural fibre composites (NFCs) are also increasingly being adopted [32, 33] in the automobile and construction industries worldwide. The NATEX (Aligned Natural Fibres and Textiles for Use in Structural Composite Applications) project [34] funded by the European Union was a recent example of an effort to progress the knowledge on bio-based resins and natural fibre fabric precursors. Coordinated efforts were made by research centres and commercial institutions spread across several countries to develop materials and manufacturing techniques for NFCs which could be readily adopted by commercial operations for structural applications. In this context, NFCs are ideal candidates for application of multiscale techniques to predict mechanical properties, so that greater confidence in their application in structures can be established.

This paper explores the application of the computational homogenization technique of Smit et al. [35-37] to predict the behaviour of flax fabric composites subjected to three point bending. Panels of four different flax-based material systems will be manufactured, and their bending behaviour determined through tests on specimens from the panels.

Materials and Methods

In this work, we reinforced a thermoplastic polymer and a thermoset polymer separately with continuous flax fibre reinforcement. For the thermoplastic, Polypropylene (PP) (Moplen RP241G, Lyondell Basell, New Zealand) was obtained in the form of 0.38 mm and 0.6 mm sheets. Prime 20 LV, Gurit, New Zealand was used for the thermoset composites. The matrix polymers properties are listed in Table 1. Unidirectional flax fibre fabric (Belgian flax, *linum usitatissimum*) of areal density 190 g/m² was obtained from Libeco, Belgium.

Glass/epoxy composites were also manufactured and tested to validate some of the numerical models. Glass unidirectional (UD) fabric of areal density 250 g/m² obtained from Gurit, New Zealand was used to manufacture these composite panels.

Table 1. Properties of the matrix polymers.

Material	Property	Value
PP	Density	0.9 g/cm ³
	Tensile modulus, E	1.1 GPa
	Yield strength, σ_y	10 MPa
Epoxy (Prime 20 LV)	Density	1.089 g/cm ³
	Tensile modulus, E	3.2 GPa
	Yield strength, σ_y	73 MPa

Manufacture

Flax/PP panels of two volume fractions, 0.22 and 0.41 were manufactured using a 100 tonne press. PP sheets and layers of flax fabric were assembled and compacted in a die which was pre-heated to 190 °C. A pressure of 0.5 MPa was applied for 10 minutes initially, which was then increased to 0.94 MPa over 5 minutes. The die was then cooled to 100 °C, maintaining the pressure at 0.94 MPa.

The vacuum-assisted resin transfer moulding (VARTM) process was used to manufacture the flax/epoxy panels of two different volume fractions, 0.41 and 0.51, and glass/epoxy panels of volume fraction 0.51. The mould was heated prior to placing the fabric inside. On completing injection, the mould was heated to 60°C to ensure complete curing of the resin.

For further details of both manufacturing processes, the authors' earlier paper can be referred to [35]. Designations have been assigned to the composite materials for ease of reference. The flax/PP composites with 0.22 and 0.41 volume fractions will be referred to as FLPP22 and FLPP41 respectively. Similarly, the flax/epoxy systems composites with 0.41 and 0.51 volume fractions will be referred to as FLEP41 and FLEP51 respectively, and the glass/epoxy system as GLEP51.

Mechanical tests and analysis

Three-point bending tests were performed to study the macroscale behaviour of the composite materials, following the ASTM D790 standard. Rectangular specimens 84 mm long by 16 mm wide for three-point bend (flexural) tests were extracted from the panels produced. As per the ASTM D790 standard, the strain rate applied to flexure specimens was decided based on initial testing done on a sacrificial specimen. The support span and strain rates for all material systems were calculated from the dimensions of this specimen, as per the procedure specified in the standard, and these values are listed in Table 2.

Table 2. Parameters used for three point bend/flexure tests on composite specimens

Material system	V_f	Designation	Support span [mm]	Strain rate [mm/min]
Flax/PP	0.22	FLPP22	47.85	1.315
Flax/PP	0.41	FLPP41	47.85	1.315
Flax/epoxy	0.41	FLEP41	63.74	1.737
Flax/epoxy	0.51	FLEP51	63.74	1.750
Glass/epoxy	0.51	GLEP51	63.64	1.728

As part of a previous study [35], we also characterized mechanical properties at the microscale. The strength of flax fibres under tension was determined using single fibre tensile tests (SFTTs) performed according to the ASTM C1557 standard, with specimens of four different gauge lengths of 10, 15, 20 and 25 mm being tested. The normal strength of the interface between fibres and matrix polymer was also studied for flax/PP and flax/epoxy using the microbond technique. Transverse tensile strength of composites was determined following the ASTM D3039 standard to estimate the shear strength of the fibre-matrix interface.

Macroscale properties

Tensile testing of rectangular specimens was performed using a 30 kN Instron 5567 UTM with a video extensometer to measure strain. As per the ASTM D3039 standard, testing was performed at a crosshead rate of 2 mm/min. Compressive tests were performed either using the Instron 5567 UTM or a 100 kN Instron UTM depending on the final load required for the material system. The ASTM D6641 standard using a combined loading fixture was employed, and the test was performed at a crosshead speed of 1.3 mm/min. Rail shear specimens were tested as per the ASTM D4255 standard. Composite specimens were also tested in flexure by performing three-point bending tests following the ASTM D790 standard.

Fabric geometry characterization

A Leica MZ16 microscope with a maximum magnification of 220x was used for optical microscopy measurements. Various optical measurements, such as the centre-to-centre distance between individual warp yarns and individual weft yarns, were performed on the fabrics. These parameters are required for the modelling of the yarn paths, crossovers and spacings, which constitute the architecture of the fabric. For the warp yarns, the distributions of yarn heights and yarn widths were measured for suitability of fitting to normal, log-normal and Gumbel distributions. The fabric geometry distributions were applied to construct the geometric model of the flax fabric using varying yarn geometries [36].

Numerical Modelling

The approach taken to implement the multiscale coupling for this study is illustrated in the flowchart (Figure 2), including the models involved and the exchange of information between them.

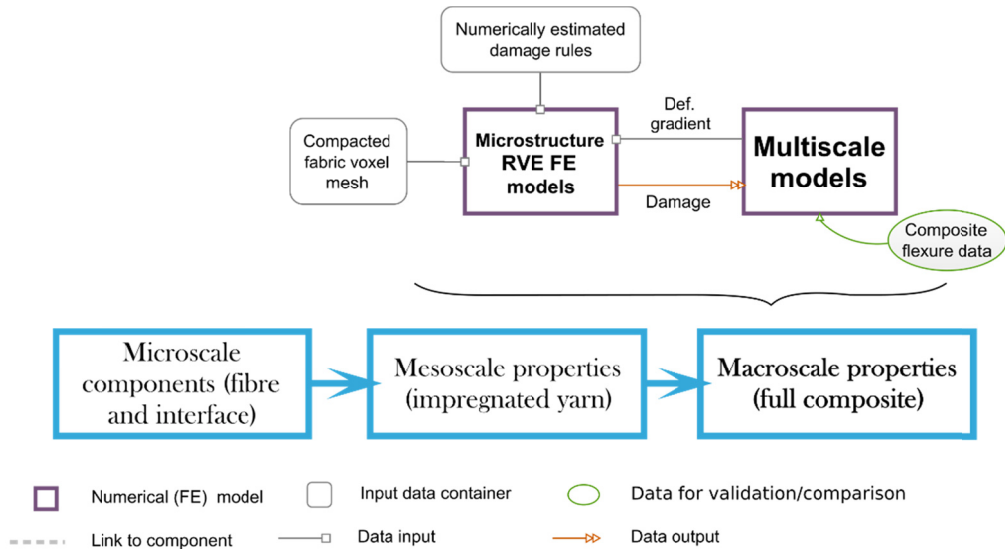


Figure 2. Data exchange involved in the multiscale FE framework

Macroscale model of three-point bending

The multiscale coupling in this study was used to analyse the behaviour of flax composite specimens under three-point bending. As shown in Figure 3, the three-point bending (TPB) test consisted of a rectangular specimen balanced on top of two steel rollers, and a loading roller bending the specimen by pushing down its middle section. The test configuration was represented by a half-symmetry model with solid elements for the beam, and analytical rigid parts representing the loading and support rollers (Figure 3). The ABAQUS CAE python scripting interface was used to parametrically generate the models for the different material systems, and the *getInputs* command enabled the collection of the geometry, material and boundary condition parameters.

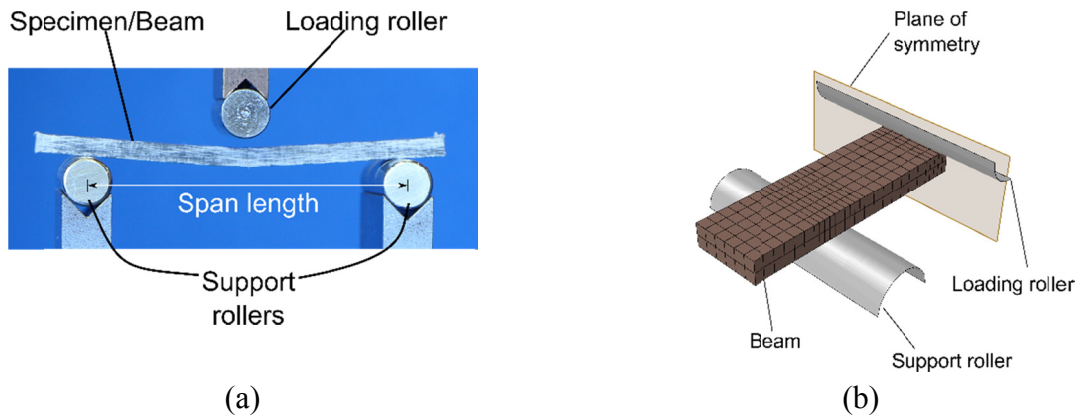


Figure 3. (a) Configuration of the three-point bending test, and (b) representation of the geometry in the FE model of the system

The rollers were positioned so as to replicate the span length used for the particular material system (Table 2). Contact definitions were made between each roller and the beam, and increased mesh refinement was used in areas of the beam closer to the rollers (Figure 3). An initial step was created, displacing the support roller upwards by a small displacement of 10^{-3} mm to establish contact with the beam. In the subsequent step, the loading roller was displaced downwards to deflect the beam.

Microstructure RVE FE models

In a previous study [36], we established a methodology to estimate the damage evolution behaviour in flax/PP and flax/epoxy systems using numerical representative volume element (RVE) models of flax yarns impregnated with either PP or epoxy. The damage rules obtained from these models were then combined with *meso-FE models* [37], which would then be able to provide the macroscale models with a damage response given the deformation gradient. The meso-FE models in this case had discrete representation of the fabric geometry in the flax fabric composites. These are the models referred to as “Microstructure RVE FE models” in Figure 2. The damage rules which were estimated flax/PP and flax/epoxy systems are as in Table 3.

Table 3. Polynomial fit parameters for numerical damage evolution

Material system	Initial strain	a_0	a_1	a_2	a_3
Flax/PP					
d11	0.0	0.0466	-11.326	6899.025	-177591.065
d22	0.002	0.1594	-84.802	21261.602	-1007230.523
d33	0.002	0.1594	-83.350	20546.880	-986335.330
d12	0.002	0.1592	-84.657	21233.65	-1005611.546
d13	0.0	0.04	-0.581	5227.291	-111522.741
d23	0.002	0.1594	-84.807	21262.992	-1007340.688
Flax/epoxy					
d11	0.008	3.950	-1148.067	104071.570	2752583.210
d22	0.0065	0.0627	-14.474	670.675	18751.528
d33	0.0065	0.0672	-15.484	731.900	17979.630
d12	0.0065	0.072	-13.301	207.513	41757.377
d13	0.01	2.644	-578.704	39704.457	-803754.974
d23	0.0065	0.091	-21.596	1217.099	5652.155

An example of flax composite geometry in the RVEs is shown in Figure 4. From the compacted geometry representing stacks of flax fabrics (Figure 4), FE models were generated with elements filling in spaces between the fabric layers to represent the polymer resin (Figure 4).

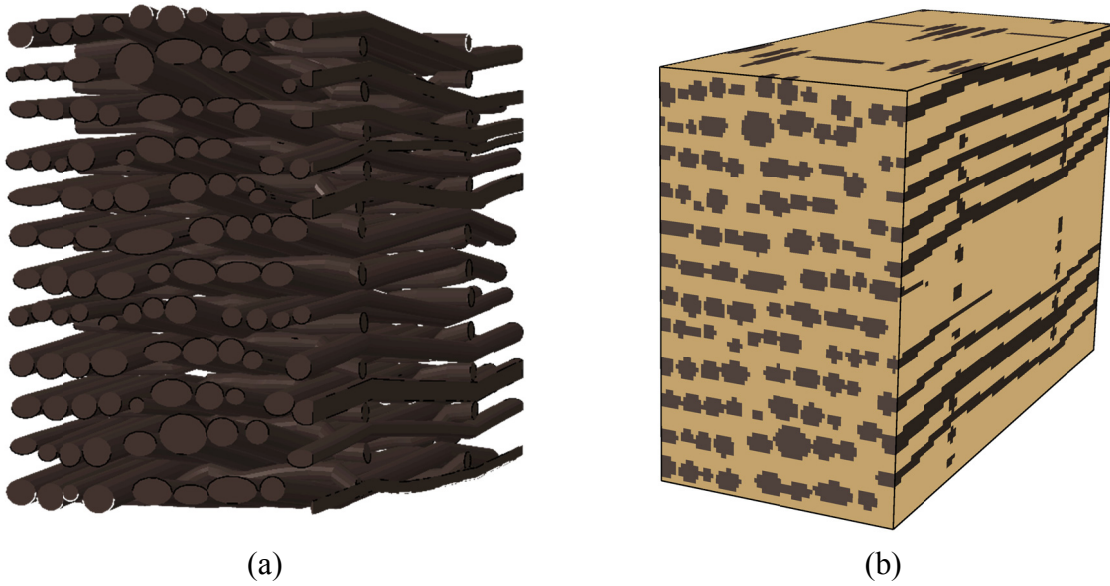


Figure 4. Illustration of (a) section of compacted fabric geometry and (b) RVE of flax fabric composite

Multiscale coupling

The coupling of scales was performed using a coupled computational homogenization method [31] because it is relatively easy to implement, and is also non-intrusive. The term “*non-intrusive*” indicates that the method does not require access/modifications to the numerical solver code itself.

In the method, the macroscopic model is solved first, and for each iteration and each integration point in the elements, the macroscopic deformation gradient F_{macro} is obtained. In the ABAQUS UMAT user subroutine, the deformation gradient is stored in the variable DFGRD1. The RVE model displacements were calculated from the deformation gradient using equations of the form

$$[X - x] = [F_{macro} - I] \cdot [x] \quad (1)$$

where X and x are the positions of a point in the RVE model in the deformed and original configurations, respectively, and I is the identity matrix. Periodic boundary conditions were imposed on the boundary nodes of the RVE models. Pilot nodes were used to control the boundary nodes, due to which displacements only needed to be specified at the pilot nodes. The RVE model was then solved, followed by volume-averaging of the stress or damage values in the RVE. The damage in the RVE was set by using the numerically estimated strain-damage rules (Table 3). The average values were then updated in the corresponding integration point in the macroscale model. The flow of data between the macroscopic and microscopic meshes is illustrated in Figure 5.

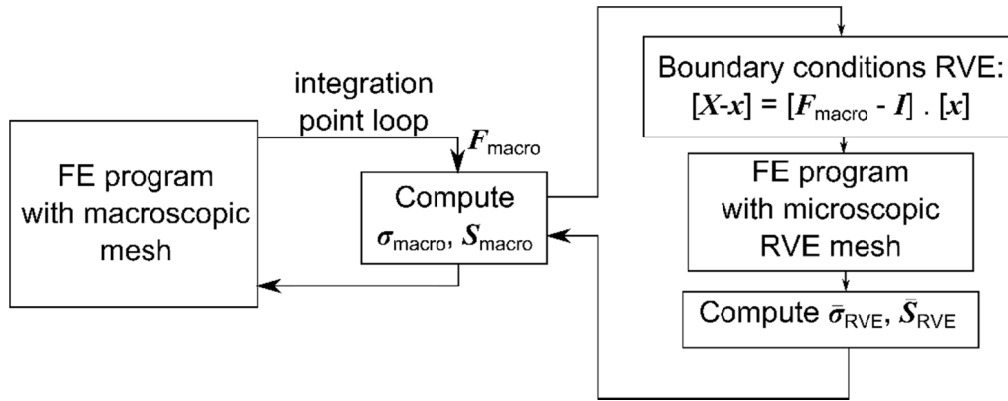


Figure 5. Flowchart of steps involved in the computational homogenization scheme of Smit et al. [31]

The macroscopic meshes used as part of the multiscale coupling in this work were from the FE models representing the three-point bending test (Figure 3). The composite beams were modelled as rectangular blocks with isotropic material properties in these. The microscopic or lower scale models were comprised of the meso-FE RVE models.

Results and Discussion

A hierarchical approach was applied, which includes obtaining material properties at the micro-scale, linking them to analytical or numerical models, and then feeding the micro-scale properties into a larger scale model. Fibre properties were determined first, fit to Weibull distributions, and compared to predictions using the classical laminate theory. The fibre properties were then used in simulations of the microbond test performed to determine properties of the interface. The properties of the fibre and the interface were subsequently used to model the failure of matrix-impregnated flax yarns.

Fabric geometry characterization

The measurements of the cross-sections of flax fibres and flax yarns are indicated as distributions in Figure 6. Concerning the flax fabrics, the term “crossover” will from now on mean when a warp yarn passes over a weft yarn or vice-versa. On observation of the fabrics, it was found that the average distance from the edge of one warp flax yarn to the next was 0.005 mm (within the same crossover), while the corresponding value for weft yarns was 2.85 mm. For the glass fabric, the glass tows were closely spaced, and were 2.08 mm wide and 0.24 mm high on average. The tows were bound by polyethylene yarns 0.3 mm wide and 0.2 mm high, which were spaced 37 mm apart.

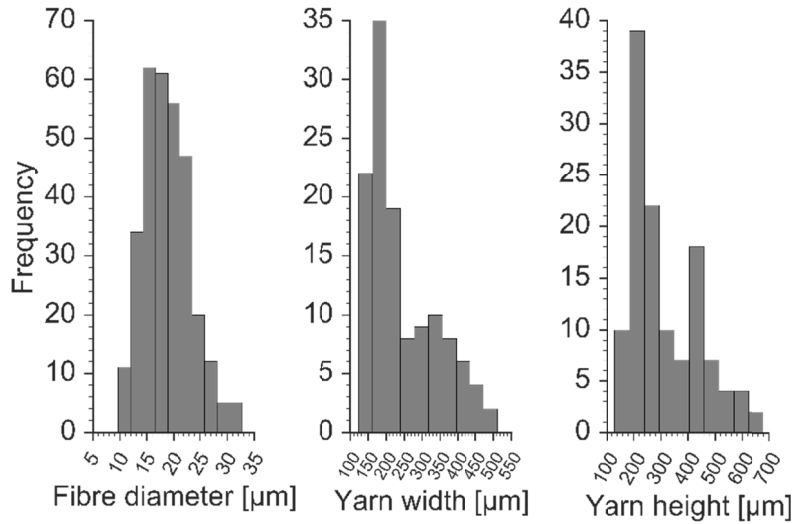
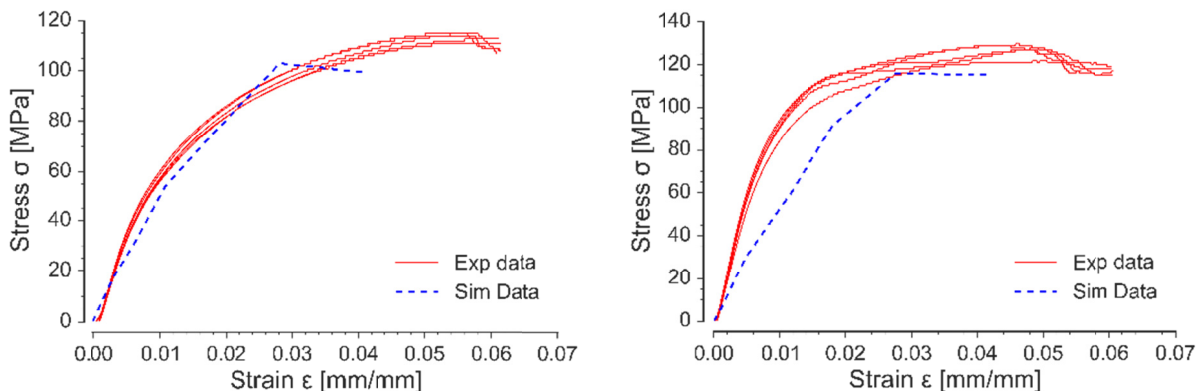


Figure 6. Distributions of fibre and yarn geometries, from optical measurements

Multiscale model results

The simulation results, presented in Figure 7, display reasonably good agreement with the experimental data for all four systems. For the flax/PP systems, the predicted stresses at failure matched closely with those obtained from the tests. For the FLPP22 model, with the numerical damage rule applied, the predicted failure was at 103.10 MPa, which is 8.3% lower than the average experimental value of 112.40 MPa. However, the failure does seem to occur early, with the failure strain of 0.028 being less than half the average strain of 0.060 in the tests. The higher failure strain in the test specimens could be due to the unwinding of the fibres in the yarn helix, and of the cellulose microfibrils in the fibres. Fractographic analysis could confirm the occurrence of such a process. However, fractographic analysis of micromechanics is not a focus of this thesis, and is something that could be investigated as part of another research work.

The prediction for FLPP41 was at a value of 115.78 MPa, which is 7.7% lower than the test value. The failure strain of 0.027 was again lower than the average strain of 0.050 in the tests. Both the stress and strain behaviours predictions were much closer to the average experimental values for the flax/epoxy systems. In the case of the FLEP41 model, the predicted strength was 236.36 MPa, 11.2% lower than the experimental average of 296.40 MPa, while for the FLEP51 model, the strength was predicted to be 299.14 MPa, only 7.5% lower than the test average of 323.4 MPa. The corresponding failure strain prediction of 0.032 was only 7.4% lower than the test average strain of 0.030.



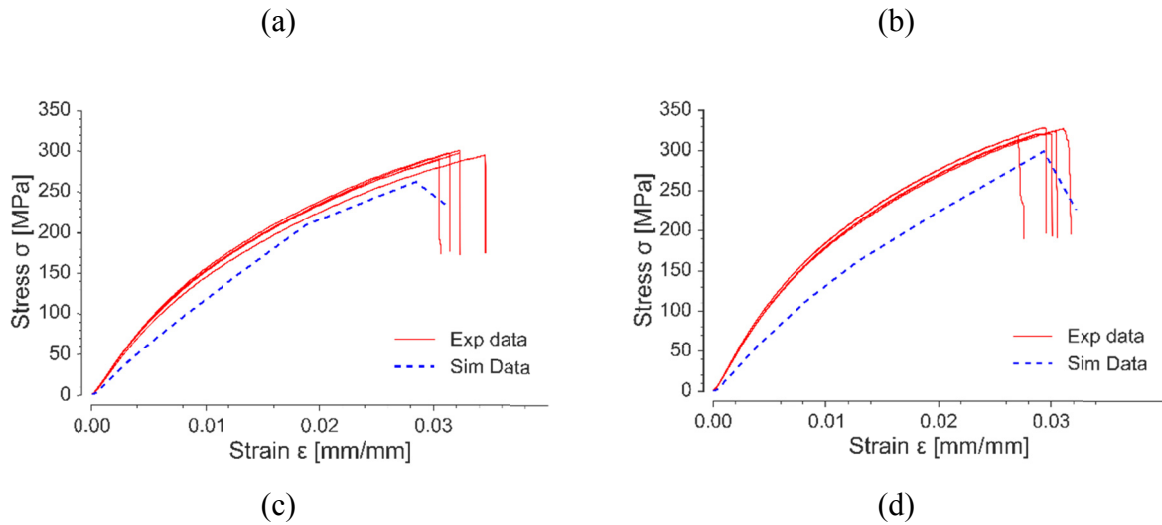


Figure 7. Results from multiscale simulations of three-point bending behaviour of (a) FLPP22, (b) FLPP41, (c) FLEP41 and (d) FLEP51

Validation with Experiments

The multiscale modelling framework here was demonstrated to yield good results for the four flax composite systems studied. To ascertain the applicability of this method to other material systems, the entire set of experiments and simulations leading up to the multiscale simulations has been repeated for a glass/epoxy material system (GLEP51). Glass/epoxy panels with a fibre volume fraction of 0.51 were manufactured using resin transfer moulding, as mentioned before and subjected to three-point bending tests.

Impregnated yarn RVE and damage rule

Using a similar method as described in our previous paper [36] to establish damage evolution rules for fabric composite systems, discrete models of glass tows impregnated by epoxy resin were constructed and samples cut out to identify RVEs. Damage was calculated based on the values of strain with respect to strain values expected at the plastic yield point of the epoxy and glass fabric. The strength values used for the glass fabric were [38]:

- Longitudinal tensile strength = 2000 MPa
- Longitudinal compressive strength = 1000 MPa
- Transverse tensile strength = 80 MPa
- Transverse compressive strength = 250 MPa
- Shear strength = 100 MPa

The strength used to calculate the yield strain for epoxy was 73 MPa (Table 1). The damage evolution from tensile and compressive loading of the glass/epoxy RVE model were calculated, and best-fits for their evolution obtained by curve-fitting. The parameters for the curves thus obtained for the evolution of damage in the glass elements are specified in Table 4.

Table 4. Polynomial fit parameters for numerical damage evolution of glass/epoxy impregnated yarn

Damage variable	Start strain	a_0	a_1	a_2	a_3
Tensile					
d11	0.005	-1.386	365.707	-19,101.067	333,478.08
d22	0.005	-1.080	284.011	-14,597.98	248,129.125
d33	0.005	-1.077	283.139	-14,553.156	247,367.21
d12	0.002	0.149	-81.43	19,771.52	-865,447.951
d13	0.002	0.153	-86.178	20,474.23	-904,093.44
d23	0.002	0.148	-83.358	18,810.508	-796,770.62
Compressive					
d11	0.0	0.043	-39.310	3,324.354	1,196,582.368
d22	0.0	0.106	-5.213	1,040.962	956,921.524
d33	0.0	0.105	-3.460	1,163.322	908,359.764
d12	0.0	0.110	-5.382	1,074.801	988,028.244
d13	0.0	0.038	-40.331	3,855.576	1,052,958.606
d23	0.0	-0.009	25.622	-2,732.963	1,046,119.704

Multiscale three-point bending

A meso-FE model was constructed for the GLEP51 system using the same methodology utilized to construct meso-FE models for the flax-based systems [37], and the geometry parameters obtained for the glass fabric. The uncompacted and compacted glass fabric geometries used in generating the meso-FE models are shown in Figure 8.

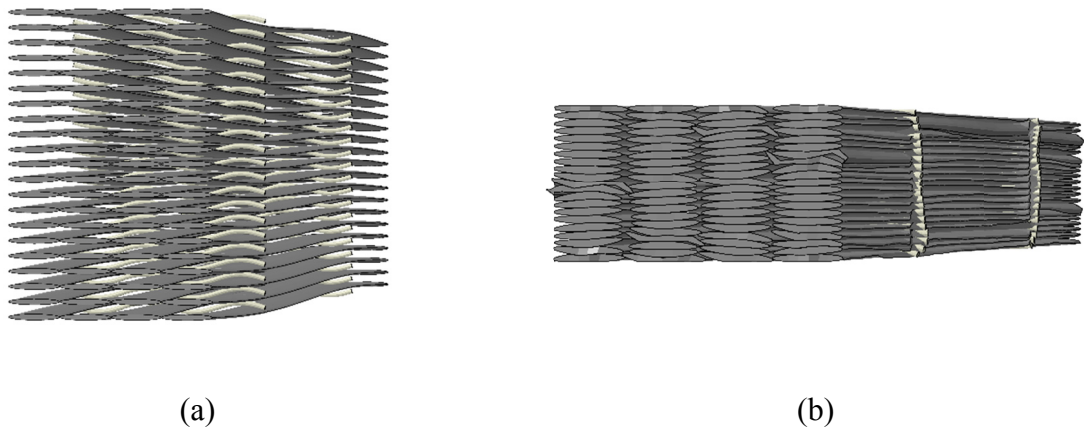


Figure 8. Glass fabric stack (a) before and (b) after compaction

A model representing the three-point bending of glass/epoxy specimens was constructed, with the initial elastic stiffness of the glass/epoxy composite set to 42.58 GPa. Subsequently, coupled multiscale studies were undertaken to simulate the three-point bending behaviour of GLEP51 models. The bending model was coupled with the meso-FE RVE model for glass/epoxy using FORTRAN code and the UMAT interface in ABAQUS. The bending behaviour predicted by the model was close to the experimental results in terms of the

stiffness behaviour and the failure strength value, as shown in Figure 9. The failure was predicted at 898.26 MPa, which is only 12.6% lower than the test average of 1028.17 MPa.

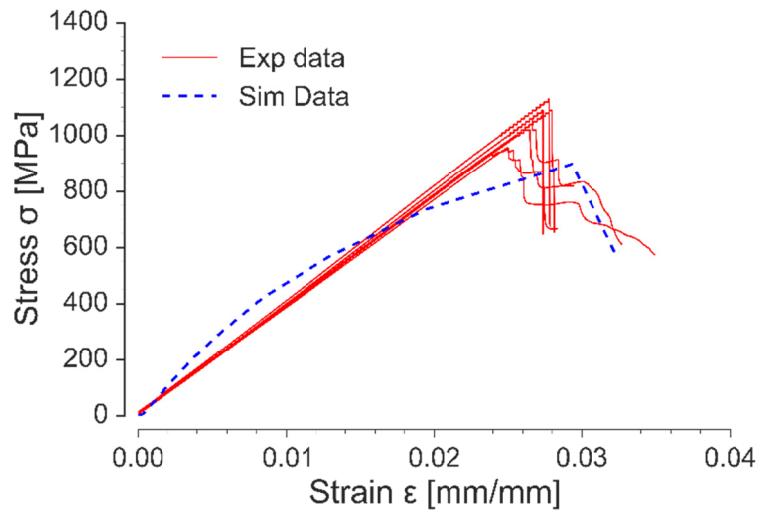


Figure 9. Stress-strain response from the multiscale simulation of three-point bending behaviour of GLEP51, compared to bending test data

Conclusions

A coupled multiscale homogenization technique was implemented to analyze the mechanical behaviour of flax/PP and flax/epoxy fabric composites using ABAQUS and FORTRAN code. Two scales were considered, one of which was an FE representation of a three-point bending specimen. The coupled microscopic meshes or lower scale models were FE models containing discrete representations of the fabric geometry. The failure of the impregnated yarn elements in the lower scale models was implemented using strain-damage evolution laws estimated numerically from the fibre and interface properties. The implementation was used to simulate the bending behaviour of the composites. The results obtained for the failure strength were between 7.5-11.2% lower than the test average values.

Validation studies for the technique were also performed, using geometry and material parameters for a glass/epoxy composite, combined with flexure test data. Numerical damage evolution laws were obtained for the glass/epoxy system. The coupled homogenization technique was then applied to the glass/epoxy system, as performed for the flax composites. The stiffness and strength behaviours were reasonably close to those of the test specimens, and the predicted failure strength was only 12.6% lower than the experimental average for the glass/epoxy three-point bending specimens.

Finite element models containing discrete representations of the fabric geometry (meso-FE models) can be used to predict the tensile failure of natural fibre composites, as demonstrated using flax fibre-based composites. This was demonstrated for four flax/polymer composite systems by combining the fabric geometry with the strain-damage evolution rules. These rules were obtained by constructing FE models from cuboid samples extracted from the impregnated yarn geometry. Equipped with the fibre, interface and matrix polymer properties, damage evolution laws can be obtained for any fibre-polymer combination by following the approach in this work.

The present work establishes the reliability in applying numerical damage rules for multiscale modelling of natural fibre-based composites. To do so, a coupled multiscale homogenisation technique was implemented to establish its capability to predict the mechanical behaviour of

natural fibre thermoplastic and thermoset composite materials. This has been demonstrated by applying the multiscale model to the bending analysis two flax/PP and two flax/epoxy composite systems, with reasonably accurate results obtained. The multiscale framework was two-scale, consisting of a homogenised material model at the highest level coupled with a microstructure model. A validation study was performed to establish the reliability of the same framework using a glass/epoxy composite material system, which was also able to predict the composite failure with good accuracy. Overall, these sets of results establish the confidence in the potential of this multiscale framework implementation in relation to its applicability for different composite material systems.

Acknowledgements

The authors would like to thank the Ministry of Business, Innovation and Employment, New Zealand, for the funding provided to support this project through grant no. 3625485 titled “Sustainable Composites”. The authors also wish to acknowledge the contribution of NeSI high-performance computing facilities to the results of this research.

References

- [1] Matzenmiller, A., Lubliner, J., and Taylor, R.L. (1995) A constitutive model for anisotropic damage in fiber-composites, *Mechanics of Materials* **20** (2), 125-152.
- [2] Nairn, J.A. (2000) 2.12 - Matrix Microcracking in Composites, *Comprehensive Composite Materials*. 403-432.
- [3] Talreja, R. (1985) A Continuum Mechanics Characterization of Damage in Composite Materials, *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **399** (1817), 195 -216.
- [4] Llorca, J., Gonzalez, C., Molina-Aldareguia, J.M., Segurado, J., Seltzer, R., Sket, F., Rodríguez, M., Sadaba, S., Munoz, R., and Canal, L.P. (2011) Multiscale Modeling of Composite Materials: a Roadmap Towards Virtual Testing, *Advanced Materials*.
- [5] Nguyen, B.N., Bapanapalli, S.K., Holbery, J.D., Smith, M.T., Kunc, V., Frame, B.J., Phelps, J.H., and Tucker, C.L. (2008) Fiber Length and Orientation in Long-Fiber Injection-Molded Thermoplastics - Part I: Modeling of Microstructure and Elastic Properties, *Journal of Composite Materials* **42** (10), 1003-1029.
- [6] Xiao, Q.Z., and Karihaloo, B.L. (2010) Two-scale asymptotic homogenisation-based finite element analysis of composite materials, *Multiscale Modeling in Solid Mechanics: Computational Approaches*. 43-100.
- [7] Llorca, J., González, C., and Segurado, J. (2007) Finite element and homogenisation modelling of materials, *Multiscale materials modelling*. 121-147.
- [8] Smit, R.J.M., *Toughness of Heterogeneous Polymeric Systems : A Modeling Approach*, Eindhoven Technical University, 1998.
- [9] Noor, A.K. (1986) Global-local methodologies and their application to nonlinear analysis, *Finite Elements in Analysis and Design* **2** (4), 333-346.
- [10] Whitcomb, J.D. (1991) Iterative global/local finite element analysis, *Computers & Structures* **40** (4), 1027-1031.
- [11] Fish, J., Belsky, V., and Pandheeradi, M. (1996) Composite grid method for hybrid systems, *Computer Methods in Applied Mechanics and Engineering* **135** (3\textendash4), 307-325.
- [12] Haidar, K., Dubé, J.F., and Pijaudier-Cabot, G. (2003) Modelling crack propagation in concrete structures with a two scale approach, *International Journal for Numerical and Analytical Methods in Geomechanics* **27** (13), 1187-1205.
- [13] Mote, C.D. (1971) Global-local finite element, *International Journal for Numerical Methods in Engineering* **3** (4), 565-574.
- [14] McCormick, S., and Thomas, J. (1986) The fast adaptive composite grid (FAC) method for elliptic equations, *Mathematics of Computation* **46** (174), 439-456.
- [15] Fish, J., Suvorov, A., and Belsky, V. (1997) Hierarchical composite grid method for global-local analysis of laminated composite shells, *Applied Numerical Mathematics* **23** (2), 241-258.
- [16] Fish, J., Shek, K., Pandheeradi, M., and Shephard, M.S. (1997) Computational plasticity for composite structures based on mathematical homogenization: Theory and practice, *Computer Methods in Applied Mechanics and Engineering* **148** (1-2), 53-73.

- [17] Feyel, F., and Chaboche, J.-L. (2000) FE2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre SiC/Ti composite materials, *Computer Methods in Applied Mechanics and Engineering* **183** (3-4), 309-330.
- [18] Peng, X., and Cao, J. (2002) A dual homogenization and finite element approach for material characterization of textile composites, *Composites Part B: Engineering* **33** (1), 45-56.
- [19] Fang, Z., Sun, W., and Tzeng, J.T. (2004) Asymptotic Homogenization and Numerical Implementation to Predict the Effective Mechanical Properties for Electromagnetic Composite Conductor, *Journal of Composite Materials* **38** (16), 1371-1385.
- [20] Ghosh, S., and Mukhopadhyay, S.N. (1993) A material based finite element analysis of heterogeneous media involving Dirichlet tessellations, *Computer Methods in Applied Mechanics and Engineering* **104** (2), 211-247.
- [21] Sanchez-Palencia, E. (1987) General introduction to asymptotic methods, *Lecture Notes in Physics*. 121-136.
- [22] Bramble, J.H., and Pasciak, J.E. (1990) A domain decomposition technique for Stokes problems, *Applied Numerical Mathematics* **6** (4), 251-261.
- [23] Farhat, C., Wilson, E., and Powell, G. (1987) Solution of finite element systems on concurrent processing computers, *Engineering with Computers* **2** (3), 157-165.
- [24] Borgers, C. (1989) The Neumann-Dirichlet domain decomposition method with inexact solvers on the subdomains, *Numerische Mathematik* **55** (2), 123-136.
- [25] Toselli, A., and Widlund, O. (2005) *Domain decomposition methods: algorithms and theory*, Springer.
- [26] Ladeveze, P., Loiseau, O., and Dureisseix, D. (2001) A micro-macro and parallel computational strategy for highly heterogeneous structures, *International Journal for Numerical Methods in Engineering* **52** (1-2), 121-138.
- [27] Boso, D.P., Lefik, M., and Schrefler, B.A. (2006) Homogenisation methods for the thermo-mechanical analysis of Nb₃Sn strand, *Cryogenics* **46** (7-8), 569-580.
- [28] Renard J, M.M. (1987) Etude de l'initiation de l'endommagement dans la matrice d'un materiau composite par une methode d'homogenisation, *Aerosp Sci Technol* **6**, 37-51.
- [29] Kouznetsova, V., Geers, M.G.D., and Brekelmans, W.A.M. (2010) Computational homogenization for non-linear heterogeneous solids, *Multiscale Modeling in Solid Mechanics: Computational Approaches*. 1-42.
- [30] Carvelli, V., and Poggi, C. (2001) A homogenization procedure for the numerical analysis of woven fabric composites, *Composites Part A: Applied Science and Manufacturing* **32** (10), 1425-1432.
- [31] Smit, R.J.M., Brekelmans, W.A.M., and Meijer, H.E.H. (1998) Prediction of the mechanical behavior of nonlinear heterogeneous systems by multi-level finite element modeling, *Computer Methods in Applied Mechanics and Engineering* **155** (1-2), 181-192.
- [32] Bismarck, A., Baltazar-Y-Jimenez, A., and Sarikakis, K. (2006) Green Composites as Panacea? Socio-Economic Aspects of Green Materials, *Environment, Development and Sustainability* **8** (3), 445-463.
- [33] Suddell, B.C., and Evans, W.J. (2005) Natural Fiber Composites in Automotive Applications in Natural Fibers in Biopolymers & Their BioComposites, *Natural Fibers, Biopolymers, and Biocomposites*. 231-259.
- [34] Commission, E.
- [35] Panamoottil, S.M., Das, R., and Jayaraman, K. (2016) Towards a multiscale model for flax composites from behaviour of fibre and fibre/polymer interface, *Journal of Composite Materials*.
- [36] Panamoottil, S.M., Das, R., and Jayaraman, K. (2016) Experimentally quantified and computational anisotropic damage rules for flax fabric composites, *International Journal of Damage Mechanics*.
- [37] Lomov, S.V., Ivanov, D.S., Verpoest, I., Zako, M., Kurashiki, T., Nakai, H., and Hirose, S. (2007) Meso-FE modelling of textile composites: Road map, data flow and algorithms, *Composites Science and Technology* **67** (9), 1870-1891.
- [38] Gan, J.M., Modelling the initiation and evolution of damage within GFRP by including real geometric variability. *18th International Conference on Composite Materials*, 21-28 August, Jeju Island, Korea, 2011.

Cosmotic, Aquatic.
Exploring the Potential of Computational Design in the Preservation of
Aquatic ecotones.

Aya mohanna

Department of Architecture and Project, La Sapienza University of Rome, Italy.

Author: aya.mohanna@uniroma1.it

Abstract

This paper looks at the possible role of computational design ecologically in the fight against the loss of the aquatic Ecotone. As climate change keeps altering all the natural aspects of our planet, and as our kind continues to sabotage its ecologies, coral reefs come in focus. Aquatically, coral reefs count as a fertile zone for biodiversity. Usually being the Ecotone between land and sea, these barriers host many species and riches. However, due to the excessive abuse caused by human activity be it world-wide pollution or direct human contact, these reefs are constantly bleaching and breaking. In 2016, the Architecture Association gathered a group of international architecture students and professionals in a visiting school in Jordan titled “Hyperbolic Reefs” looking at the possibility of recruiting new computational methods to preserve and possibly regenerate the Ecotone. It was considered that new simulation techniques along with parametric design could contribute into the assessment and prevention of the catastrophic results.

The two-week event was divided into chapters and was initiated by a series of lectures and discussions conducted by worldwide leading architects and experts who presented an important material to build upon. Then, the participants underwent a site visit to the coral reef of Al-Aqaba, collecting data, samples and media and recording insights and local testimonies. The third step of the experience was to assimilate the material and data and discuss openly the ways that computation could lead to a better coral life. Several software and tools were assigned to produce a design that would help attenuate the compromise of the coral reef through computation. An archive of data was produced and exhibited to the public.

The results of this brief exercise was a number of suggestions and future aspirations triggered solely towards revitalizing the Ecotone. Issues such as the abundance of irresponsible snorkeling and diving, many governments’ indifferent policies towards the coral reefs, global warming, climate change, coral bleaching and aquatic architecture were confronted through parametric projects ranging from purely architectural to abstract human capsules. Computational tools allowed the reproduction of the whole system digitally, the precise tracing of the corals’ patterns, dimensions and colors, simulation software predicted the role of light and heat in certain zones, and parametric programs provided an incomparable flexibility in the designing process, going completely in sync with the fragile and intricate aspect of a coral unit. 3D printing was also an integral factor in the presentation and study of the presented models.

This study’s scope was to expand the use of computation in a theoretical way to reach new and creative prospects, and to raise awareness to the situation of the coral reef and the risks facing its degradation.

Keywords: Computation, Artificial reefs, Performance Architecture, Aquatic, Coral resilience, Regeneration, Detoxification.

Introduction

The discussion of coral reefs has been repeatedly cast on an international level in the recent decade. Following the repercussions of El Niño and what it caused of coral bleaching worldwide, the attention was then moved to the global threat to coral reefs due to human activities. Coral reefs maintain an integral importance environmentally and economically for bearing the richest ecosystems worldwide and for their role in several fields in human activity. However, due to continuously increasing climate issues along with many other damaging factors, phenomena of bleaching and coral mortality have been widespread globally. Events of skeletal destruction and loss of coral reefs are compromising the rich Ecotone and endangering the local species. This was confronted by relatively scarce attention and weak policies to fight against the menacing situation.

On an international scale, some countries are starting to look into alternative solutions to preserve the coral diversity and help save the inhabiting species under danger. This comes in both direct and indirect ways; such as enclosed protected zones and the implementation of complex structures on the shallow end of the sea called artificial reefs. These reefs are executed in many ways and forms and with a wide spectrum of materials and theoretically offer an alternative skeletal structure to corals and their dwelling species. Applications of these artificial installations are still very minimal and preliminary with small effect on a macro scale, but they offer an assuring start to the regeneration process.

On a more promising scope, scientists have noted the existence of coral zones and societies that have been naturally fighting off bleaching and resisting the toxic environment that would otherwise contaminate the reef. This led to the extensive study of the potential factors that would enforce the corals and make them resilient to bleaching elements. These factors range from naturally intrinsic factors to extrinsic factors related to the surroundings of the zone or the external effects of the undamaged area. A list of key components was produced to better understand the dynamic of corals and how they manage to self-preserve.

This paper suggests the idea of implementing the resilience factors of corals and their preserving environments in the production of artificial coral reefs, in a way that would take the idea to an advanced level and to further expect a better outcome of these structures, instead of their current state as discarded non-degradable materials in the ocean. The use of computation is thus suggested and encouraged, due to its rising importance and convincing effects on an architectural platform. Computation will function here on many steps of the design, from conceptualization to implementation with the use of software that will generate possible solutions, tools that will simulate natural elements and therefore grade the spatial performances of the volumes and programs that will chronologically predict the aspect of the skeletal structure. All these steps will be showcased in a case study underwent by a group of design students and young professionals in 2016 that put computation into testing in order to come out with potential designs for the artificial reef of Al Aqaba, Jordan.

This paper will showcase therefore the big potential computation has in the synthesis of natural factors in the architectural design of artificial reefs with an amplified performance.

Aquatic Ecotones

Aside from their striking vista and attractive features, coral reefs are the bearers of the richest ecosystems on earth. These structures that are situated between the shore and the ocean, are very important for the sustainability of our environment and life. On a natural level, they house a great variety of species and aquatic animals and form the architectural skeleton for their habitat and their first defense from direct exposition to the natural dangers from both the shore and the deeper side of the ocean. Coral reef importance lies in the fact that it's the natural Ecotone between land and sea; it thrives with very few nutrients, it protects the shore and the frontier cities from possible tsunamis and mega-waves, and forms a rich environment for developing businesses such as fishing and local craft-making. Also, with the fast rise of the tourism business as an important pole of world economy, coral reefs have become an essential attraction for divers and tourists around the world. Along with the traditional benefits, corals are recently being tested and used for medicine production after the decline of antibiotics due to the secretion of self-defensive chemicals by corals. [2]

Unfortunately however, human activity has seriously damaged a considerable part of this natural asset; the pacific aside, 70% of total coral reefs are endangered. The main components that jeopardize corals are those that cause pollution and disease to the ecosystem. This includes mainly over-fishing, unstudied tourism, shore constructions which lead to sedimentation and coral suffocation, freshwater flooding and more. However, the main precursor of coral damage nowadays is climate change and its symptoms such as increased hear, changing water levels, changing percentages of carbon which affects the algae-coral balance, and severe exposition to UV light and chemicals. [1]

Ultimately, the most affected areas under high risk of climate change are regionally the southeastern Asian reefs, the Caribbean and the Indian Ocean. Other regions with considerable threat are the Middle East and some zones in the Atlantic and the Pacific. The most damaged areas are namely the shores of Indonesia and Australia, with small to no protection of the reefs. Today, protected areas are spreading; however, the serious effort of enclosing and shielding the reefs from the elements is still preliminary and falsified. Over 56% of the coral reefs are under low to high threat of damage and bleaching, and face life-threatening risks.

If not promptly healed, the bleached corals due to hazardous components never regain their health and color and face unavoidable mortality. In a span of 2 years after, dead corals decompose and the skeletal scene of the once diverse coral reef is drastically changed. Coral reduction compromises the ecosystem; the absence of architectural complexity for the dwelling of species will finally limit their survival chances and larvae cultures will reign the shallow waters. Additionally, the nearby towns will lose a critical aspect of their economical state and will be exposed to serious threats from mega-waves and deep water predators. [8]

Artificial Reefs, Current State and Potential

In order to tackle the menacing factors that cause coral mortality, nations worldwide have been seeking methods and policies to help regenerate and revitalize the aquatic Ecotone. Some interventions include the enclosure and the shielding of the coral zones from direct contact, making them inaccessible to tourists, fishermen and other elements. Other countries are taking it a step ahead and are attempting to quicken the regeneration process of the damaged ecosystem by the implementation of artificial reefs.

Basically, they're the addition of particular structures and volumes with features close to coral skeletons in hopes to encourage the regeneration of the area and to quicken the rehabilitation process of the corals [4]. As a spectrum of used materials, artificial reefs are often non-degradable structures made out of concrete, castoff tires, discarded automobiles, and shipwrecks. Less often, plastic, PVC, metal and ropes are installed. The main issues when setting an artificial reef are its environmental effects and expected performance; materials should not produce any chemicals, should not decompose, and assigned forms for the new host structure should be complex, void from the inside, with a certain amount of complexity in its entrances in a capsule-like volume [5].

International laws have been put together to guide the execution of artificial reefs, but it still faces issues of corruption, lack of effective management, and the results for these structures are still quite minimal. Studies suggest a need for alternative artificial reefs, ones that would be less superficial and that would look into depth at the performance of the potential solutions [5]. These solutions would take into consideration the repetitive contact of human elements and limit it, while understanding completely the nature of the local species and their specificity in order to better produce a unique and positive result. This would change drastically the reality of artificial reefs that would therefore be offered as an efficient solution to coral mortality. In order to understand the specificity of corals and what harms them, a thorough study on coral weakness and defense mechanisms should be underwent.

Resilience in Corals

Despite the global devastation of the state of coral reefs, experts have noted a repetitive pattern of resilience in some zones that would otherwise be affected. A detailed and close study to the main influences behind the defiance to bleaching in these small localities helped gather a series of factors, intrinsic and extrinsic, that helped the corals thrive under compromising conditions. A study gathered the case studies and organized a set of circumstances in which corals are less affected by means of pollution and disease.

This list includes factors that naturally attenuate bleaching precursors such as UV light exposure, light deficiency, water warming, and varying water levels for long spans of time. The defense mechanisms tackling light exposure were selectively, the existence of a natural shading structure on the beach, cloudy regions and generally turbid waters. Those that tackled warmer waters however were areas with exceptionally strong currents and channels, areas exposed to water upwelling, brief exposure to air, the interchanging stream of water that would flush toxins out, and generally colder regions. As to intrinsic factors, corals that were repetitively exposed to changing temperature for short spans of time and that get exposed often to air and interchanging water levels have learned to overcome these factors and grew unaffected by them on the long term [1].

This study could lead into a design that would revitalize the coral reefs. A structure could ecologically and sustainably reinforce the resilience of corals and help them fight against the elements. Just like detoxification architecture that is aiming to reintegrate nature with the urban landscape through effective biomorphic structure [6], this policy could be extended to sub-aquatic structure that would "renaturalize" the damaged areas and tackle the menacing damage done by tourists, climate change and other causes mentioned above. Examples of detoxification architecture in packed cities have proved that an efficient choice of materials and policies along with good management could help in mending the damage done by humans, leaving the rigid forms and presenting a new experimental spirit [6].

A synthesis is thus proposed, to take the artificial reef domain a step forward, forming structural hubs in the shallow waters that would both have the shielding features of artificial installations and train the damaged corals for a better defense mechanism and a stronger resistance on an extended time span. This would help reintegrate the species with the existing corals and potentially help the regeneration of these corals which will subsequently restore the ecosystem and quicken the healing process of the region.

Establishing quite a design and form would require advanced technologies and techniques, for this matter is very fragile and trial and error are not an option in many cases. This promotes the need of new ways of design that would optimize the performance of the forms and generate a series of solutions that are in grade of testing before implementing them on site.

Computational Possibilities in the Regeneration of Corals

This employs the idea of a structure that would use the analogy of corals in order to come up with the solution. Going from the bottom up, studying the specificities of the corals and their construction, extracting every detail concerning patterns, textures, skeletons and internal spaces.

A call for an experimental design in a fragile nature also requires the employment of techniques that would generate parametric shapes following biomimetic principles of design. Understanding biomimetic design is to understand the way biological elements function; a heterogeneous structure at best, multilayered and offering a differentiation of scale and shape, biological elements are a set of composite materials that could be translated into high-end materials in execution, heavily hierarchal and marked by their multilevel structure ranging from Nano to Macro, suggesting fractal design and parametric analogy [9].

Fortunately, computation offers the biggest chance of the best solution due to the tools it offers through the whole designing process. In fact, computational architecture has been on the rise recently and has produced a collection of experimental outcomes that otherwise would not be executable by traditional design techniques. The rising wave has been present in difference phases of the design process from conceptualization to solution generation up until performance testing. It is helping redefine some once concrete norms and notions of spaces, culture, human dynamic and environment. Faced with a challenge to create a both aesthetically pleasing structure and one that is also high in performance, computation comes in handy in offering a variety of solutions and the ability to test each of these solutions' performance and efficiency [6][8]. In our case of advanced artificial reef design, the set of characteristics and leading factors are human-decided. However, the outcome is ought to be a production of a computational model to be tested depending on the context.

The scope of this process is to extract the biggest number of solutions and proposals by using parametric design and fractal digital productions and then employ other computational tools to extract knowledge from these outcomes. A main benefit of this method, aside from the possibility to explore many designs choices, is to allow the designer to review the possible configurations in an early stage of the design. Afterwards, computation would go hand in hand with the design process to get to a concrete point following the abstract phase; genetic algorithms could be installed, and additional software could mimic the natural surroundings of the project and allow a systematic prediction of the parametric design's performance [8]. This process follows four phases described below.

At first, computation allows the extraction of preliminary data such as environment and other information provided by the designer. This allows the virtual recreation of the current state of the project and allows the withdrawal of some complex natural volumes otherwise impossible to regenerate.

Secondly, a list of criteria is offered to the software and an extensive search in volumes and parameters allows the origination of parameterized attributes that would offer a unique set of projects with experimental volumes and diverse aesthetics. The third step then follows; a step both personal and technical: it involves the subjective selection of the designer for the more aesthetically pleasing designs and then involves simulation and virtual reality prediction to assess the expected performance of each and every configuration offered.

The last step would be to run a processing tool equipped with the right parameters and equations that would deduce the future aspect of the artificial reef after the colonisation of corals in a specified time frame, something otherwise unfeasible without the available computational apparatus.

Application

To showcase this analogy in a practical frame, this chapter will present the proceedings of the 2016 AA visiting school in Jordan, under the title *Hyperbolic Reefs*, which looked closely into addressing the issue discussed in this paper [10].

The beginning of the summit included a basic briefing of the programs and digital instruments to use in the workshop; including Autodesk 123 Catch for 3D scanning and shape extraction, Rhinoceros for modeling, along with grasshopper (parametric design plugin) and beehive, Autodesk Maya for simulation and selection, Processing for the chronological review of the chosen attributes and lastly Keyshot for virtual representation and rendering.

The figures provided below showcase a collection of works produced by the participants of the visiting school that are clear images of the design process [10].

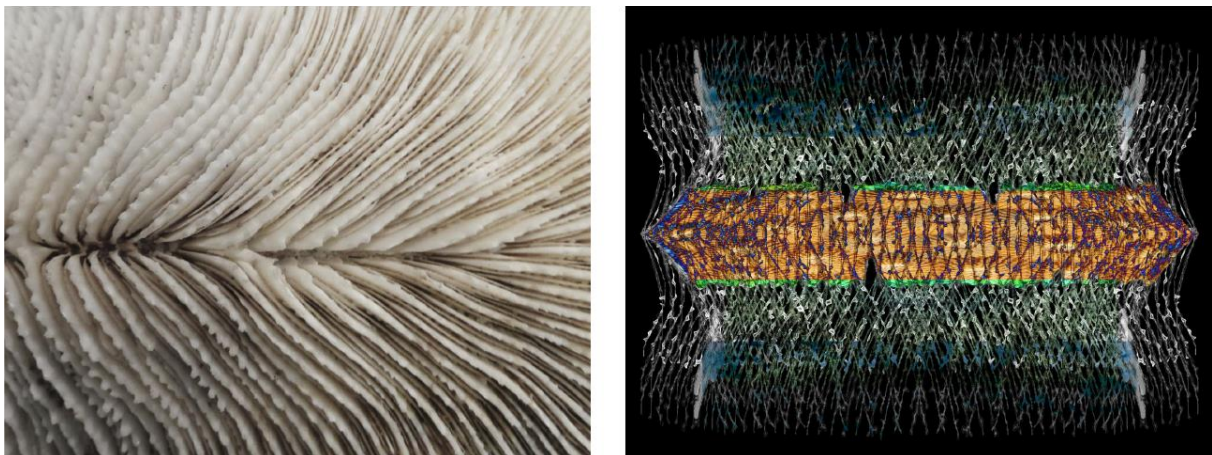


Figure 1. Example of the preliminary scanning and reproduction of a coral unit.

Figure 1 shows the computational regeneration of a skeletal coral structure. 123 Catch was used to trace the voids, protruding branches, shadows and texture of the coral and to map it digitally.



Figure 2. Phase 1: Primary generation of a virtual biomimetic version of the coral.

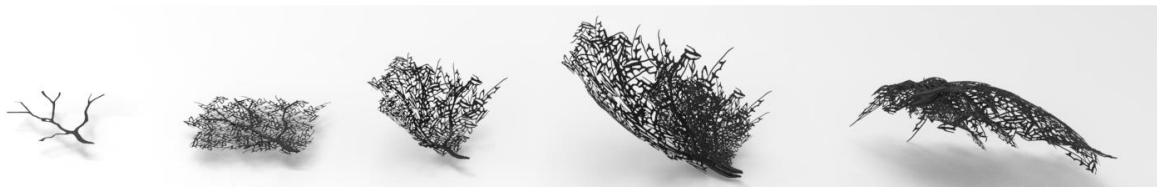


Figure 3. Phase 2: Extracting the coral's modular unit and experimenting with potential generations.

The first phase involved the selection of a primary coral unit that is believed to be persistent and resilient in shape and texture in the studied area. The next step was to scan and regenerate the architectural structure of the species and model it in different ways in order to create an archive of potential generic volumes and shapes for further study. (Figure 2)

Following the first phase, it was possible to excerpt the founding unit of the volume and to experience with it, creating sets of fractal repetitions of the same nucleus and producing potential versions of possible applications to the intended project. This resulted in a wide and diverse collection of choices and shapes that challenged the designers in their quest towards an optimized and positively performing outcome. (Figure 3)

In the work presented in this chapter, the group's concept was to create an artificial reef that would fragment the experience between user and coral, shielding the indigenous species from direct contact and extreme UV light, while visually exposing the tourists to the coral through an inter-system hub, one that would double the notion of the Ecotone it being from shore to sea, and from water to air.

This required undergoing a simulation effort to understand the required volumetric composition following the wanted facilities and to include corals in the colder areas and the ones less exposed to solar radiations and warmer water. Figure 4 shows the application of this simulation on the most efficient volume retrieved.

Figure 5 showcases the intervention that was applied using beehive and rhinoceros to attenuate the opacity of the chosen shape and to allow a better performance to the complex. It studies openings for human facilities and coral capsules. The resulting interference was the application of a porous structure that would diffuse natural light and wind.

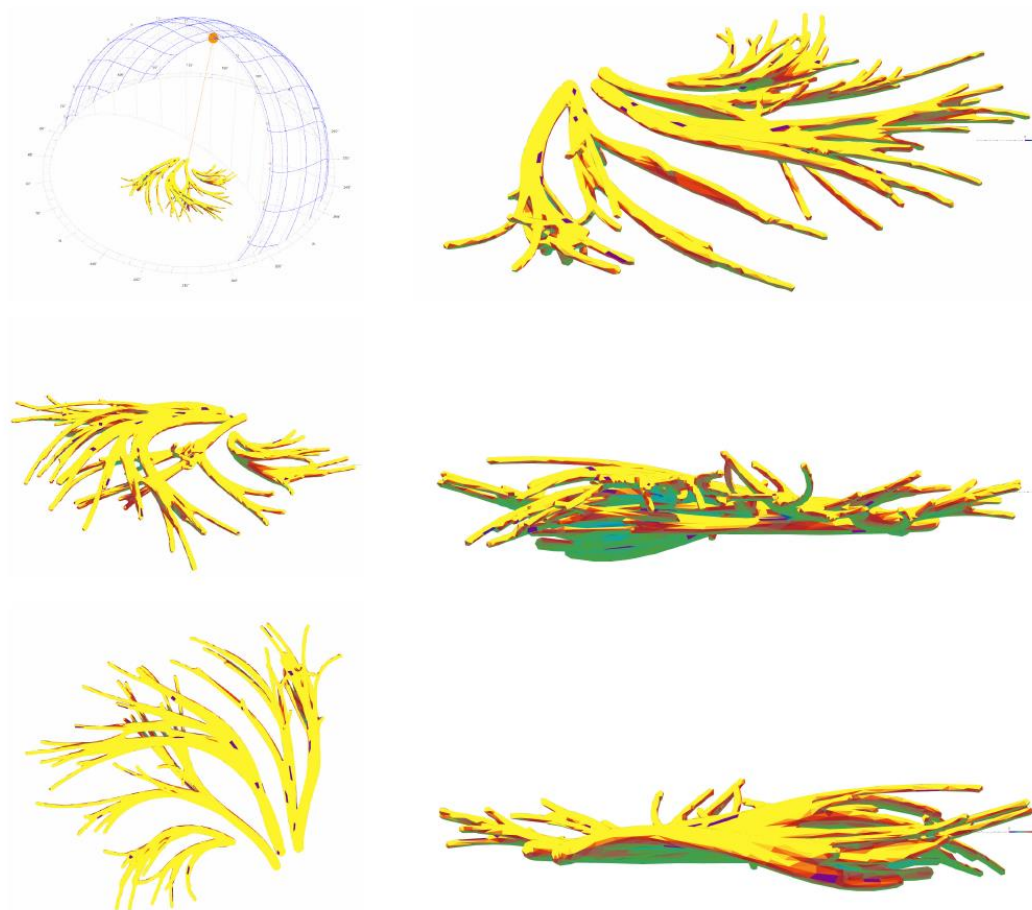


Figure 4. Phase 3: Solar Radiation, Temperature Simulation and Water level diagrams on a selected attribute for performance review

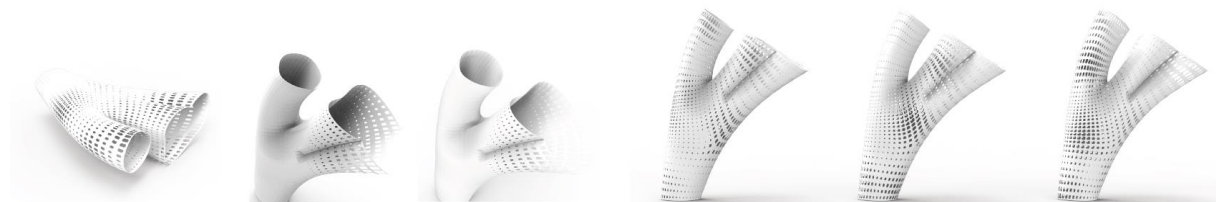


Figure 5. Phase 3: Porosity study following radiation diagrams to amplify the internal performance.

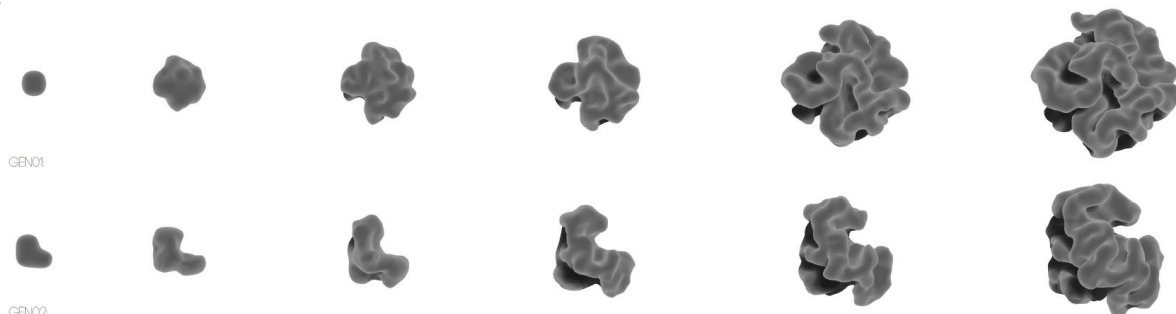


Figure 6. Phase 4: Simulation of growth patterns in Corals.

Following the generated new shape that was now equipped with openings, it was time to predict the growth patterns that would affect the general shape of the structure in a given time frame. The results were generated using Processing and were then showcased in the final outcome. (Figure 6)



Figure 7. Spatial distribution in the new proposed artificial hub.

Figure 7 clarifies how space is recruited and distributed following openings and temperature exposure along with light and sea slope to help attenuate the direct contact between man and corals and to enclose the endangered species in a shielded zone that is semi-open and efficiently lit and exposed to a continuous current.

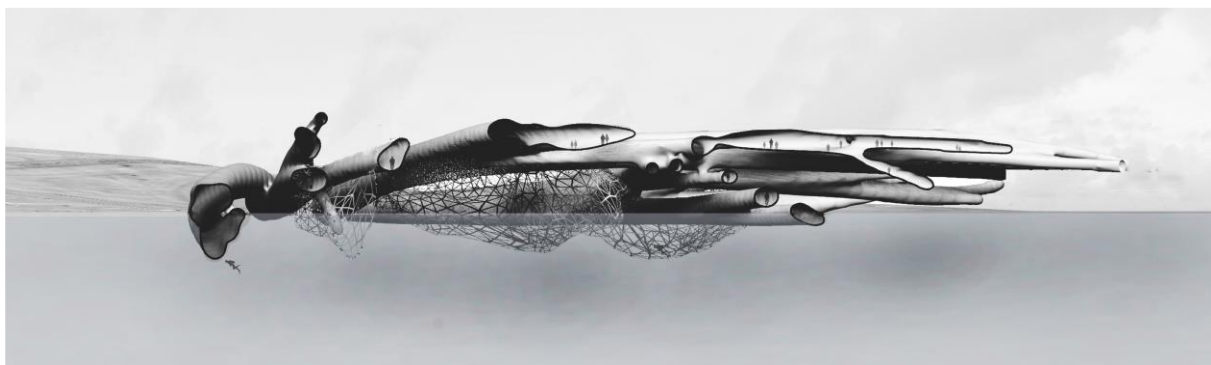


Figure 8. Group 1: Final outcome, cross section.

The final outcome showed in figure 8 emphasizes the importance of porosity natures and openings in the performance of the hub; it also sheds light on the importance of the exchange between the different realms such as the passages between sea and land and the translucent lighting to be diffused to corals through the porous fractal design generated from the experimentation of the founding nucleus.

Lastly, Figure 9 presents an imaginative scene in which the artificial reefs would have been completely developed and grown following the growth patterns studied in figure 6. The once porous links are expected to be colonized by corals once again and to host species that were once menaced in the zone. The figure also displays the user-coral interface in a protected environment that would help corals thrive in an interactive and healthy ambiance.



Figure 9. Final outcome: imaginative scheme of future forms.

This was one of many case studies underwent in the visiting school and was directed primarily to the effect of light, radiation and temperature on coral health and survival. Other proposals revealed ideas and concepts such as semi-submerged touristic complexes, human capsule units that would double as exploration pods and artificial coral cultures, museums and awareness centers, and diving facilities. All these projects and ideas were led by a similar analogy of design based on the recruitment of computational methods and simulation in the generation of the general shape and the forecast of its functioning in a specific given context.

Conclusion

This paper investigated the potential development of the concept of artificial reefs in a way that would raise the expected performance of the now weak attempts of requalification, through the employment of computation method as a promising instrument in this quest. It briefly reviewed the importance of coral reefs as crucial economic poles for neighboring cities and vital architectural habitats for a wide spectrum of species. It then discussed the main threats and the resilience factors that should be listed as criteria for any future artificial reef planning.

In this exploration, computation presents itself as a driver to experimental and unorthodox development that would detoxify and revitalize the fragile Ecotone. Computation methods, so far mostly enrolled in mono-disciplinary domains such as math, civil engineering and electrical engineering, has proven to have a potential role in poly-disciplinary practices as well such as architecture and environmental studies tackling global warming repercussions. This is a new horizon that we think should be further carried out in the near future in order to optimize the efficiency of once challenging and obsolete issues in architecture and complex industrial design. This however calls for the adoption of a new attitude towards computation and experimentation and for the setting out of a detailed plan with optimized management that would bring out the finest result and most effectual solutions for today's problematical issues.

References

- [1] West, J. M., & Salm, R. V. (2003). Resistance and resilience to coral bleaching: implications for coral reef conservation and management. *Conservation Biology*, 17(4), 956-967.
- [2] Bryant, D., Burke, L., McManus, J., & Spalding, M. (1998). Reefs at risk: a map-based indicator of threats to the world's coral reefs.
- [3] Sherman, R. L., Gilliam, D. S., & Spieler, R. E. (2002). Artificial reef design: void space, complexity, and attractants. *ICES Journal of Marine Science*, 59(suppl), S196-S200.
- [4] Baine, M. (2001). Artificial reefs: a review of their design, application, management and performance. *Ocean & Coastal Management*, 44(3-4), 241-259
- [5] Cekić, N. (2018). DETOX TECHNOLOGY IN NEW ARCHITECTURE BUILDINGS. *CONTEMPORARY MATERIALS*, 1(8), 59-66.
- [6] Carrara, G., Kalay, Y. E., & Novembri, G. (1994). Knowledge-based computational support for architectural design.
- [7] ALVAREZ-FILIP, L., Côté, I. M., Gill, J. A., Watkinson, A. R., & Dulvy, N. K. (2011). Region-wide temporal and spatial variation in Caribbean reef architecture: is coral cover the whole story?. *Global Change Biology*, 17(7), 2470-2477.
- [8] Turrin, M., Von Buelow, P., & Stouffs, R. (2011). Design explorations of performance driven geometry in architectural design using parametric modeling and genetic algorithms. *Advanced Engineering Informatics*, 25(4), 656-675.
- [9] Knippers, J., & Speck, T. (2012). Design and construction principles in nature and architecture. *Bioinspiration & biomimetics*, 7(1), 015002.
- [10] Caprani, M., Khatib, M., Mustapha, S. (2016), Morjan, collaborative work during the proceedings of *Hyperbolic Reefs*, AA Visiting School, Amman, Jordan.

The topological principles in the contemporary architectural design process

Vladimir Lojanica¹, †Maja Dragisic¹

¹Department of Architecture, Faculty of Architecture, University of Belgrade, Serbia

†Corresponding author: majchilo@gmail.com

Abstract

Continuing a chronological line of mutual influences of architecture and geometry, where geometry is perceived as an inextricable part of the syntax of architectural space, this paper focuses on the clarification of a specific position which mathematical topology takes within contemporary architectural discourse. The understanding of topology within architectural design process is based on the mathematical theoretical framework in which the term of continuous deformation of geometric shapes is specified, whose subsequent occurrence in architectural creative work is linked to the increasing use of digital tools in design process and the shift of the dominant philosophical influences in architectural theoretical research. In order to completely perceive the topological method, the theoretical framework ranges between the area of architectural theory of form and architectural design theory, firstly through the explanation of three basic design principles of topological method: *deformability*, *openness* and *continuity*, and secondly through the representation of the models through which the principles occur in the architectural design process. The first part of this work will introduce and analyse the transition of concepts of deformability, openness and continuity, from mathematical topology through philosophy to architecture emphasizing the computational shift in architectural design, while the second part of the work will explain the modalities through which the principles are applied in several architectural design practices. Generally, the paper is conducted in order to determine whether the development of the topological method, as a creative tendency, resulted in forming a unique design strategy due to transformations and adaptations through some authorial design approaches. The topological method design strategy, which involves a complete design approach, is identified as a result of an in-depth research of distinguished methods through three case studies, taking into consideration the complexity of topology within the mathematical area and a complex transition towards the area of architectural theory. The final question returns to the primarily theoretical framework, seeking to set operating platform for development and use of three strategic principles, which simultaneously indicate the possible directions of future development.

Keywords: topology, deformation, continuity, openness, digital tools, design theory, design methodology

Transition of topology from mathematics to architecture

In current theoretical studies of architecture, there are numerous references to a branch of mathematics i.e. notably higher geometry, which is called topology, but it is difficult to detect more precise and detailed elaboration of the analysed in-depth and somewhat hidden properties of geometrical objects that topology is generated for. The analysis of the term topology points out the problem of ambiguity, which occurs due to imprecise and frequently loose interpretations of terms which belong to the field of the exact science disciplines. In the widest sense of the word, one can say that mathematical topology does not make distinction between two shapes or two spaces, if it is possible to shift from one to another under continuous deformation. When it comes to these spaces, size and shape are irrelevant if they

can be changed by, for instance, stretching. The difference between two spaces is primarily related to those components which remain unchanged when deformation occurs. The relevant literature in the field of mathematical topology explains that, generally speaking, topology studies the properties of geometrical objects which remain preserved under continuous deformations, such as connectedness or compactness. Geometrical objects that topology studies are usually manifold, but set theory enabled the studying of both general and abstract objects, the so-called topological spaces. Some of the typical examples of topological spaces are Möbius strip, Klein bottle, tori, different knots, etc. In the outline of the history of mathematics, Morris Kline indicates that the first ideas about topology can be found in the works by Gottfried Wilhelm Leibniz, in his book “*Characteristica geometrica*” from 1679, in which Leibniz introduced the concept of *Analysis situs* (Analysis of position) to counter size and form, highlighting the lack of adequate language when talking about form [1]. Also, in a letter addressed to Christiaan Huygens, Leibniz accentuated that we need “another, strictly geometrical analysis which can directly express *situm* /position/ in the way algebra expresses the Latin *magnitude* /magnitude/” [2]. The first precise setting of topological spaces was conducted by Leonard Euler in the period around 1736. In an attempt to solve the problem of *The Seven Bridges of Königsberg*¹ he made the first topological diagram. What is essential for understanding the problem which Euler reduced to the diagram is the cognition that, regardless of the quantitative characteristics of the diagram, the shown topological structures, as well as a solution to the problem given remains the same. By changing the approach Euler has predominantly pointed out to the nature of the problem, placing it in the field of autonomous, qualitative properties of geometric shape, ones that remain unchanged under certain conditions. Euler explains this as follows: “The branch of geometry that deals with magnitudes has been zealously studied throughout the past, but there is another branch that has been almost unknown up to now; Leibniz spoke of it first, calling it the “geometry of position” (*geometria situs*). This branch of geometry deals with relations dependent on position alone, and investigates the properties of position; it does not take magnitudes into consideration, nor does it involve calculation with quantities” [3]. Sergei Petrovich Novikov underlined that it was even intuitively clear that the cognition of geometric properties of shapes was not exhausted by data on their metrical characteristics, such as length, height, angles etc, i.e. “there is something more beyond the limits of the old geometry” [4]. Regardless of length, a line can be open, closed, knotted, several lines can be linked in different manners, shapes can contain holes etc. The characteristic of these and similar properties of geometric shapes, as well as of different mathematical objects that do not have geometric realisations, is that they do not change upon continuous deformations. The invention of precise calculus i.e. the part of mathematics with its exact terms, methods and formulas describing topological properties lasted for a long time. Throughout the 19th century, it was developed, among others, by Karl Friedrich Gauss and Bernhard Riemann, but it is deemed that topology, as an autonomous branch of mathematics, was established at the end of the 19th century by Henri Poincaré. During the following decades, its internal tasks were being resolved and only in the 1970’s did the topological methods more intensively infiltrate into the apparatuses of contemporary physics and chemistry and they were more generally interpreted through discourses of social sciences and humanities, particularly through philosophy and therefrom spreading the influence to different branches of art.

By analysing the transition of topology from mathematics to architecture, one can detect certain influences which result in its more intense presence in architectural discourse around 1990’s. There are two streams of influence, the first one being streamlined through philosophical discourse in specific methods and through work of certain authors, and the other

¹ Königsberg is the name of a former city in Prussia, now Kaliningrad in Russia.

stream being reflected in the change of tools used in the process of architectural design induced by emergence of digital tools and intensive development of software for drawing and modelling. One of the most significant influences on adoption of mathematical terms and concepts, notably those in the field of topology, was realized in the 20th century through philosophical theory of Gilles Deleuze. Taking distance from the predominant thought of the period, where language became the fundamental problem of philosophy, Deleuze insisted on philosophical creativity which enables the formulation of new concepts instead of exclusively describing the existing appearances and states. Basing his philosophical theory on creation of concepts through experimental thinking, Deleuze stressed that there were no simple concepts but instead that they were complex, multi-layer structures, figures, metaphors, individual elements etc. [5]. His overall approach to philosophy defines him as a more progressive materialist, who based his materialism on science and its discoveries and does not observe matter exclusively as essence but also addresses its genesis and the genesis of its form. Matter does not have an inert but rather an active character, and its form is shaped primarily by generic processes, which results in concepts that merge scientific knowledge with philosophy. The link of philosophy with scientific knowledge, primarily with that of mathematics and mathematical topology, gave a fundamentally spatial character to Deleuze's numerous philosophical concepts, and thus he defines the differences between continuity and discontinuity, smoothness and folding, topological and metrical, large and small, stable and nomadic etc. Through his philosophical materialism which relied on mathematical terms and their interpretation, Deleuze made topological concepts accessible to public. But Deleuze's contribution to topology was somewhat greater than mere interpretation of mathematical discoveries. He applied topological discourse to his philosophical concepts falling within the domain of philosophy, such as the issues of ontology and the nature of being, metaphysics etc, and thus he gave additional meaning to classical philosophical terms, attributing them the properties such as continuity, deformability, curvature, smoothness, folding, bending etc. The impact of his work thus became important for theoretical discourses apart from philosophy, notably for architecture, since he used dominant spatial characteristics to interpret the issues of individuals, societies, relations within social groups.

One can observe that the methodology of applying mathematical concepts to wider scope of knowledge frequently relies on specific knowledge that define different areas, which are defined by Arkady Plotnitsky by reciprocity of mutual influence of mathematics and philosophy known as "quasi-mathematics" [6]. Although he does not question philosophical influence of mathematics on the development of civilization, he claims that quasi-mathematics enables the spreading of certain mathematical terms and principles which are not defined exclusively by mathematical tools, although deriving therefrom, and therefore they become feasible and applicable beyond its disciplinary margins. Through term quasi-mathematics, Plotnitsky explains the difference in interpreting algebra, geometry and topology in general. He interprets algebra as an ultimate concept of formalisation, whether formalising a system in sciences, conceptual systems like those in logic or philosophy, or language system existing in linguistics. In this manner, "algebra" means a set of certain formal elements and their relations. On the other hand, "geometry" and "topology" have different mathematical backgrounds although they both deal with the issues of space. "Geometry" deals with space measuring as *geo-metry*, whereas "topology" disregards sizes and deals exclusively with the structure of space (*topos*) and the essence of a shape. Putting them in a philosophical discourse, Plotnitsky explains the difference between these two theoretical aspects with Derrida's "algebra", which referred to writing, characters, and form dislocated in negation, and Deleuze's "topology", through which he insists on the continuity of folding.

Referring to the previously given elaborations, one can conclude that topology was difficult, incomprehensible and entirely abstract for architecture, and that it emphasised certain differences in mathematical and architectural perception of space. On the one hand, mathematics brings abstraction to its extreme, which exceeds architectural perception of spatial relations. On the other hand, the methodology used in mathematics for solving its internal tasks is exceptionally precise and exact, which is not characteristic for the process of architectural design. It appeared that philosophical texts, which were already highly positioned in the theory of architecture, managed to overcome this discrepancy between architecture and mathematics by interpreting certain mathematical terms using language that was much comprehensible for architectural discourse.

Simultaneously with the change of philosophical influences, the presence of topology in architectural discourse was also registered in the change of working tools used in the process of architectural design. Digital tools and the development of modelling software changed the position of classical drawing where space was displayed through projections during the design process. Even greater influence is resulting from the knowledge of software and their intensive upgrading, which introduces algorithm logic for design problem solving into architectural discourse. Computer software had an option to generate geometry of topological characteristics, not only by means of equation, but also through parametric functions that provided numerous variants for continuous curves. Already in mid-90's, the computers with software for modelling the desired curves became affordable because their price was drastically decreasing. However, in the context of this paper, there is a more important thesis that states that digitalisation in architecture implies a more drastic progress towards a new architectural paradigm i.e. a new way of thinking where use of digital technology does not only imply the use of digital tool, but also the theory of algorithm as the main creative postulate, way of thinking, special thought and creative form. In early 1990's, Peter Eisenman introduced a new term: "During the fifty years since the Second World War, a paradigm shift has taken place that should have profoundly affected architecture: this is the shift from the mechanical paradigm to the electronic one" [7]. During the nineties, theoretical papers in the field of architecture and the related discussions began to see the positions that digital principles started to transform the paradigmatic framework and that they were growing from technological fascination into the way of thinking. Word has it that algorithmically-generated space indicated fundamental, ontological change of basic elements of architecture and that the appearance of digital tools and the specific logic for their use in architectural theory and practice became a reality. The Deleuze's philosophy in the theory of architecture definitely appeared at the moment when digital technology was already well developed. At the same time, this is a basis for debating whether Deleuze's philosophical platform found a tool for its realisation in digital means i.e. whether it would have such an impact on architecture if there were no technological prerequisites for its visualisation. Anyhow, the presented comparative analysis of the impact of philosophical thought on the one side and the development of computer tools for modelling on the other, proves that the connection between Deleuze's philosophical theory and digital tools in architectural discourse is undeniable. By joint action – that of Deleuze through philosophical terms based on mathematical topology and that of digital technologies that enabled the manifestation of certain abstract mathematical concepts expressed exclusively by calculus – the idea of topological tendencies in architecture is actualized.

In the context of in-depth elaboration of different modes of use of topological principles, one can understand why certain historical overviews contain the term "topological architecture". The clarity of visual expression, which was present at the very beginning of use of topological principles, led to the denial of claim that certain fields of art must first decide how to present their final product in relation to the process of its generation. It appeared that quite the

opposite was in the case of architecture – the form of final product was known, with increasingly clear picture of the possibilities for its realisation through faster development of technological means and applied materials, but once the manifestation became clear its actual meaning came into question.

Three topological methodology principles

The discussion on positions of different scientific disciplines in architectural discourse tells us that certain parallels can also be made with topology, primarily in the context of relation between architecture and sciences. When it comes to methodologies of architectural design, it is clear they can be different, but they usually do not imply the exactness and fixed language for solution of individual problems such as other scientific disciplines. Therefore, for the sake of more efficient link between contemplation and creation leading to ultimate result – the work of architecture, architecture freely adapts specific methodologies of other disciplines. As regards topology, it is clear that in architectural discourse it cannot be formally considered as mathematical topology. Adjustment of knowledge in topology for the purpose of forming topological principles in architecture is explained by philosopher and architectural theorist Manuel De Landa through term “topological thinking”, based on the idea of research of system potentials and the manner in which the potentials may generate certain forms, whereas he treats form as a system of elements with capacity to influence other system elements [8]. Relying on De Landa’s positions, in the upcoming text we will use three topological principles - continuity, openness and deformability - to explain transition and transformation of topological properties, from mathematical definitions to segments of individual project methodologies, and to explain their potential for creation of a wider design platform.

Principle of continuity

Generally speaking, the main idea that defines and specifies mathematical topology is the idea of continuity, which in topology primarily refers to continuity of mapping. Continuous mapping can be explained by the idea that “close” points of one set are transferred to the “close” points of the other set. Intuitive explanation of continuity implies that, upon mapping of figure *A* into figure *B* there are no sudden rises, hence upon “slight” changes of the original its picture is also “slightly” changed. The term homoeomorphic mapping that can be found in architectural theory texts comes as a more precise definition of mathematical topology, and it can be perceived as mapping of one set of elements into another, without tearing or subsequent gluing together. If we presume that it is possible for figures *A* and *B* to be stretched and bent so that we bring figure *A* to translate to figure *B*, we can generally say that they are homoeomorphic. For instance, the perimeter of triangle is homoeomorphic to a circle, the surface of sphere is homoeomorphic to the surface of cube or cylinder and it is not homoeomorphic to torus etc. Also, line segment can not only be stretched and shrunk, but also bent and straightened.

The principle of smooth continuous stretching contains deep spatial references and, interpreted by the continuity of architectural space, it demonstrates a necessary degree of flexibility of spatial framework. It can be interpreted through continuous circulations, implying that the architectural structure has continuous trace of movement and continuous flows of different information. With regard to the type of the observed trajectory, it is possible to treat continuity as a spatial characteristic that includes and spreads within an architectural structure, or more narrowly observed as a continuous planar communication visible at the architectural plan. Continuity of architectural structure reflected through superficial continuous movements is conditioned by predominantly organisational solutions, whereas spatial continuous movements can be achieved by the continuous void within architectural

structure. The principle of continuity of spatial voids is closely linked with interpretation of and linkage with the principle of free plan, since both of them rely on acceptance of basic architectural postulates, as defined back in the modernism.

The analysis of this specific principle tells us that it comprises of two terms that need to be elaborated: *continuity*, which is closely linked with the mentioned principle of modernism and which can be partly interpreted through forms of movement within space, and *void*, the manipulation whereof can be used to define the structure of the work of architecture. Specific continuity of inner space in terms of volume relates to a more general perception of continuous flows including, beside movement of users, visual, information and other spatial circulations. However, the issue of spatial articulation, empty space within certain form, represents one of the key issues of architecture that can be interpreted both as a relation and as mutual action of internal and external space. The origin of these contemplations dates back in the 19th century, when space i.e. void had a sort of metaphysical significance, but the overall methodological basis was developed by Raumplan concept at the beginning of the 20th century through a complex system of interior development by Adolf Loos. Although connected with the development of open plan principle, Raumplan builds on Loos' design methodology based on the idea of designing space instead of plan. The basic idea of interior space segmentation is achieved by dividing different floors in to the several levels, so that continuous space spreads within a building. Although this system is close to Le Corbusier's methodology "from the inside out", Loos' perception of space incites a volume-based modality of creative thinking, relying on enclosing skin as a structural element. Charles Jencks indicated the existence of another approach to modern space, based on the tradition of rational Chicago school, and Le Corbusier further developed it through structural skeleton of Maison Domino, where all future principles of modern architecture can be observed [9]. Space is here perceived as being homogenous in every direction, but segmented as skeleton at right angle to the façade plane. Although such interior space is characterised by vast and open structures, it is nevertheless limited by edges, the enclosing skin is clear, ultimately rational and feasible.

It was only with contemporary definitions that the terms of *continuity* and *void* were brought into clear connection for the purpose of explaining topological properties of architectural works. Methodological postulates of architectural theorists such as Greg Lynn, Lars Spuybroek et al. clearly rely on previous researches, they underline the significance of Raumplan concept and take a distance from Le Corbusier's open plan, whilst topological continuity of space is described by emphasising the potential towards more flexible connection between internal elements. The logic of fluidity supports the thesis on intensive mobility, which implies vast and easy deformability because the interior structure is such that small forces lead to large deformation. The issue of continuous interior discussed herein is maybe most relevantly referring to the logic of continuity elaborated by Spuybroek in his book *The Architecture of continuity*. In the first place, he writes that buildings are made of elements does not mean that architecture should be based on elementarism; we should rather strive for an architecture of continuity that fuses tectonics with experience, abstraction with empathy and matter with expressivity [10]. Spuybroek recalls the logic of continuity that is philosophically developed by Charles Sanders Pierce on basis of topology explanation [11]. Elaborating Pierce's notion "structure of vagueness", Spuybroek explains that the relation between elements is always vague since they are at the same time elements and parts of a whole. Vagueness does not represent the absence of logic, on the contrary, the logic of vagueness is what constitutes the relations. In this manner, using Pierce's conclusions, Spuybroek establishes equivalence between continuity and relations, strictly opposing the idea that architectural space can be classified as space and void. He thus concludes that void needs to be interpreted as a spatial structure and not as air between the walls. Thus observed,

an architectural work implies that emptiness has a temporal aspect as well, since it can change iterations over time. The definition of this type is significant in the domain of continuity research, since it indicates that it is possible to create a specific design methodology which defines space through manipulation of the absence of space. As it is completely unnatural in built structures to conduct construction and subsequently take out its parts, the design methodology building on emptiness implies that the most significant spaces within buildings are created either by elimination or omission in the phase of architectural concept design. Strategically, such design methodology treats void as an integral part of spatial complex and builds upon it in the design process.

In her experimental work, Dagmar Richter used the model of Le Corbusier's Maison Domino as a subject of research and a mechanism for verifying her own hypothesis that topological principle of continuity can be largely defined through changed attitude to *surface*, which becomes the ultimate structural concept [12]. Richter tests her explicitly topological position starting from a paradigmatic definition of five architectural principles, where the connections between reinforced concrete column and beam bear the ceilings on six offset points and represent the only load-bearing elements of architectural assembly. The research process implies that the structure of Maison Domino is analysed by layers, where the first layer in the structural hierarchy – slab – is treated as a series of fields with local characteristics in points of bearing, which breaks up the primary form of universal form skeleton. During the first phase of research, the skeleton evolves, during the second phase it develops spatially under the influence of spatial connectivity principle, non-hierarchical spatial relations and adaptability, all for the purpose of controlling the variable nature of the newly-designed prototype. By application of animation technique, Richter's research team developed several simulations of the process, creating a collection of possible prototypes which typologically still refer to housing. After the second phase, the change of the structural skeleton is visible, as it is now constructed from a single continuous plane, the mass whereof becomes an important factor for construction of the new model of topological structure (Figure 1).

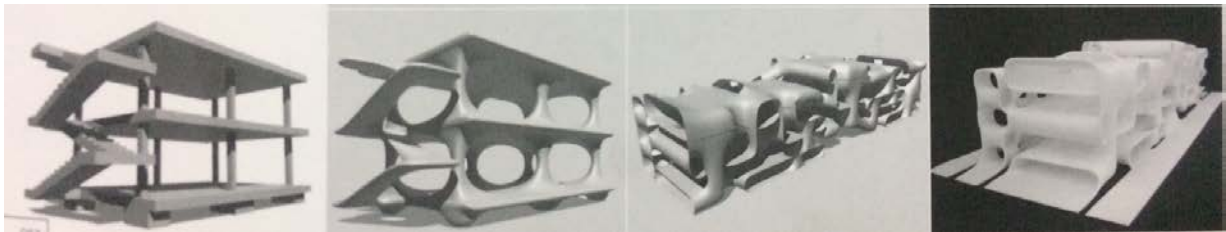


Figure 1. Dagmar Richter, new structural role of surface - transformation of Le Corbusier's Maison Domino model

The aspect of deconstruction of façade shell, which drastically differs from Le Corbusier's free façade, can be interpreted as a prototype of essential intertwining of interior and exterior space of the house. An important conclusion of Dagmar Richter's research is that topological deconstruction of an architectural work can only be achieved by changing the structural system i.e. its major transformation relative to the existing assemblies. Throughout the project research, she demonstrates how this can be realised, proving her own assumption that the treatment of space, relying on topological principles during the design of architectural works, is founded on the new structural role of *surface*.

The establishment of the role of topological principles in the design and research done by Dagmar Richter can be monitored through all phases of architectural work design. In addition to structural system deformation and creation of a new type of topological architectural structure, Richter directly refers to topological principles in her theoretical essays, as well as

to mathematical discipline. Broadly observed, the entire research process, which starts from the paradigm of modernism, additionally refers to post-modernistic references, where a new structure can never be clear and original in a classical sense, but it is always an interpretation. Thus perceived, the design methodology is defined as a strategy for adapting previous systems to new conditions i.e. rather as a process of reinterpretation than that of creation. The design methodology of Dagmar Richter directly relies on chronological references to the notion of space continuity, including the questions of what was before and what would follow, as a constant historical postulate, that the science has been striving to overturn since the beginning of the 20th century. Cyclical history, fundamentally different as compared with the evolutionistic, is a history of one multi-dimensional space where texts and writing overlap, making a single network of “diversity” of events. Each text has a reading history because different societies re-write the text by reading it and unconsciously attribute it with different meanings. No one can read a text without making an image of the context in which it was made and what its actual aim is.

The principle of openness

Generally speaking, the notion of limits of space is related to the notion of distance between two elements, which in mathematical sense implies that elements can be attributed real number and declare its distance between them. Metrical space reflects natural ideas about the notion of distance, relying on understanding of spatial relations where i.e. distance between two points is always positive – distance between x_1 and x_2 is always equal with the distance between x_2 and x_1 etc., which simultaneously most corresponds to the perception of Euclidean space. The notion of openness becomes very important for the topology induced by metrical space, because mathematical definition of surface relies on understanding of the surrounding of points that it comprises of, so that surfaces with and without boundaries are being distinguished. Surfaces with boundaries are, for instance, circle or sphere where several openings have been cut out, while surfaces without boundaries are generally classified according to the number of holes they contain.

Contemplation on the notion of boundary in architectural discourse can be extremely broad, but in the context of the postulate of topological principle of openness, it primarily implies the research of properties of space at the limits of structure, where topological character is displayed in its imprecision. As originally explained in topology, the focus on *surroundings* of a point can be perceived as a small shift in the surroundings so that it is never even abandoned, which complicates the relation inside-out and the question of defining the limits and bordering areas relating to the observed point surroundings. Openness thus starts to refer to the property of architectural work which relativizes the treatment of exterior and interior of architectural structure and is manifested through the weakened attitude to the object limits. The characteristic of this type implies the research of classical spatial duality interior/exterior, by methodologies where architectural work is designed by intertwining structure with imminent surroundings. Research can be done from two aspects, by analysing façade plane and ground floor of the architectural work, in order to clarify possible approaches relative to the shell and to the ground. Transition of openness term from mathematics towards architecture is largely founded on Deleuze's interpretations, both through specific spatial indications and through more direct elaborations of philosophical postulates. Architectural discourse of the 1990's recognised the importance of interpretations of Deleuze's positions stating that exterior is not a fixed limitation but rather a mobile matter, animated by movement, folds and bending, which defines the interior: it is nothing else but exterior, or more precisely the interior of the exterior [13]. Research of the architectural work boundaries demonstrates that folding goes from outside to the inside and vice versa, through different scales and regardless of distances, where nothing is fixed but rather in constant change.

Architectural treatment of façade through theoretical assumptions of the 20th century, formed within the two previously noted and somewhat opposed architectural principles: the principle of structure and the principle of membrane. Design methodologies of architects such as Walter Gropius and Le Corbusier affirm the idea of structural façade plane which is based on the plan of structural columns (*pilotis*) by means of which façade can be treated apart from the spatial assembly of objects, which is fully described by Le Corbusier's fifth postulate of architecture. The structure principle enables the façade plane to exist irrelevant of the interior space organisation, which provides it with certain autonomy in display of exterior/interior relation, so that the limit can be treated both as open and closed, depending on different external parameters. On the other hand, Bernard Cache explains that, as opposed to the principle of structure the principle of membrane is being developed, which is particularly affirmed in theoretical platform of Adolf Loos where façade is an element conditioned by internal space organisation and the membrane indicates the differences between internal spaces. This gradual movement toward interior can also be understood as an attempt to create an extrovert architectural work. Although the principle of membrane in Loos's realised works is displayed through organisation of façade openings, a significant deviation in façade treatment is represented by the attempt of its deconstruction, which leads to the nearing of interior and exterior of architectural object. Italian architectural theorist Alicia Imperiale explains that modernistic design methodologies generated in the first half of the 20th century strived to present a tension between deep interior space and façade surface, using glass or similar transparent façade planes, which actually indicated a dialectic difference between interior and exterior [14]. Certain contemporary approaches use different techniques for processing façade plane in order to allude to this principle by means of delayering the surface with specific materials or patterns with different transparency. However, Imperiale deems that the openness principle can only be achieved by substantial merging of the interior and exterior, inseparably from the continuity principle, and that the relation between these two dualities can only be significantly changed by continuous plane treatment. If the weakening of limit in the façade plane is not restricted exclusively to the impacts that the membrane receives from the inside and/or outside along its surface, but it is rather interpreted through more intensive inside-outside relation, the perception of membrane surpasses the plane structure and the limit along which the membrane extends becomes a zone on the inside-outside crossing. Blending of interior and exterior is in this case slowly shifting, according to Loos' attitude to structural shell, which in its final form can become the entire structure. In theoretical postulates where façade is comprehended as a certain area, where width and position of the area fluctuate to the outside or to the inside, one can record significant use of topological surfaces such as Möbius strip, Klein bottle etc. Möbius strip is a surface with one side and only one boundary, but its direction covers both the inside and the outside through orientation reversal. Although it is clear that the system of Möbius strip and notably that of Klein bottle do not have three-dimensional realisations, they cannot be directly applied, in the context of architectural design the non-orientability of surface indicates membrane treatment where certain bending can provide for continuity of interior and exterior. If the notion of non-orientability in architectural discourse starts to refer to Deleuze's notion of fold, it can also be interpreted as a process of continuous and homogenous transformation that manages to preserve integrity, continuity and uniformity of parts. Intuitively, this process can be perceived as bending of surface, smoothly and without pulling, whereas after several variations the limit between the outside and the inside and between full or empty would disappear. Therefore, topological property of openness can be estimated on basis of recognition of the open façade interspace which displays the causative relation that affects its deformation. The weakening of opposites outside/inside does not imply the vanishing of architectural structure, but it requires theoretical analysis of elements affecting the limit

behaviour. Through such postulates, the space between interior and exterior is treated as space between, as defined by Andrew Benjamin when he says that the space of difference is not just “between” but the interspace of the boundary becomes a third segment that cannot be predefined and it is directly comprised of the inside and outside elements that generate it [15]. Direct application of openness principle can be analysed on the design for Eyebeam museum, the award-winning work of the architecture studio Diller Scofidio + Renfro at the international competition realised in 2002 by Eyebeam foundation, which is an art organisation oriented towards research of technology impact on the development of different art practices. The call and programme definition of the competition required a facility that combined the purposes of museum, theatre, education and production, wherefore the museum part was intended for exhibition of modern art works generated under the influence of new media, in form of performance, video works, 2D and 3D digital imaging, sound installation etc. The second part implied laboratory for art production, as a requirement of contemporary art institutions, for the purpose of displaying the works of art whilst transparently presenting the process of their generation.

The conceptual presentation of Diller Scofidio + Renfro is based on the use of pliable strip the disposition of which separates two museum segments: presentation and production. The strip starts to bend from the ground floor level, it extends along the entire building width so as to form a continuous plane of the floor, wall and ceiling. Each bending and change of direction opens either presentation or production zone, which additionally combines the movements of visitors and staff. Activities within the space can best be seen at the section that displays strip bending and intertwining of interior and exterior space, which is entirely transposed to the façade plane. The approach where exterior of a building is conditioned by the internal organisation, which is displayed on the façade, represents a significant departure from the traditionally closed facades of museum buildings. Spatial relations are becoming even more complex in the parts where the strip splits towards upper and lower level, which leads to additional overlapping of presentation and production spaces and announces additional thematic dualism. One can say that the concept of pliable strip permanently indicates the existence of different binary pairs within the designed space, but with the aim of their essential overlapping instead of distancing. Duality is particularly emphasised by the two-ply treatment of the very strip, which comprises of a smooth concrete ply with precast service jacks or “smart pores” line exhibition levels and a second ply of lightweight removable panels of non-conductive composite material line production/education levels. The interstitial space is an installation base running through the interior and exterior of the building so that the open structure of the strip enables subsequent additions to the necessary installations, which was described by Diller Scofidio + Renfro as the building’s “nervous system” [16]. The building orientability in terms of outside/inside relation is additionally complicated by double-layer treatment of plane, where specific colours underline the change of position (Figure 2).

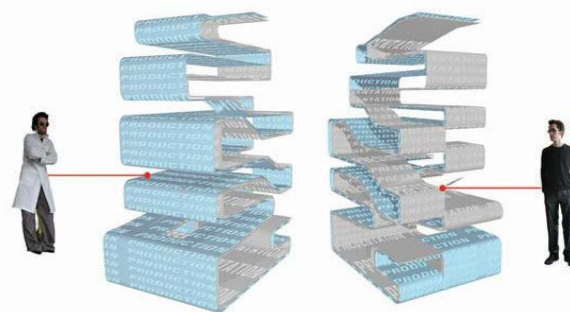


Figure 2. Diller Scofidio + Renfro, double-layered folding plane for Eyebeam museum

The presented building, generated by bending of surface that twists freely under different angles, is now indicated as paradigmatic example of architectural theoretical standpoints based on Deleuze's theory of folding. Although Diller Scofidio + Renfro do not elaborate the design in this context, other segments of their work, particularly the initial period, indicate direct influence of French post-structuralism philosophers. Direct application of topology in their work is confirmed along with the position of Deleuze's philosophical platform, so that different topological properties can be observed in the continuous and smooth deformation of plane. Generally speaking, openness that relates to the weakening of limits is reflected in specific vertical plane deformation, which results in weakening of boundaries at the positions opposite to bending.

The principle of deformability

Although the previous presentation of topology might lead to the conclusion that in the essence of homoeomorphic translation contains the deformability of a certain figure in order for it to translate to another figure, the introduction of deformation notion in topology primarily implies the deformability of the translation itself. More precisely, the term deformation is used in order to describe the relation between two continuous functions and not between two figures. Off the record, the function parameter can only be perceived as *time*, whereas the notion of time in this case does not imply any of the precise mathematical definitions, but it underlines that the process of deformation primarily displays over a certain interval and the display of change happening within such interval. It appears that the relation to form change temporality in architectural works can be built and displayed by application of the same principles. On the other hand, in architectural design methodology deformability usually implies the analysis of potentials of spatial structure for deformation, but not as a property which explicitly regulates formal rules for architectural buildings. Deformability implies that an architectural work was created by application of design methodologies where elements build structure by examination of internal relations instead of according to pre-defined order principles. Generally speaking, deformability property is preserved through a specific space building logic and it can be analysed on basis of two key issues: *what* has the potential for transformation of the main form, which is known and metrically defined, and *how* the change occurs. The first issue indicates the object of deformation i.e. specific characteristics of architectural work and it pertains to the research of deformable potentials of the existing elements. The second issue focuses on the display of the change of form of architectural structure i.e. the process that deforms it. Accordingly, the deformability of architectural structure speaks about the potential for topological logic of construction to be manifested in the ultimate form of architectural work.

Anas Alfaris deems that the form is a set of elements and their structure, where elements represent parts of a whole, and structure regulates formal relations between such elements [17] Referring to this definition, it is possible through the history of architecture to follow the development of the principle of architectural form creation *from general to individual*, which is rooted way back in the architecture of ancient Greeks and Romans, where the system of proportions was based on the golden ration, symmetry and examination of ideal relations between parts and the whole. The first examples of use of *individual to general* system in form development can be found way back in the Islam architecture, where mathematical formulas were used for repetition of geometric figures to obtain a complex pattern surface. Generally speaking, both systems of form development – from general to individual and from individual to general – rely on the idealised perception of the world. Geometrical systems of proportion, perspective, typology, geometrization of tiles in Islam architecture etc. are *a priori* based on ideal principles which either multiply elements or separate the whole, but the internal principles of the given system remain unchanged. Considering that both parts and the

whole are displayed as fixed, unchangeable geometric forms, they cannot be combined so as to make any impact on or to modify one another. The presented systems of composition of structural assemblies neglect the character of relation between elements and such relation remains quite simplified, even in the treatment of structure as developed by Metabolists or members of Archigram group. Against these two systems, Farshid Moussavi's book *The function of form* suggests a different system for construction of architectural functionally-formal assemblies which she called *transversal system* [18]. The system implied that the main constructive element is not geometrically pre-defined but that it comprises of multiple causatively-complex systems. The principle of element combination arises from their specific nature, which causes form to be generated from proto-geometrical characteristics that are physically and geometrically specific but are not necessarily specified. The elaboration of proto-geometrical characteristics of basic constructive elements is close to definition of topological invariants, and, as explained by Moussavi, these are the characteristics with capacity to be the constants in any form they generate. Topological character of deformable spatial assembly essentially implies the same as in mathematical discourse i.e. that architectural work is accessed at the moment of generation as to a system subject to free deformation, the formal display whereof can be freely changed if the elements, which are close in the initial disposition, remain close after the deformation. On the other hand, the limits of deformation and constant provision of architectural part are conditioned by external restrictions and they are never the result of a universal internal rule.

The principle of deformability in creation of an architectural work, including a temporal component in the process, implies the display of all phases of deformation and not only the original figures. The idea of direct display of deformation process is based on the display of change in architectural work in certain time intervals, while the transposition from one form to another is done by small deviations from the previous spatial determination. Visual research can rely on the perception of time continuum through strip element, which also represents a change graphic. This type of display implies that it is possible to read movement or motion through such formed conceptual aspect, whereas this type of minor departure in time definition is only defined when there are no fixed reference points or suggested identity, but only when relations based on uncertainty and certain differences are established, instead of those based on traditional attempts of order and its repetition. During the project procedure, what mattered was transformation of previous step into the following through current state, and in order for each deformation step to remain within the limits of topological method, the elements must not be subject to tearing and subsequent connection. Ultimately, the use of deformation principle can be tested on architectural works by application of standard architectural techniques, such as architectural drawing or three-dimensional picture, since the displayed time interval provided the projected methodologies and it refers exclusively to the design process. Elaborating the kinetic form, Kostas Terzidis added that movement was an act or process which changes position or place over time, hence the movement includes temporal component which actually represents a unit of change [19]. Referring to similar definitions, Greg Lynn indicates a need for systematic inclusion of time and movement upon form definition, identifying cinematic model when movement indication in architecture is concerned [20]. Cinematic model implies the multiplication of static film sequences that simulate movement, and the displayed frames create a memory of form which is spatially and temporally simultaneous. In fact, it generates an idea of architecture that creates temporal component through memory of time. It is based on animation, morphing and similar techniques based on the display of several isolated pictures over a defined time period. In the context of creation of architectural works, Kostas Terzidis explains that morphing is a term used to describe a process where an object changes form to obtain another shape. Although this is a gradual transition, it can result in significant change of appearance, character, state or

function. Morphing is a significant formal means and it refers to one of the most significant matters of architectural objects: possibility to express and identify itself through own form. The interior design of Miran Galerie by architectural studio dECOi is explained by Mark Goulthorpe as a process of membrane generation and the beginning of research of architectural surfaces by Rhinoceros computer programme, for the purpose of creating the impression of morphed three-dimensional shell within a static architectural building. Homogeneity of continuous curved surface that fills the existing spatial frame by unique treatment of floors, walls and ceiling, had been preserved by bending and folding of plane within the spatial limits. A series of analytic diagrams displays how dECOi follow the change of main form through deformation of longitudinal axis, which can be interpreted according to mathematical definition as translation deforming the displayed structure in the interval starting from initial to the final position (Figure 3).

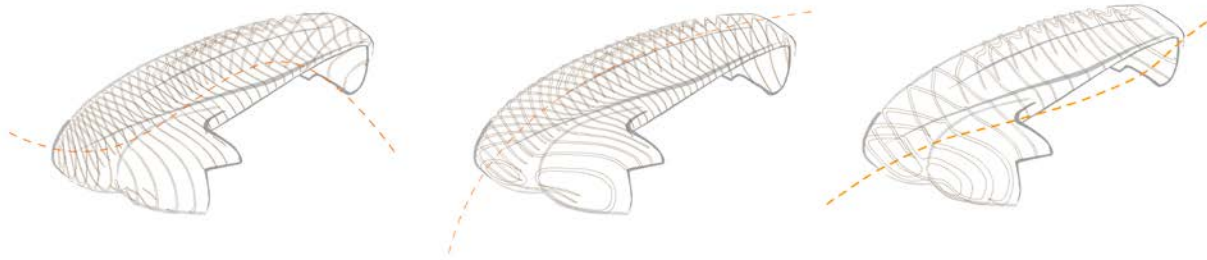


Figure 3. dECOi, series of analytic diagrams showing the process of deformation

The designers' intention to implement deformation process through a series of cross-sections implied that dECOi had to develop a sophisticated technique of computer modelling – in order to present a three-dimension form as a collection of two-dimension elements, to obtain structurally stable frames and to develop encrypted tools generating sections on the desired positions. During the elaboration phase, the problem with main figures in generation of hybrid form was detected. Their manipulation uses two curves of *nurbs* type, where one controls the angle of cutting and the other controls the assumed density of crossing, hence the final figures can be harmonised with the primary volume. An important part of the developed methodology relates to the development of diagram approach to the control of these two lines – where definition of the assisting cross-section and line for density regulation, located outside the building, can be used for testing of variations of temporal axis. The displayed methodology implies changes in the manner of element production, along with the development of special computer programme with clear optimisation parameters. The recognition of use of topological principles in the development and realisation of the internal membrane of Miran Galerie is reflected in the clear intention to present the process of deformation by specific cross-section system. The design methodology, based on the development of programmed tools used to control the complex geometry of the designed membrane, takes into account the change in curvature of the local plane parts. The significance of Goulthorpe's opus is reflected in the intensified use of certain topological principles in project development phase by using certain topological methods to resolve explicit problems in realisation and use of specific materials. As the focus from its initial work was shifted towards the research of new materialisation models and creation of theoretical platform relying on specific computer logic for resolution of architectural details, Goulthorpe retains priority in the realised works as well.

Conclusions

Based on the presented topological principles, illustrated by examples of architectural works where they can be observed, it becomes clear that contemporary architectural paradigm has,

by registering key principles of topological method in design process, acknowledged the presence of topology as an integral part of a wider design strategy. During the review of statement that geometry conditions a part of relations within architectural space but it needs to be integrated in a much more comprehensive theory so that her proper place and significance would be assigned, it becomes clear that topology, as a creative tendency, contains design principles based on which it builds its own framework and which ensure its position in contemporary architectural paradigm. Although the first part of the paper indicated a more complex theoretical position of topology in architectural discourse, it followed on to observe three topological properties and principles of architectural structures, demonstrating that recognising the use of topological principles in designing architectural works can be done on basis of basic architectural postulates. These principles are not new, but in previous design methodologies they were not observed in this manner i.e. so that openness is reflected in a weakened relation to façade shell, continuity in internally empty and full spaces, and deformability is interpreted by the specific principle of building an architectural structure and as a process causing deformation. The presented principles indicated a wider context of application previously elaborated through the idea that architecture and sciences contribute to creation of a specific system of world perception, while it is insisted on flexibility of structure susceptible to easy changes dictated by a moment. If we assume the most general position that architecture, as well as sciences, speaks about the creation of the world which is inhabited by subjects and objects, whereas the “manner” is always historically determined, we can understand the position of Antoine Picon who refers to relation between a subject and its environment, while architecture and sciences are the fields which define this relation, primarily in the domain of the created environment. It can be noted that certain historical processes hold this relation to be more significant and productive, and Picon relates it to the periods when architecture and sciences contribute to creation of a specific system of world perception [21]. The manner of perception and image of the environment refers primarily to the environment as a cultural category.

This paper is an attempt to clearly outline the causative link between digital technologies and complex geometry reflected in topology, but also to emphasise the additional problem arising from the approach of topological treatment of form which refers to the absence of aesthetic valorisation of the deformed amorphous architectural forms. A part of problem also stems from the absence of clearly defined system of evaluation of aesthetic characteristics of new forms, and another part is in the process of their generation which is underlined by question of what is it in the process of form transformation that determines when the form would end. Similar observations are made by Michael Meredith who states that the results of use of topological principles during the 1990's are reflected in isolated physical and aesthetic models that have no wider impact but instead remain within their own limits [22]. In this context, we can contemplate whether literal application of topological principles in architectural design favours the formal-shaped aspect of architectural work, which transfers the matter of use of topological principles from the domain of architectural theory of form closer to the architectural design theory. The three suggested strategic principles of application of topology: principle of continuity, openness and deformability represent certain characteristics of architectural works and indicate methodological processes rather than precise designing recommendations, which makes the defined platform for use of topological principles less determined and looser for interpretation.

References

- [1] Kline, M. (1972) *Mathematical Thought from Ancient to Modern Times*, Vol.1, Oxford University Press, New York, 370-378.
- [2] Kantor, J. M. (2005) A Tale of Bridges: Topology and Architecture, *Nexus Network Journal*, vol. 7, no. 2, Birkhäuser Verlag, Basel.

- [3] Euler, L. (1956) The Seven Bridges of Königsberg, *The World of Mathematics*, Vol. 1, James Newman, ed., Simon and Schuster, New York, 573.
- [4] Novikov, S. P. (1984) Foreword in *Očigledna topologija*, Zavod za udžbenike i nastavka sredstva, Beograd.
- [5] Deleuze, G. (1994) *Difference and Repetition*, The Athlone Press, London, 14.
- [6] Plotnitsky, A. (2003) Algebras, geometries, and topologies of the fold: Deleuze, Derrida, and quasi-mathematical thinking (with Leibniz and Mallarmé), *Between Deleuze and Derrida*, Paul Patton & John Protevi, ed., Continuum, London, New York, 98-119.
- [7] Eisenman, P. (1992) Visions unfolding: architecture in the age of electronic media, *Domus* no. 734, 17-21.
- [8] De Landa, M. (2002) *Intensive science and virtual philosophy*, Continuum Books, London, 26.
- [9] Dženks, Č. (2007) *Nova paradigma u arhitekturi*, Orion Art, Beograd.
- [10] Spuybroek, L. (2008) *The Architecture of continuity*, NAI Publishers, Rotterdam, 23.
- [11] Buchler J. (1955) *Philosophical Writings of Peirce*, Dover Publications Inc, New York.
- [12] Vidler, A. (2001) The Future is a Graph: Dagmar Richter's Diagrammatic Practice, *X Y Z: The Architecture of Dagmar Richter*, Princeton Architectural Press, New York, 15.
- [13] Delez, Ž. (1989) *Fuko*, Izdavačka knjižarnica Zorana Stojanovića, Sremski Karlovci, 99.
- [14] Imperiale, A. (2002) Digital Skins: The Architecture of Surface, *Skin: Surface, Substance, and Design*, Ellen Lupton ed., Princeton Architectural Press, New York, 56-57.
- [15] Bendžamin, E. (2011) *Filozofija arhitecture*, Clio, Beograd, 37.
- [16] Dimendberg, E. (2013) *Diller Scofidio + Renfro: Architecture after Images*, The University of Chicago Press, Chicago, 146-147.
- [17] Alfariš, A. (2009) *Emergence Through Conflict, The Multi-Disciplinary Design System (MDDS)*, doctoral dissertation, Department of Architecture, MIT, Massachusetts, 49.
- [18] Moussavi, F. (2009) *The function of form*, Actar and Harvard University Graduate School of Design, New York, 31-32.
- [19] Terzidis, K. (2003) Expressive Form. A Conceptual Approach to Computational Design, Spon Press, London and New York, 33.
- [20] Lynn, G. (1999) *Animate form*, Princeton Architectural Press, New York, 11.
- [21] Picon, A. (2003) Architecture, Science, Technology and The Virtual Realm, *Architecture and The Science, Exchanging Metaphors*, Antao Picon and Alessandra Ponte, eds., Princeton Architectural Press, New York, 294.
- [22] Meredith, M. (2008) Never Enough (transform, repeat ad nausea), *From control to design: Parametric/Algorithmic Architecture*, Tomoko Sakamoto ed., Actar, Barcelona, 7-9.

Application of Uniform Design on Improvement Design of Detector Slides in Switch Machine system

†*Yung-Chang Cheng, and Jin-Long Liou

Department of Mechanical and Automation Engineering, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan.

*Presenting author: yccheng@nkust.edu.tw

†Corresponding author: yccheng@nkust.edu.tw

Abstract

The purpose of this paper is to present the use of uniform design of experiments method in improving the von Mises stress of the detector slides in the switch machine. Four system parameters of the detector slide are selected as the control factors to be improved. Uniform design of experiment is applied to create a set of simulation experiments. Applying ANSYS/Workbench software, the finite element modeling is investigated and the von Mises stress of each detector rod is calculated under fatigue testing. From the numerical results, the best detector slide of all the experiments which causes the smaller von Mises stress is selected as the improved version of design.

Keywords: Uniform design, Detector slide, ANSYS/Workbench, von Mises stress

Introduction

Railway turnout consists of switch machines and crossings with specific complexity which is exposed to several defects. When the train through the turnout area, the safe and reliable operation of switch machines and crossings must be assured by high levels of routine maintenance. Figure 1 shows the all elements in the turnout area. [1]

Some literatures have presented the dynamic analysis of the switch machine system. By using the dynamic switch machine model, Xu et al. [2] studied the lock calculation of nose rail after conversion and presented the stress and deformation for the nose rail and wing rail. Wang et al. [3] investigated the effect of different bedplates, different friction coefficient, lateral stiffness of switch rail end, stroke error and performance of fasteners on switching force and deviation. Based on envelope and morpheme match algorithm, exact curve matching method is used to match the detected current curve with the reference curve by Mo et al. [4].

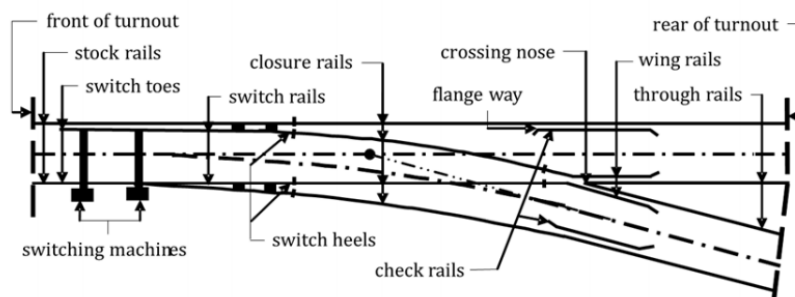


Figure 1. Turnout system and its components [1]

In this article, the fatigue analysis of detector slides under the dynamic forces is investigated. The uniform design of experiment is used to create a set of simulation experiments. According to the GB/T 25338.1-2010 fatigue testing standard, the maximum von Mises stresses of the upper and lower detector slides are obtained by using ANSYS/Workbench and HyperMesh software. Finally, from the numerical results, the best detector slide of all the experiments which causes the smaller von Mises stress is presented.

Finite Element Modeling for Detector Slide

The S700K switch machine system, including motor, clutches, control system, throw bar, detector slides, gear reducer system and case, is shown in Figure 2. [5] The throw bar, which is connected to the points, is held with a defined force in the end positions by the trailing clutch. During the trailing of a trail-able point machine, it is released once the retention force is exceeded. For fail-safe detection of the blade end positions, the point machine is equipped with detector slides. The detector slides are linked to the point blades via the detector rods and prove whether the blades have reached the end position. The end position of the point blades is detected continuously.



Figure 2. S700K switch machine system

Fatigue testing standard

Figure 3 illustrates schematically the detector slide fatigue test. The upper and lower detector slides are assembled in the guide sleeve as shown in Figure 3. Then, the upper and lower detector slide connects to a fork joint to translate the force to the lock member linked the switch rails. The guide sleeve is fixed on the case and the case is fixed on the ground. Next, the dynamic loads 6000 N act on the lock member in the horizontal direction as shown in Figure 3. The dynamic loads of the upper and lower detector slides are given in Figure 4. According to the GB/T 25338.1-2010 fatigue testing rule [6], the maximum endured loading is conducted to 10^6 cycles.

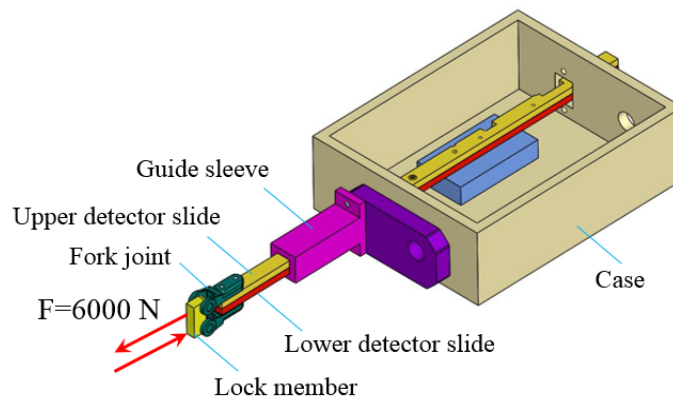


Figure 3. S700K switch machine system

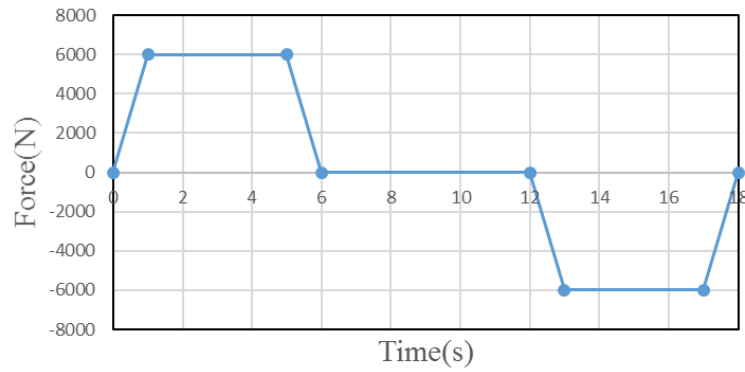


Figure 4. Time history of dynamic loads in one period

Finite element modeling

In the solving stage, HyperMesh software is used for the pre-processing. The boundary and contact conditions are given as shown in Figure 5. ANSYS/Workbench software is applied to calculate the maximum von Mises stress of the detector slide. The material properties of the detector slides, fork joint, lock member, guide sleeve, and the case are given as shown in Table 1. In addition, the S-N curve for the detector slide is given as shown in Figure 6.

In FE analysis, a convergence test is necessary for verifying the mesh quality because the solution is approximate rather than an exact solution. Consequently, each FE model executed the simulation test and the convergence condition is determined to be less than 5% of changes in the maximal magnitudes of von Mises stress with varying element sizes. Figure 7 shows the convergence curve of the obtained maximal magnitude of von Mises stress with meshing different element size. The optimized element size is 4 mm because the difference of the simulation result between element size of 4 and 3.5 mm is less than 5%. Therefore, element size is determined to be 4 mm to mesh all FE models.

The maximum von Mises stress and fatigue safety factor of each detector slide for the fatigue testing are shown as Figures 8 and 9. From the numerical results, the maximum von Mises stresses of the upper and lower detector slide are given as 117.16 MPa and 121.5 MPa, respectively. In addition, the minimum fatigue safety factors of the upper and lower detector slide are given as 1.34 and 1.29, respectively. Figure 10 shows the deformation of each detector slide. From the numerical results, the maximum deformations of the upper and lower detector slide are given as 0.84 mm and 0.52 mm, respectively. It means that the strength performance of the upper detector slide is better than the lower detector slide.

Table 1. Material types and properties

Body	Young's Modulus (Pa)	Poisson's Ratio	Density (g/cm ³)
Detector slide	21×10^{10}	0.29	7.9
Fork joint	20×10^{10}	0.2	7.85
Lock member guide			
Sleeve the case			

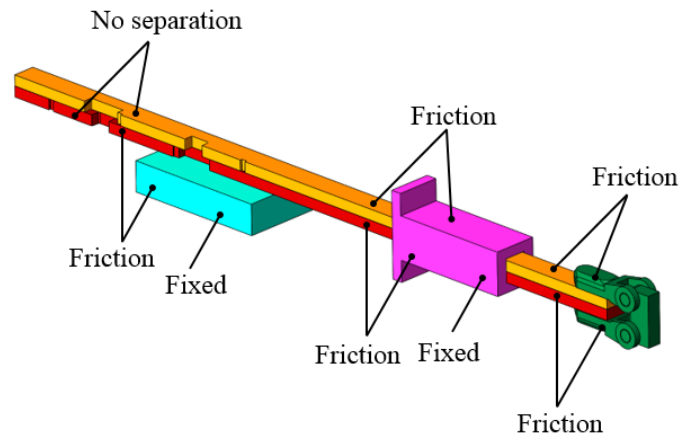


Figure 5. Boundary conditions for detector slides

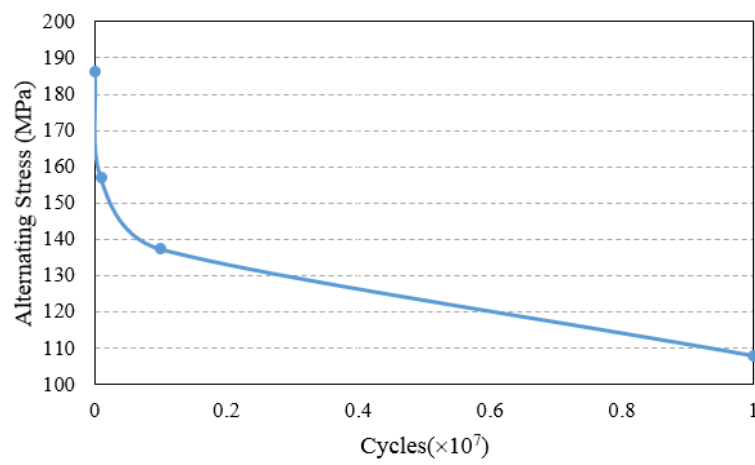


Figure 6. S-N curve for detector slide

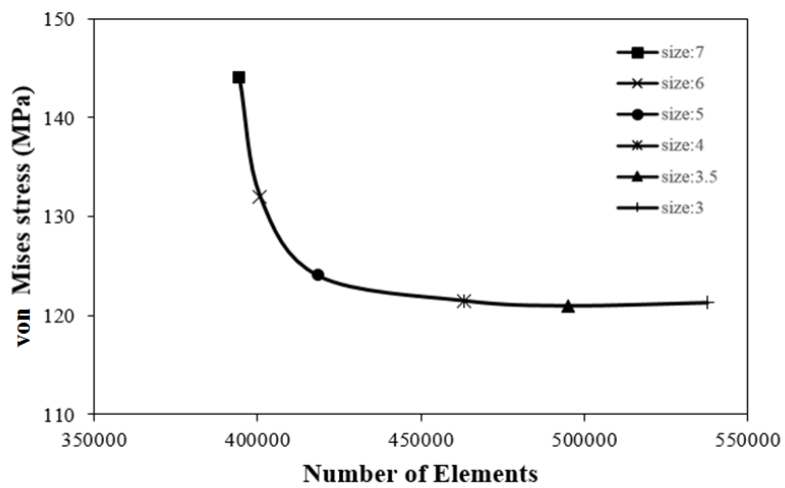


Figure 7. Convergence curve of the obtained maximal magnitude of von Mises stress

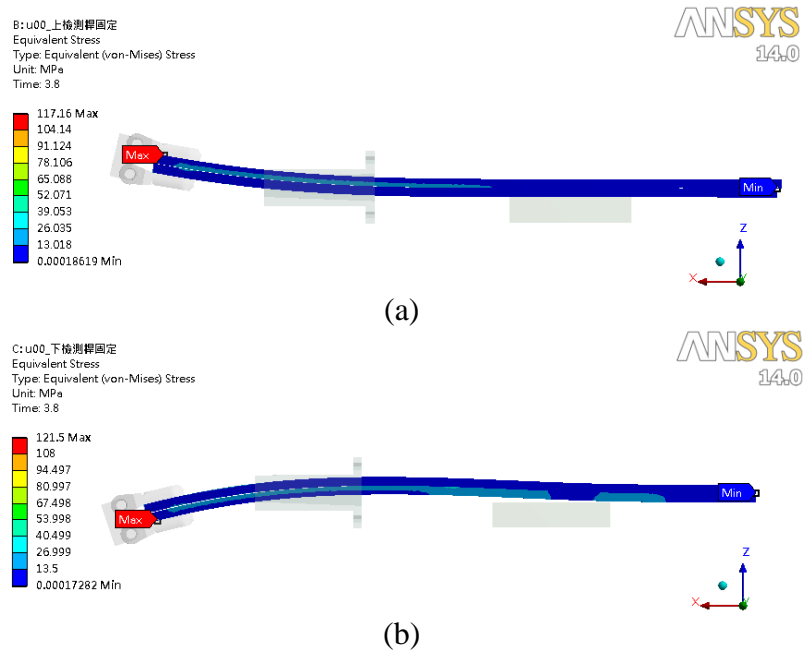


Figure 8. von Mises stress of the (a) upper detector slide and (b) lower detector slide.

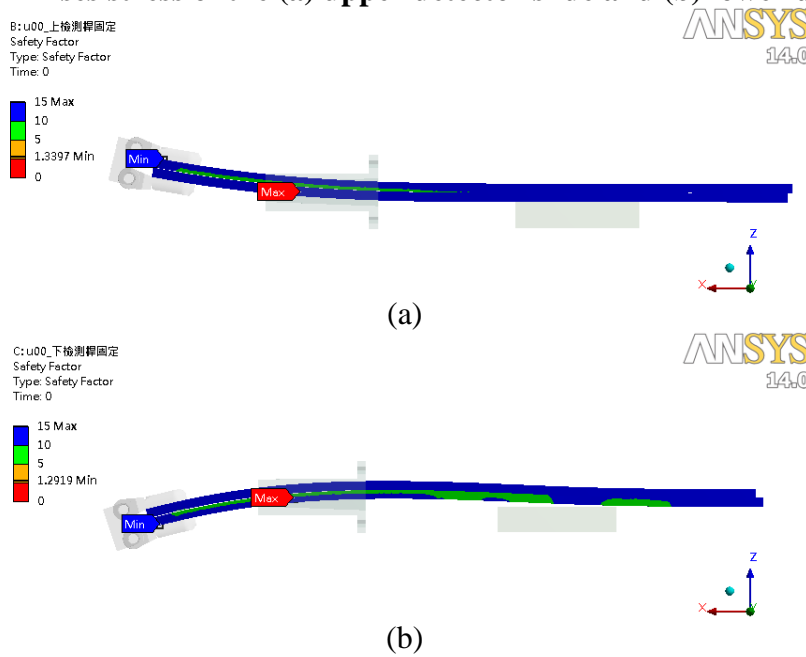


Figure 9. Fatigue safety factor of the (a) upper detector slide and (b) lower detector slide.

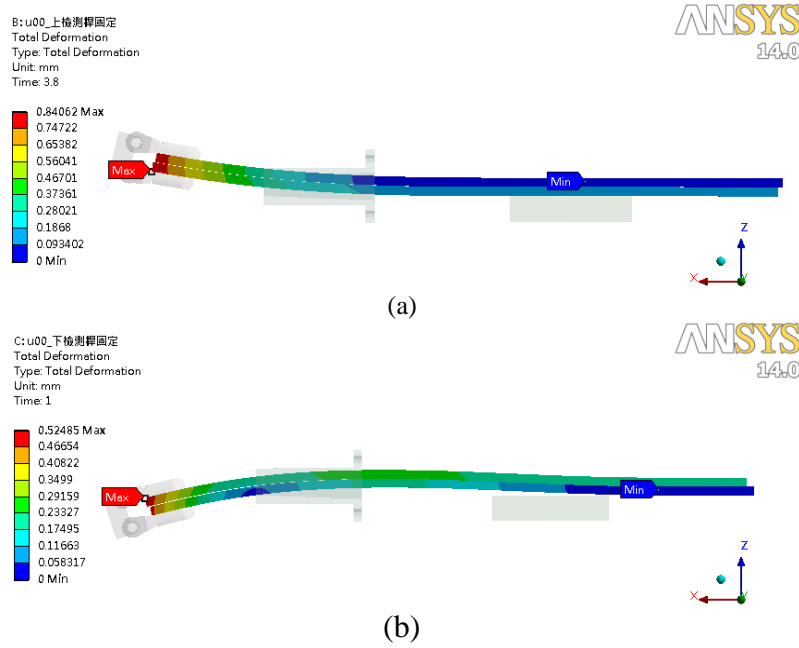


Figure 10. Deformation of the (a) upper detector slide and (b) lower detector slide.

Uniform Design of Experiments for Reducing von Mises Stress of Detector Slide

As shown in Figure 10, the detector slide studied in this paper has four dimensional control factors. The original design and variation ranges of control factors are shown in Table 2. Since all factors are continuous, the design space is also a continuous space. For a continuous design space, design points are infinite and evaluation of all design points is impossible. Therefore, this study applies uniform design method proposed by Fang and Wang [7] to construct a set of sample points which are scattered uniformly in the continuous design space. Uniform design has been widely used for optimization in many engineering applications [8-10].

In our study, due to the limitation of computational resource, each factor is divided into 12 levels and the uniform table $U_{12}^*(12^{10})$ is utilized to construct 12 experiments. As shown in Table 3, the uniform table $U_{12}^*(12^{10})$ has 12 rows and 10 columns (Fang and Wang [7]). Since the detector slide has four control factors, columns 1, 6, 7, 9 should be used according to the use table of $U_{12}^*(12^{10})$. The constructed experiments are shown in Table 4. Each experiment denotes a specific design of detector slide. For each detector slide, SolidWorks is applied to build the geometric model of frame and ANSYS/Workbench is used to simulate the behavior of detector slide undergoing fatigue testing.

From the numerical results in Table 4, the minimum von Mises stress of the upper detector slide occurs at the 3rd experiment. However, at 7th experiment, it has the minimum von Mises stress of the upper detector slide. It is difficult to select an optimal solution. Considering the lower von Mises stress for each detector slide, the 7th experiment is regarded as the improved version of design. The improved version causes the von Mises of 104 MPa and 110.2 MPa for the upper and lower detector slide, respectively.

Factor Selection of Upper Detector Slide:



Factor Selection of Lower Detector Slide:

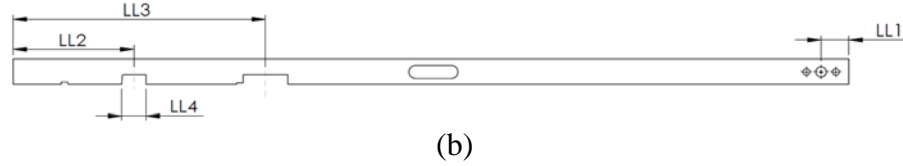


Figure 10. Control factors of the (a) upper detector slide and (b) lower detector slide

Table 2. Variation ranges of control factors

Control factors	Lower bound	Original design	Upper bound
LL1 (mm)	30.6	34	37.4
LL2 (mm)	120.47	133.85	147.24
LL3 (mm)	253.8	282	310.2
LL4 (mm)	27	30	33

Table 3. Uniform table $U_{12}^*(12^{10})$

Experiment No.	1	2	3	4	5	6	7	8	9
1	1	2	3	4	5	6	8	9	10
2	2	4	6	8	10	12	3	5	7
3	3	6	9	12	2	5	11	1	4
4	4	8	12	3	7	11	6	10	1
5	5	10	2	7	12	4	1	6	11
6	6	12	5	11	4	10	9	2	8
7	7	1	8	2	9	3	4	11	5
8	8	3	11	6	1	9	12	7	2
9	9	5	1	10	6	2	7	3	12
10	10	7	4	1	11	8	2	12	9
11	11	9	7	5	3	1	10	8	6
12	12	11	10	9	8	7	5	4	3

Table 3. Constructed experiments and results

Experiment No.	von Mises stress of upper detector slide (MPa)	von Mises stress of lower detector slide (MPa)
1	111.3	113.9
2	118.4	114.4
3	97.5	117.4
4	110.0	150.4
5	112.2	112.2
6	115.0	120.9
7	104.0	110.2
8	113.3	129.3
9	113.8	121.1
10	113.0	118.2
11	118.2	121.0
12	116.0	120.7

Conclusions

This paper has completed the improvement of von Mises stress for the detector slides under the fatigue testing simulations by using uniform design method. Uniform design method is used to build a set of experiments and finite element modeling is fulfilled by employing ANSYS/Workbench. For the original design, the von Mises stress for the upper and lower detector slide are 117.16 MPa and 121.5 MPa, respectively. After executing uniform design improvement, the von Mises stresses of the improved version will go down to 104 MPa and 110.2 MPa. This paper has shown that uniform design is a useful tool to reduce the von Mises stress of detector slides.

Acknowledgment

The authors gratefully acknowledge the financial support from the Ministry of Science and Technology of Taiwan under Grant no. MOST 106-2221-E-327-016.

References

- [1] Hassankiadeh, S. J., Failure Analysis of Railway Switches and Crossings for the purpose of Preventive Maintenance, MS Thesis, Royal Institute of Technology, 2011
- [2] Xu, J. M., Wang, P. and Xu, H. (2013) Improvement and lock calculation of electric switch machine of railway turnout, *Applied Mechanics and Materials*, **409-410**, 496-1501.
- [3] Wang, P., Chen, R. and Xu, H. (2013) Conversion and its deviation control of electric switch machine of high speed railway turnout, *Journal of Vibroengineering*, **15 (3)**, 1513-1525.
- [4] Mo, X. M., Fang, Y. and Yang, Y. G. (2013) Method on the fault detection and diagnosis for the railway turnout based on the current curve of switch machine, *Applied Mechanics and Materials*, **427-429**, 1022-1027.
- [5] <https://www.mobility.siemens.com/mobility/global/SiteCollectionDocuments/en/rail-solutions/rail-automation/signaling-components/s-700-k-en.pdf>
- [6] http://www.gbstandards.org/GB_standards/GB_Search.asp

- [7] Fang, K. T. and Wang, Y. (1994) Number-Theoretic Methods in Statistics, Chapman & Hall Press, London, U.K.
- [8] Liang, Y. Z. Fang, K. T. and Xu, Q. S. (2001) Uniform Design and its Applications in Chemistry and Chemical Engineering, *Chemometrics and Intelligent Laboratory Systems*, **58**, 43-57.
- [9] Tang, M., Li, J., Chan, L. Y. and Lin, D.K.J. (2004) Application of Uniform Design in the Formation of Cement Mixtures, *Quality Engineering*, **16** (3), 461-474.
- [10] Gao, N., Starink, M. N. and Sinclair, K. I. (2007) Application of Uniform Design in Optimization of three Stage Ageing of Al–Cu–Mg Alloys, *Journal of Materials Science*, **42** (12), 4398-4405.

Numerical simulation of Two-dimensional risers under oscillatory flows with low Reynolds and KC for predicting the involve forces response

*M. Catalina Valencia-Cárdenas and C. A. Riveros

School of Engineering, Research group GEOc, Universidad de Antioquia, Colombia.

* Presenting author: mcatalina.valencia@udea.edu.co

ABSTRACT

This research implements a numerical simulation of flexible risers under oscillatory flows initially at rest. A single freedom degree spring-mass-damper system is employed with low mass allowed to move only in cross-flow direction. Two-dimensional incompressible Navier-Stokes equations are discretized using the Finite Volume Method (FVM). To resolve the pressure-velocity linkage, an iterative solution strategy SIMPLEC algorithm for transient problems is used. To study the influence of smaller Reynolds values on the dynamic system, a critic value $KC=7,9$ is fixed. For Reynolds less than 300, the system behavior is in agreement with the literature, where a direct relationship between the force and the vortex shedding is presented. An almost null cross-flow force appear for $Re=40$ which becomes important as increased the Reynolds value, changing between fluctuating and stabilizing force. From Reynolds 300 the cross-flow force is very chaotic and some discrepancies with literature appear in the system dynamics. The numerical results obtained from the proposed numerical base approach show good agreement with experimental data collected from a flexible riser model according to the spectral analysis.

Keywords: *Flexible riser, Keulegan-Carpenter number, oscillating flow, numerical simulation, vortex dynamics*

1. Introduction

The structural movement by external variable velocity flow is a very common phenomenon in flexible risers but there is still a limitation to accurately predicting the response of these structures. This is due to most of the prediction models rely heavily on large experimental databases. In addition, flexible and light materials have been developing for marine applications leading to slender structures with low mass ratios. This characteristic facilitates the structure movement caused by vortex shedding for incoming flow. In many real cases, the flow can be oscillating in a sinusoidal mode and Reynolds (Re) and Keulegan-Carpenter (KC) dimensionless numbers can describe this behavior. Here $Re = U_m D / \nu$ and $KC = U_m T / D$, where ν is the kinematic viscosity, U_m is the amplitude of the oscillatory flow velocity, T is the oscillatory flow period and D is the diameter of the cylinder. Their importance regards in the relationship between these numbers and the forces that causing the structure movement. The cross-flow force exerted onto the structure is particularly harmful, exerting a significant dynamic stress, increasing the damage accumulation and eventually causing structural failure.

The frequency at which the vortices are shear is known as vortex shedding frequencies (f_v); a regular pattern of vortices induces fluctuating lift and drag forces on the cylinder [1]. The shedding frequency and both forces become constant for Reynolds number intervals, which means the heavy influences of this number on the cylinder maximum response ([2] [3]). Many authors have been researching its influence until a value of 500, considering a fixed value for KC ([4], [5], [6], [7], [8], [9], [10], [11]). Tatsuno and Bearman [12] analyzed $1.6 < KC < 15$ and $5 < \beta < 160$ ($\beta = Re/KC = D^2/\nu T$), including three-dimensional features, provided the identification of eight regimes denoted from A to G in a plane (KC, Re). This classification has become the standard description for the associated flow regimes. The regimen A corresponds to the Williamson's symmetrical regimen [13], that is also similar to Regime B but with an

axial direction three-dimensional structure. Regime C corresponds to vortices of opposite rotation senses in the same fashion of a Von Kármán vortex street. Regime D exhibits a symmetrical V-pattern around the transverse axis, very similar to regime E, however here the V-pattern changes intermittently its direction from one side to the other. Regime F describes the Williamson's double pair regime [13], whereas the Williamson's transverse street is similar to Regime G. Finally, the Tatsuno and Bearman [12] regimen classification suggests that a KC such as 7,9 that crosses five different regimens, can be assumed as critical and important to be studied.

This study describes and discusses the resulting forces by oscillating fluid flow effect around a cylinder under the influence of a fixed KC (7,9) and Re between 40 and 1000, considering lower mass ratio and covering the most Tatsuno and Bearman regimes (A, D, E, F and G). The cylinder-oscillating period is implemented for a long time (more than fifty cycles) in order to analyze the flow patterns in each regime. A single degree of freedom system with a spring-mass-damper is implemented, where the mass is allowed to move only in cross-flow direction. To discretize the transport equations, the Finite Volume Method (FVM) is used and to resolve the pressure-velocity linkage, an iterative solution strategy SIMPLEC algorithm for transient problems is used. Moreover, a bi-dimensional model is established using OpenFOAM simulations and employing a single desktop computer. The results presented here consider an experimental validation [14] in order to demonstrate the use of numerical based approaches to predict the response of flexible risers.

In what follows, Section 2 presents the numerical method description with the respective equations modeling, parameters taking account, computational domain, boundary conditions and model validation. Section 3 put forward a discussion of the force characterization according to the transitions between regimes established for different intervals of Reynolds values. In addition, the results are compared with experimental data in this section. A study summary is presented in Section 4.

2. Numerical method

To model the vortex generation around a structure, commonly a sinusoidal oscillatory flow is considered to represent a realistic phenomenon representation. The flow in the longitudinal direction is given by:

$$U_1(t) = U_m \sin(2\pi t/T) \quad (1)$$

The oscillating flow considered is controlled by 40, 100, 150, 200, 250, 300, 500 and 1000 as Reynolds numbers and 7,9 as Keulegan Carpenter number, so eight different regimens were simulated. Other parameters (mass, damping, reduced velocity) are set up to allow the cylinder movement according to Valencia-Cardenas, M. [15].

A discretized form of two-dimensional incompressible Navier-Stokes equations must be defined at a nodal point placed within each control volume in order to solve the problem. OpenFOAM, an open source solver, is used to solve the governing equations selecting adequate solution schemes in order to achieve reliable results. To reach it, a second-order central difference scheme is used for the convection and diffusion terms. A stable and accurate simulation is obtained by choosing an implicit second-order scheme for temporal discretization. For the numerical procedure in the simulation, to improve the pressure and velocity coupling, the PIMPLE algorithm is utilized [16].

The structure is allowed to move only perpendicularly to the flow direction. To apply the transport equations to the inertial system, time to time and according to the cylinder movement, the numerical grid is moved and adjusted. For that reason, a mesh dynamic motion solver is implemented in the model, where the cylinder is constrained to only move along “y” and cannot rotate. Finally, the total force per unit length by a stationary cylinder under a oscillatory flow F_{osc} is known as Morison’s equation [17], written as:

$$F_{osc}(t) = \rho C_m \frac{\pi}{4} D^2 \dot{U}(t) + \frac{1}{2} \rho C_D D |U(t)| U(t) \quad (2)$$

Where ρ represents the fluid density, C_m the inertia coefficient and C_D the drag coefficient. The last two are functions of Re and KC .

The computational domain is a cylinder in a channel represented using two-dimensional numerical simulations as shown in Figure 1. The cylinder is represented as a circle with diameter D submersed in an incompressible fluid, represented here as a rectangular flow domain. As the simulation begins, the center body is located at the center of the coordinate’s axis, $10D$ from the horizontal walls and $20D$ from the vertical walls. The domain areas around the cylinder, where the vortices are shed, contains a higher cell density in order to obtain a better resolution. This region is shaped by four arcs whose radius equal $2.5\sqrt{2}D$.

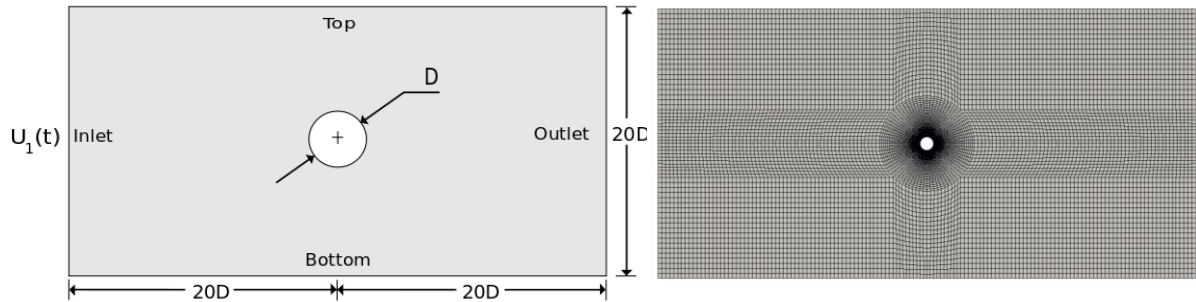


Figure 1. Sketch and computational mesh of oscillating flow around circular cylinder

To guarantee the smallest numerical errors, it is necessary to proof the meshing independently. The test is developed from a course mesh established, using a non-dimensional time step $U_1 \Delta_t / D = 0,1$ (where Δ_t is time step) as sufficient condition to ensure coefficients with three significant digits [18]. Then, the mesh is refined consecutively and the time step is determined from the Courant number (c_o) expression, $c_o = |U_1| \Delta_t / \Delta_x$ where Δ_x is the smaller cell size in the velocity direction and c_o is defined as 0.2 [19]. Finally, the appropriate mesh is selected taking into account the fitting between results and literature and the tradeoff between precision and computational cost (see Figure 1).

Table 1. Comparison of drag force coefficient ($\bar{x}=1.307$, 95% CI [1.281, 1.332]) and Strouhal number ($\bar{x}=0.193$, 95% CI [0.187, 0.198]) at $Re = 200$ and $KC = 10$

	\bar{C}_D	S_t
Guilmineau and Queutey (2002)	1.286	0.195
Cao et al. (2010)	1.300	0.186
Cao and Li (2015)	1.343	0.191
Present work	1.331	0.192

In order to guarantee an accurate solution, the model is setting up at values of Reynolds and Keulegan Carpenter well studied in the literature (e.g. [18], [20], [10], [21]). The drag and lift coefficients time history and the Strouhal number are analyzed, considering 20 vortex shedding periods once the periodic flow is established. These values have been compared with published results ([10], [21], [22]) and shown in Table 1.

The drag coefficient (mean value of the in-line non-dimensionalized force) and the Strouhal number, $S_t = f_v D / U_m$, are obtained from the frequency of vortex shedding f_v , which is calculated with the period measured from velocity time history. At $Re = 200$ and $KC = 10$, \bar{C}_D is equal to 1.331 and S_t is equal to 0.192, meaning that the vortex natural frequency shedding is $f_0 = 0.192$. Results are in good agreement with those published in the literature (see Table 1).

3. Results and discussions

The results of direct numerical simulation are presented in this section considering the effect of $KC = 7,9$ and Reynolds values equals to 40, 100, 150, 200, 250, 300, 500 and 1000. In this work, the regimes are defined from the flow structure and force behavior.

3.1. In-line and transverse forces

Drag and lift coefficients time histories in an oscillating flow are estimated using force coefficients function library by OpenFOAM. Vortex shedding frequencies and Strouhal numbers obtained for different Reynolds values are shown in Figure 2. The time history frequencies are verified for both drag and lift coefficients in order to obtain the vortex shedding frequency, using the Fast Fourier Transform method (FFT) [20]. The dominant frequency corresponds to the oscillating frequency.

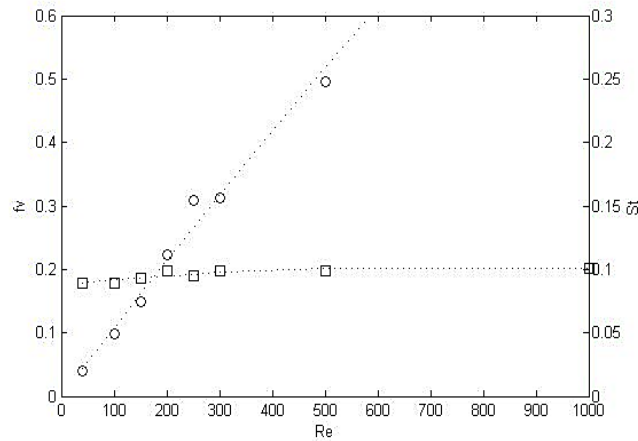


Figure 2. Vortex shedding frequencies (circle) and Strouhal numbers (square) by Reynolds number

The semi-empirical Eq. (2) estimate the in-line force on a cylinder divided in two forces: 1) The drag force, proportional to the flow instantaneous velocity square and 2) the inertial flow coupled with the local flow acceleration [22]. The drag force coefficient C_d and the inertial force coefficient C_m can be obtained by last square fitting on the time history of F_{osc} , or calculated using the equations below [23].

$$C_d = \frac{3}{4} \int_0^{2\pi} \frac{F_D \sin \theta}{\rho D U_1^2} d\theta = \frac{3}{8} \int_0^{2\pi} C_D \sin \theta d\theta \quad (3)$$

$$C_m = \frac{2U_1 T_f}{\pi^3 D} \int_0^{2\pi} \frac{F_D \cos \theta}{\rho D U_1^2} d\theta = \frac{U_1 T_f}{\pi^3 D} \int_0^{2\pi} C_D \cos \theta d\theta \quad (4)$$

Where, C_D represents the mean drag coefficient.

According to Cao and Li [22], the time history of drag force can be acceptably approximated by Morrison's equation. Using the same equations is possible to obtain the cross-flow force. Figure 3 shows an interval of in-line and cross-flow forces for every Reynolds value studied.

The drag and the inertial forces share a direct relation with the vortex behavior as shown in Figure 3. A uniform in-line force with almost constant frequency and period can be visualized in Figure 3(a). This regimen A does not exhibit vortex detachment, but it does show a vortex formation totally symmetric in the direction of the flow. This is why in-line force predominates and the transverse force is almost null.

A representative regimen D is shown in Figure 3(b), where the V-pattern symmetrical starts to dominate and the cross-flow force becomes important. The symmetric pattern becomes predominant when the cross-flow force amplitude reduces. The fluctuation force classifies this case of Reynolds value in a symmetric regime group.

The Re=150 case presents amplitude instability due to irregular vortex shedding in Figure 3(c). After a while, the fluctuations tend to become regular and the in-line force stabilizes. Here regimen E is predominant with intermittent changes of direction, related with the fluctuating amplitude in-line force because of the action of cross-flow force. When cross-flow force presents a peak, vorticity pattern tends to be transversal.

For Re= 200, an amplitude stability is presented and persists at the time (Figure 3(d)). In the same way, the Re= 250 presents a regular fluctuation force (Figure 3(e)). Also, cross-flow force presents uniform behavior for both Reynolds (200 and 250). Reynolds cases 200 and 250 are dominated by regimen F. In-line force for Re= 200 tends to be more stable in comparison with Re= 250 which means the first one presents vortex shedding more symmetrical about the cross-flow axis while the second one tends to be more transversal. In this way, the proximity to the transition range is a bit evident for Re= 250.

For cases of Reynolds 300, 500 and 1000, in-line and cross-flow forces are chaotic and strong peaks appear (see Figures 3(f)(g) and (h)). A chaotic behavior is observed here because there are no persistent vortex patterns. Sometimes, the dominant harmonic for the case of Re= 1000 is three times the frequency of oscillating fluid flow, but predominates two times the oscillating frequency. On the other hand, flow regime is dominated by the viscous drag component in all the cases.

3.2. Spectral analysis

Several oscillating fluid flow frequency peaks that are integral times the vortex oscillating frequency, are illustrated in Figure 4 and Figure 5. These figures were obtained from the time history of drag force and lift force coefficients respectively, using Fast Fourier Transform (FFT). The oscillating frequencies are graphed with the magnitude of the Fourier Transform, using the main peak as indicated by Williamson [13].

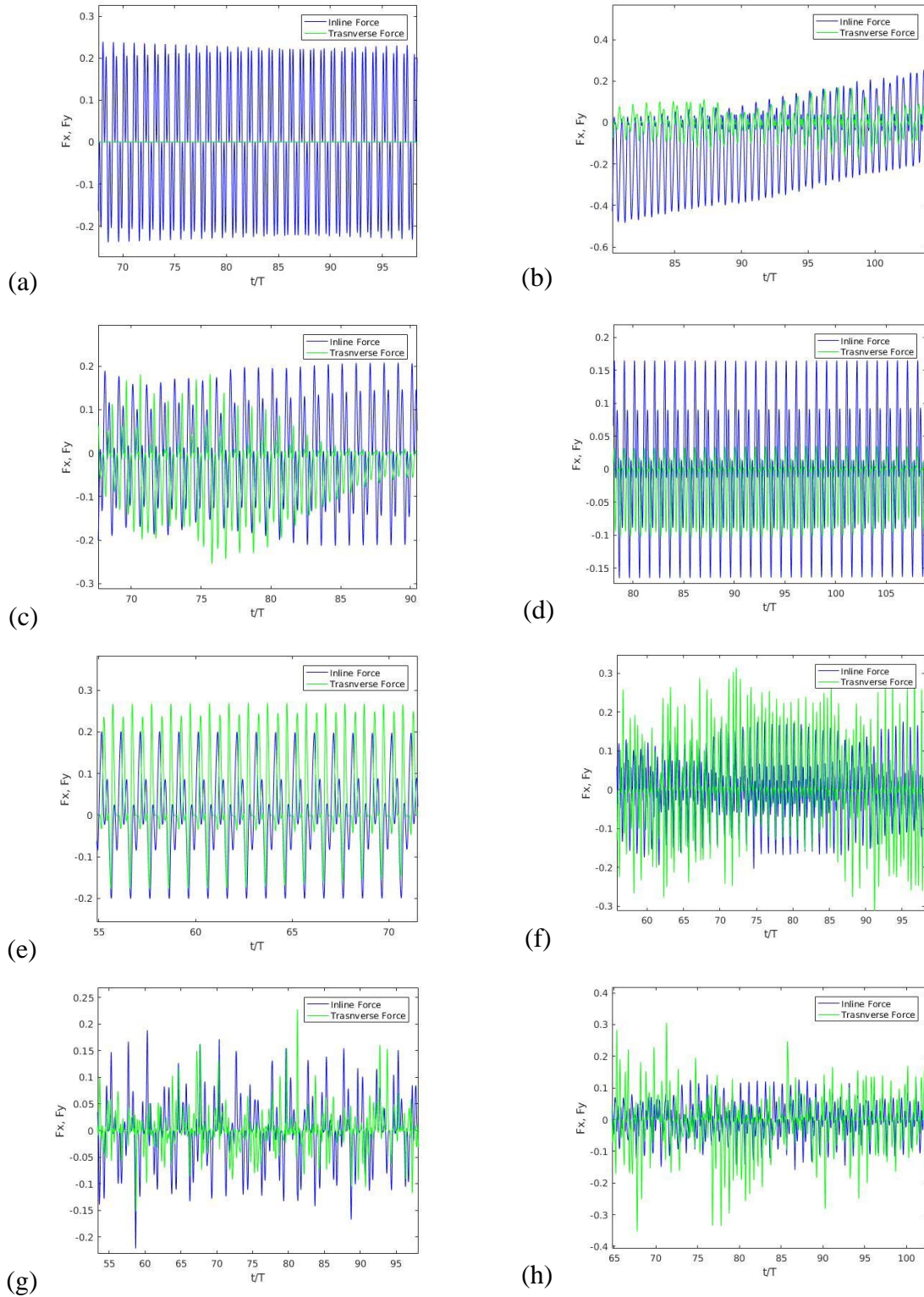


Figure 3. In-line and cross-flow forces for a) Re 40, b) Re 100, c) Re 150, d) Re 200, e) Re 250, f) Re 300, g) Re 500 and h) Re 1000.

Spectral analysis of drag force is shown in Figure 4 with several peaks at frequencies. The main frequency are always the oscillating frequency and the other peaks has an increment factor, namely $3f_0, 5f_0, 7f_0$ and in that way forward (see Figure 4(a)(b)(c)), for regimes A, D, E and

F. Besides of the main frequency and the $3f_0$ peak, other frequencies peaks without multiple of f_0 appear in regimen G.

It is possible to note, the oscillating frequencies for Reynolds values less than 200 (see Figure 4) are twice the main frequency of lift coefficient. As shown in Figures 5(a)(b), for regimes A and E respectively, the main frequencies occur with an increment of $2f_0$. Otherwise, the dominant frequency for regimen F (Figures 5(c)) is three times the oscillating frequency. In this case, the case of Reynold value 250 presents main frequency at $3f_0$ and the increment is about $1f_0$ in the others frequencies peaks.

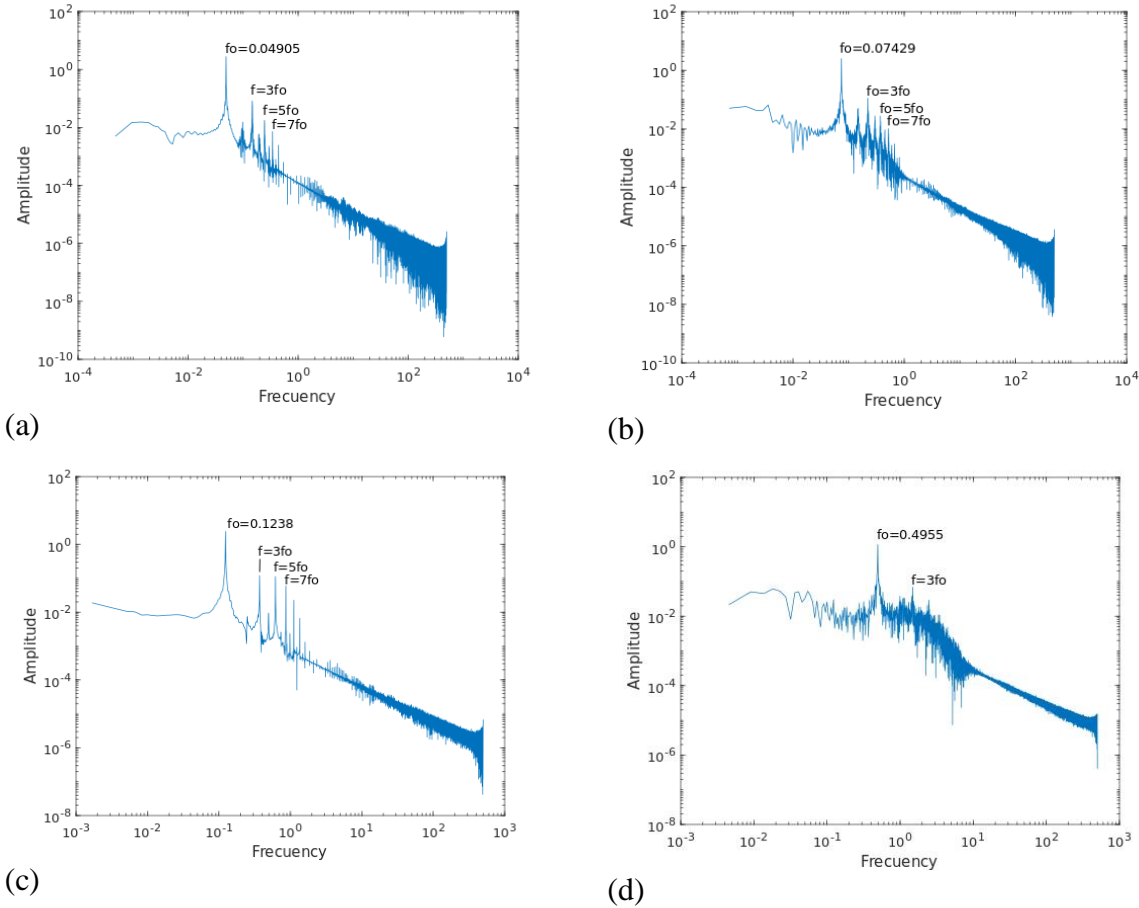


Figure 4. Spectral analysis of drag force coefficient at a) Re 100, b) Re 150, c) Re 250, d) Re 1000.

The same behavior for these regimens is shown by Duclercq et. al. [4]. Again, for Reynolds values higher than 300 the oscillating frequencies are two times the main frequency of lift coefficient. Note that the main frequency in Figures 5(d), is presented around of $2f_0$. Then appear others frequencies peaks without a clear multiple of f_0 and strong fluctuations is observed. Williamson [13] concluded that in an oscillating flow, the dominant frequency of lift force is equal to one plus the number of vortices shedding in a half period, which is evidenced in the regimes A, D and F. However, for regimen G this condition is not fulfilled as proved in this study.

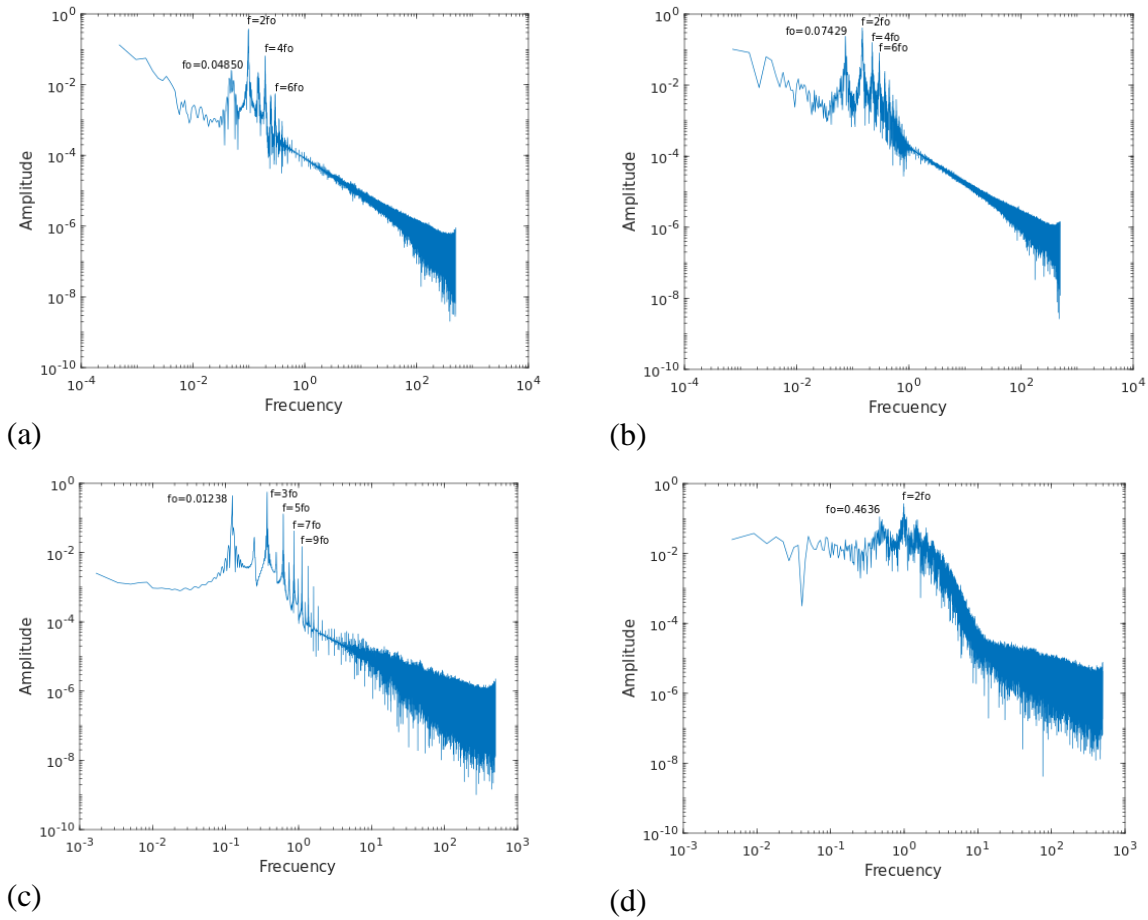


Figure 5. Spectral analysis of lift force coefficient at a) Re 100, b) Re 150, c) Re 250, d) Re 1000.

3.3. Experimental validation

Riveros et. al. [14] conducted a series of forced oscillation experiments for flexible risers where a 20-meter riser model was tested for different values of Re and KC numbers; their experimental model case 1 has the same diameter as Re (1000) and KC (7,9) numbers presented here. Although good agreement was reported by Riveros et. al. [14], it is still possible to observe some deviations between the simulation results and experimental data in the main cross-flow frequency. The model presented in this paper, as shown in Figure 4(d), overcomes this difficulty providing a value of the dominant cross-flow frequency in good agreement with the experimental value of 1 Hz presented by Riveros et. al. [14]. Likewise, the dominant inline-flow frequency presented in the experimental model by Riveros et. al. [14] is 4,9 Hz, the same visualized in Figure 4(d) using the numerical model.

4. Summary

A numerical simulation of two-dimensional risers under oscillatory flows with low Reynolds and KC for predicting the involve forces response was presented in this paper. A inverse relationship between the Reynolds number and the mean drag coefficients, is an expected behavior result [14]. Also demonstrated that the peaks shape variation of forces behaviour is related with a pressure distribution asymmetry in the flow direction, due to asynchrony in the vortex shedding. Thus, the lowest Reynolds number, the less variation in amplitude and

frequency parameters and conversely, the higher Reynolds numbers produces significant variations, especially in fluctuation amplitude.

Reynolds value 40 present a good agreement to the Williamson regimen with vortex symmetric formation in the flow direction which explains why the in-line force predominates and the cross-flow force is almost null. Instability develops when Reynolds value is incremented to $R_e = 100$ where transverse vortices appear. This is conforms to Tatsuno & Bearman regimen D behavior, where a symmetrical V-pattern began to develop and the cross-flow force becomes important. The V-pattern symmetrical regimen persists in the $R_e = 150$ case because changes intermittently from a transverse street to oblique street vortex. An amplitude instability is present due irregular vortex shedding, after a while, the fluctuations becomes regular and the in-line force stabilizes. Here regimen E is predominant with intermittently changes of direction, related with the fluctuating amplitude in-line force because the action of cross-flow force, this conforms to Tatsuno & Bearman.

In-line force for $R_e = 200$ tends to be more stable in comparison with $R_e = 250$, so $R_e = 200$ shows vortex shedding more symmetrical about the cross-flow axis while the second one tends to be more transversal. In this way, the proximity to the transition range is evident for $R_e = 250$. Both cases are conform to Tatsuno & Bearman regimen F behavior with regular fluctuation force and also, the cross-flow force present uniform behavior. Regimen instability appears as the Reynolds value increases (Reynolds value 300, 500 and 1000). Reynolds 300 and 500 is classified by Williamson [13] as regimen G, but transverse street is not the most persistent behavior. The pattern is considered chaotic for all these cases where there are ot persisten force patterns. In-line and cross-flow forces are chaotic and strong peaks appear. The dominant harmonic for the case of $Re = 1000$ sometimes is three times higher than the oscillating fluid flow frequency, but two times the oscillating frequency. The flow regime is dominated by the viscous drag component in all these cases.

Finally, the numerical results presented in this paper for the dominant in-line and cross-flow frequency shows good agreement with experimental results provided by Riveros et. al. [14].

Acknowledgements

The authors gratefully acknowledge to the reviewers, the Universidad de Antioquia and the Grupo de Estudios Oceánicos “Luis Fernando Vásquez-Bedoya” for the financing and logistic support provided to this work.

References

- [1] S. Chakrabarti, "The Theory and Practice of Hydrodynamics an Vibration," *World Scientific Publishing*, vol. 20, 2002.
- [2] C. Govardhan and R. Williamson, "Defining the Modified Griffin plot in vortex-induced vibration: revealing the effect of Reynolds number using controlled damping," *Journal of Fluid Mechanics*, vol. 561, pp. 147-180, 2006.
- [3] J. Klamo, A. Leonard and A. Roshki, "On the maximum amplitude for a freely vibrating cylinder in crossflow," *Journal of Fluid and Structures*, vol. 21, pp. 429-434, 2005.
- [4] M. Duclercq, D. Broc and O. Cadot, "Characterization of long time fluctuations of forces exerted on an oscillating circular cylinder at $KC=10$," *Journal of Fluids and Structures*, vol. 27, no. 4, pp. 18-29, 2011.

- [5] J. Lin and D. Rockwell, "Quantitative interpretation of vortices," *Experiments in Fluids*, vol. 23, no. 2, pp. 99-104, 1997.
- [6] P. Anagnostopoulos and G. Iliadis, "Numerical study of the flow pattern and the in-line response of a flexible cylinder in an oscillating stream," *Journal of Fluids*, vol. 12, pp. 225-258, 1998.
- [7] G. Iliadis and P. Anagnostopoulos, "Viscous oscillatory flow around a circular cylinder at low Keulegan–Carpenter numbers and frequency parameters," *International Journal for Numerical Methods in Fluids*, vol. 26, p. 403–442, 1998.
- [8] H. Dutsch, F. Durst, S. Becker and H. Lienhart, "Low-Reynolds-number flow around an oscillating circular cylinder at low Keulegan–Carpenter numbers," *Journal of Fluid Mechanics* 360,, vol. 260, p. 249–271, 1998.
- [9] B. Uzunoglu, M. Tan and W. Price, "Low-Reynolds-number flow around an oscillating circular cylinder using a cell viscous boundary element method," *International Journal for Numerical Methods in Engineering*, vol. 50, p. 2317–2338, 2001.
- [10] E. Guilmineau and P. Queutey, "A numerical simulation of vortex shedding from an oscillatory circular cylinder," *Journal of Fluids and Structures*, vol. 16, no. 6, p. 773–794, 2002.
- [11] D. Nehari, V. Armenio and F. Ballio, "Three-dimensional analysis of the unidirectional oscillatory flow around a circular cylinder at low Keulegan–Carpenter and beta numbers," *Journal of Fluid Mechanics*, vol. 520, p. 157–186, 2004.
- [12] M. Tatsuno and P. Bearman, "A visual study of the flow around an oscillating," *Journal*, vol. 211, pp. 157-182, 1990.
- [13] C. Williamson, "Sinusoidal flow relative to circular cylinders," *Journal of Fluid Mechanics*, vol. 155, p. 141–174, 1985.
- [14] C. Riveros, T. Utsunomiya, K. Maeda and K. Itoh, "Response prediction of long flexible risers subject to forced," *Journal of Marine and Science Technology*, vol. 15, pp. 44-53, 2010.
- [15] M. Valencia-Cardenas, *Numerical study of uniform and oscillatory flows over a circular cylinder (M.S. thesis)*, Universidad de Antioquia, 2017.
- [16] C. J. Greenshields, OpenFOAM User Guide, OpenFOAM Foundation , 2017.
- [17] J. Morison, M. O'Brien, J. Johnson and S. Schaaf, "The force exerted by surface waves on piles," *Journal of Candian Petroleum Technology*, vol. 189, p. 149–154., 1950.
- [18] S. Étienne and D. Pelletier, "The low Reynolds number limit of vortex-induced vibrations," *Journal of Fluids and Structures*, vol. 31, pp. 18-29, 2012.
- [19] J. K. Tawekal, CFD simulation of the Flow over a 2-Dimensional Pipe and Vortex Induced Vibration of the Pipe with 1 Degree of Freedom, Stavanger: University of Stavanger, 2015.
- [20] Z. Pan, W. Cui and Q. Miao, "Numerical simulation of vortex-induced vibration of a circular cylinder at low mass-damping using RANS code," *Journal of Fluids and Structures*, vol. 23, pp. 23-32, 2007.
- [21] S. Cao, S. Ozono, Y. Tamura, Y. Ge and H. Kikugawa, "Numerical simulation of Reynolds number effects on velocity shear flow around a circular cylinder," *Journal of Fluids and Structures*, vol. 26, p. 685–702, 2010.
- [22] S. Cao and M. Li, "Numerical study of flow over a circular cylinder in oscillatory flows with zero-mean and non-zero-mean velocities," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 144, pp. 42-52, 2015.

- [23] G. Keulegan and L. Carpenter, "Forces on cylinders and plates in an oscillating fluid," *Journal of Research of the National Bureau of Standards* 60, vol. 60, pp. 423-440, 1958.
- [24] T. Sarpkaya, "In-line and Transverse Forces on Smooth and Sand-roughened Cylinders in Oscillatory Flow at High Reynolds Numbers," *National Science Foundation*, 1976.

An Optimal Eighth-Order Family of Iterative Methods For Multiple Roots

Fiza Zafar, Saima Akram*, Nusrat Yasmin

Centre for Advanced Studies in Pure and Applied Mathematics,
Bahauddin Zakariya University, Multan, Pakistan

*Corresponding author:saimaakram@bzu.edu.pk

Abstract

In this paper, we propose a new family of efficient and optimal iterative methods for multiple roots with known multiplicity ($m \geq 1$). We use weight function approach involving one and two parameters to develop the new family. An extensive convergence analysis is discussed in order to demonstrate the optimal eighth-order convergence of the proposed scheme. Finally, numerical and dynamical tests are presented which confirm the theoretical results established in this paper and illustrate that the proposed family is efficient among the domain of multiple root finding methods.

Keywords: Nonlinear equations, multiple zeros, efficiency index, optimal iterative methods.

Introduction

The problem of solving nonlinear equation is recognized to be very old in history as many practical problems arising in nature are nonlinear. Various one-point and multi-point methods are presented to solve nonlinear equation or system of nonlinear equation [17, 18, 21]. The above cited methods are designed for the simple root of nonlinear equations but the behavior of these methods are not similar when dealing with multiple roots of nonlinear equations. The well known Newton's method with quadratic convergence for simple roots of nonlinear equations decays to first order when dealing with multiple roots of nonlinear equations. These problems lead to minor troubles such as greater computational cost and severe troubles such as no convergence at all. The prior knowledge of multiplicity of roots make it easier to deal with these difficulties. The anomalous behavior of the iterative methods while dealing with multiple roots is well known at least since 19th century when Schröder [20] developed a modification of classical Newton's method to preserve its 2nd order of convergence for multiple roots. The nonlinear equations with multiple roots commonly arise from different topics such as complex variables, fractional diffusion or image processing, applications to economics and statistics(Lévy distributions) etc. By knowing the practical nature of multiple root finders, various one-point and multi-point root solvers have been developed in recent past [1, 3, 4, 6-9, 11-15, 19, 24] but most of them are not optimal as defined by Kung and Traub [10] which states that an optimal without memory method can achieve its convergence order at most 2^n requiring $n+1$ evaluations of functions or derivatives. According to Ostrowski [17], if O is the convergence order of an iterative method and n is the total number of functional evaluations per iterative step then the index $E = O^{1/n}$ is known as efficiency index of an iterative method.

Sharma and Sharma [19] proposed the following optimal fourth order multiple root finder with known multiplicity m as follows:

$$y_n = x_n - \frac{2m}{m+2} \cdot \frac{f(x_n)}{f'(x_n)}, m > 1 \quad (1.1)$$

$$x_{n+1} = x_n - \frac{m}{8} \Phi(x_n) \frac{f(x_n)}{f'(x_n)},$$

$$\text{where } \Phi(x_n) = \left\{ (m^3 - 4m + 8) - (m+2)^2 \left(\frac{m}{m+2} \right)^m \frac{f'(x_n)}{f'(y_n)} \times 2(m-1)(m+2) \left(\frac{m}{m+2} \right)^m \frac{f'(x_n)}{f'(y_n)} \right\}.$$

Geum et al. in [7], presented a non-optimal family of two-point sixth-order methods to find multiple roots given as follows:

$$\begin{aligned} y_n &= x_n - m \cdot \frac{f(x_n)}{f'(x_n)}, m > 1, \\ x_{n+1} &= y_n - Q(r_n, s_n) \cdot \frac{f(y_n)}{f'(y_n)}, \end{aligned} \quad (1.2)$$

where, $r_n = \sqrt[m]{\frac{f(y_n)}{f(x_n)}}$, $s_n = \sqrt[m-1]{\frac{f'(y_n)}{f'(x_n)}}$ and $Q: \mathbb{C}^2 \rightarrow \mathbb{C}$ is holomorphic in a neighborhood of $(0,0)$.

Following is a special case of their family:

$$\begin{aligned} y_n &= x_n - m \cdot \frac{f(x_n)}{f'(x_n)}, n \geq 0, m > 1, \\ x_{n+1} &= y_n - m \left[1 + 2(m-1)(r_n - s_n) - 4r_n s_n + s_n^2 \right] \cdot \frac{f(y_n)}{f'(y_n)}. \end{aligned} \quad (1.3)$$

Another non-optimal family of three-point sixth-order methods for multiple roots by Geum et al. [8], is given as follows:

$$\begin{aligned} y_n &= x_n - m \cdot \frac{f(x_n)}{f'(x_n)}, m \geq 1, \\ w_n &= y_n - m \cdot G(p_n) \cdot \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= w_n - m \cdot K(p_n, v_n) \cdot \frac{f(x_n)}{f'(x_n)}, \end{aligned} \quad (1.4)$$

where, $p_n = \sqrt[m]{\frac{f(y_n)}{f(x_n)}}$ and $v_n = \sqrt[m]{\frac{f(w_n)}{f(x_n)}}$. The weight functions $Q: \mathbb{C} \rightarrow \mathbb{C}$ is analytic in a neighborhood of 0 and $K: \mathbb{C}^2 \rightarrow \mathbb{C}$ is holomorphic in a neighborhood of $(0,0)$. Following is a special case of the family (1.4):

$$\begin{aligned} y_n &= x_n - m \cdot \frac{f(x_n)}{f'(x_n)}, m \geq 1, \\ w_n &= y_n - m \cdot [1 + p_n + 2p_n^2] \cdot \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= w_n - m \cdot [1 + p_n + 2p_n^2 + (1 + 2p_n)v_n] \cdot \frac{f(x_n)}{f'(x_n)}. \end{aligned} \quad (1.5)$$

The families (1.2) and (1.4) require four functional evaluations to produce sixth order convergence

with the efficiency index $6^{\frac{1}{4}} = 1.5650$ and therefore are not optimal in the sense of Kung-Traub's conjecture [10].

Recently, Behl et al. [2] have proposed a family of optimal eighth order iterative methods for multiple roots given as:

$$\begin{aligned} y_n &= x_n - m \frac{f(x_n)}{f'(x_n)}, m \geq 1, \\ z_n &= y_n - u_n Q(h_n) \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= z_n - u_n t_n G(h_n, t_n) \frac{f(x_n)}{f'(x_n)}, \end{aligned} \quad (1.6)$$

where, the weight functions $Q: \mathbb{C} \rightarrow \mathbb{C}$ and $G: \mathbb{C}^2 \rightarrow \mathbb{C}$ are analytic functions in a neighborhoods

of (0) and $(0,0)$, respectively, with $u_n = \left(\frac{f(y_n)}{f(x_n)} \right)^{\frac{1}{m}}$, $h_n = \frac{u_n}{a_1 + a_2 u_n}$ and $t_n = \left(\frac{f(z_n)}{f(y_n)} \right)^{\frac{1}{m}}$,

being a_1 and a_2 are complex non-zero free parameters.

We take particular case (27) for $(a_1 = 1, a_2 = 1, G_{02} = 0)$ of the family by Behl et al. [2] and denote it by *BM* as follows:

$$\begin{aligned} y_n &= x_n - m \frac{f(x_n)}{f'(x_n)}, \\ z_n &= y_n - \left(m + 2h_n m + \frac{1}{2} h_n^2 (4m + 2m) \right) \frac{f(x_n)}{f'(x_n)} u_n \\ x_{n+1} &= z_n - \left(m + m t_n + 3m h_n^2 + m h_n (2 + 4t_n + h_n) \right) \frac{f(x_n)}{f'(x_n)} u_n t_n. \end{aligned} \quad (1.7)$$

Most recently, second optimal eighth order scheme have been proposed by Zafar et al. [22], which is given as follows:

$$\begin{aligned} y_n &= x_n - m \frac{f(x_n)}{f'(x_n)}, m \geq 1, \\ z_n &= y_n - m u_n H(u_n) \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= z_n - u_n t_n (B_1 + B_2 u_n) P(t_n) G(w_n) \frac{f(x_n)}{f'(x_n)}, \end{aligned} \quad (1.8)$$

where $B_1, B_2 \in \mathbb{R}$ are free parameters and the weight functions $H: \mathbb{C} \rightarrow \mathbb{C}$, $P: \mathbb{C} \rightarrow \mathbb{C}$ and $G: \mathbb{C} \rightarrow$

\mathbb{C} are analytic in the neighborhood of 0 with $u_n = \left(\frac{f(y_n)}{f(x_n)} \right)^{\frac{1}{m}}$, $t_n = \left(\frac{f(z_n)}{f(y_n)} \right)^{\frac{1}{m}}$ and

$$w_n = \left(\frac{f(z_n)}{f(x_n)} \right)^{\frac{1}{m}}.$$

From the eighth order family of Zafar et al. [22], we consider the following special case denoted by *ZM* :

$$\begin{aligned} y_n &= x_n - m \frac{f(x_n)}{f'(x_n)}, \\ z_n &= y_n - mu_n (6u_n^3 - u_n^2 + 2u_n + 1) \frac{f(x_n)}{f'(x_n)} \\ x_{n+1} &= z_n - mu_n t_n (1 + 2u_n)(1 + t_n)(1 + 2w_n) \frac{f(x_n)}{f'(x_n)}. \end{aligned} \quad (1.9)$$

Optimal iterative methods are more significant than the non-optimal ones, regarding their efficiency and convergence speed. Therefore, there was a need to develop optimal eighth-order schemes for finding multiple zeros ($m > 1$) as well as simple zeros ($m = 1$) because of their better efficiencies and order of convergence [17], in addition optimal schemes require a small number of iterations to obtain desired accuracy as compare to fourth and sixth-order methods of Sharma and Geum [7, 8, 19]. In this paper, our main concern is to find the optimal iterative methods for multiple root μ with known multiplicity $m \in \mathbb{N}$ of a sufficiently differentiable nonlinear function $f: I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ where I is an open interval. We, in here, develop an optimal eighth order zero finder for multiple roots with known multiplicity $m \geq 1$. The beauty of the method lies in the fact that developed scheme is simple to implement with minimum possible number of functional evaluations. The family requires four functional evaluations to obtain eighth-order convergence with the efficiency index $8^{\frac{1}{4}} = 1.6817$.

The rest of the paper is organized as follows: In Section 2, we propose a new family of optimal eighth-order iterative methods to find multiple roots of nonlinear equations and discuss its convergence analysis. Some special cases are given in Section 3. In Section 4, numerical performance and comparison of the proposed schemes with the existing ones are given, dynamical analysis is given in section 5. Concluding remarks are given in Section 6.

Development of the scheme

In this section, we propose a new family of eighth-order method for a known multiplicity $m \geq 1$ of the desired multiple root as follows:

$$\begin{aligned} y_n &= x_n - m \cdot \frac{f(x_n)}{f'(x_n)}, n \geq 0, \\ z_n &= y_n - m \cdot t \cdot H(t) \cdot \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= z_n - m \cdot t \cdot L(s, u) \cdot \frac{f(x_n)}{f'(x_n)}, \end{aligned}$$

where $t = \sqrt[m]{\frac{f(y_n)}{f(x_n)}}, s = \sqrt[m]{\frac{f(z_n)}{f(y_n)}}, u = \sqrt[m]{\frac{f(z_n)}{f(x_n)}}$, (2.1)

and the weight function $H: \mathbb{C} \rightarrow \mathbb{C}$ is analytic function in the neighborhood of 0 and weight function $L: \mathbb{C}^2 \rightarrow \mathbb{C}$ is holomorphic in the neighborhood of (0,0) and t, s and u are one-to- m multiple-valued functions.

In the next theorem, it is demonstrated that the proposed scheme (2.1) achieves the optimal eighth order of convergence without increasing the number of functional evaluations.

Theorem 1 Let $x = \mu$ (say) be a multiple zero with multiplicity $m \geq 1$ of an analytic function $f: \mathbb{C} \rightarrow \mathbb{C}$ in the region enclosing a multiple zero μ of $f(x)$. Then the family of iterative methods defined by (2.1) has eighth-order convergence when the following conditions are satisfied:

$$H_0 = 1, H_1 = 2, H_2 = -2, H_3 = 36, L_{00} = 0, L_{10} = 1, L_{01} = 2, L_{11} = 4, L_{20} = 2. \quad (2.2)$$

Then the proposed scheme (2.1) satisfies the following error equations:

$$e_{n+1} = \frac{1}{24m^7} \{c_1(c_1^2(11+m) - 2mc_2)((677+108m+7m^2)c_1^4 - 24m(9+m)c_1^2c_2 + 12m^2c_2^2 + 12m^2c_1c_3)e_n^8\} + O(e_n^9), \quad (2.3)$$

where $e_n = x_n - \mu$ and $c_k = \frac{m!}{(m+k)!} \frac{f^{(m+k)}(\mu)}{f^{(m)}(\mu)}, k = 1, 2, 3, \dots$

Proof. Let $x = \mu$ be a multiple zero of $f(x)$. Expanding $f(x_n)$ and $f'(x_n)$ about $x = \mu$ by the Taylor's series expansion (with the help of computer algebra software Mathematica), we obtain

$$f(x_n) = \frac{f^{(m)}(\mu)}{m!} e_n^m \left(1 + c_1 e_n + c_2 e_n^2 + c_3 e_n^3 + c_4 e_n^4 + c_5 e_n^5 + c_6 e_n^6 + c_7 e_n^7 + c_8 e_n^8 + O(e_n^9)\right), \quad (2.4)$$

and

$$f'(x_n) = \frac{f^{(m)}(\mu)}{m!} e_n^{m-1} \{m + c_1(m+1)e_n + c_2(m+2)e_n^2 + c_3(m+3)e_n^3 + c_4(m+4)e_n^4 + c_5(m+5)e_n^5 + c_6(m+6)e_n^6 + c_7(m+7)e_n^7 + c_8(m+8)e_n^8 + O(e_n^9)\}, \quad (2.5)$$

respectively. By using the above expressions (2.4) and (2.5) in the first substep of (2.1), we obtain

$$y_n - \mu = \frac{c_1 e_n^2}{m} + \frac{(2c_2 m - c_1^2(m+1))e_n^3}{m^2} + \sum_{k=0}^4 G_k e_n^{k+4} + O(e_n^9), \quad (2.6)$$

where $G_k = G_k(m, c_1, c_2, \dots, c_8)$ are expressed in terms of m, c_1, c_2, \dots, c_8 where the two coefficients G_0 and G_1 can be explicitly written as

$$G_0 = \frac{1}{m^3} \{3c_3 m^2 + c_1^3(m+1)^2 - c_1 c_2 m(3m+4)\} \text{ and}$$

$$G_1 = -\frac{1}{m^4} \{c_1^4(m+1)^3 - 2c_2 c_1^2 m(2m^2 + 5m + 3) + 2c_3 c_1 m^2(2m+3) + 2m^2(c_2^2(m+2) - 2c_4 m)\}, \text{ etc.}$$

With the help of Taylor's series expansion, we obtain

$$f(y_n) = f^{(m)}(\mu) e_n^{2m} \left[\frac{\left(\frac{c_1}{m}\right)^m}{m!} + \frac{(2mc_2 - (m+1)c_1^2)\left(\frac{c_1}{m}\right)^m e_n}{c_1 m!} + \sum_{k=0}^4 \bar{G}_k e_n^{k+2} + O(e_n^9) \right]. \quad (2.7)$$

By using the expressions (2.4) and (2.7), we get

$$u = \frac{c_1 e_n}{m} + \frac{(2mc_2 - (m+2)c_1^2)e_n^2}{m^2} + \psi_1 e_n^3 + \psi_2 e_n^4 + \psi_3 e_n^5 + O(e_n^6), \quad (2.8)$$

where,

$$\begin{aligned}\psi_1 &= \frac{1}{2m^3} [c_1^3(2m^2 + 7m + 7) + 6c_3m^2 - 2c_1c_2m(3m + 7)], \\ \psi_2 &= -\frac{1}{6m^4} [c_1^4(6m^3 + 29m^2 + 51m + 34) - 6c_2c_1^2m(4m^2 + 16m + 17) + 12c_1c_3m^2(2m + 5) \\ &\quad + 12m^2(c_2^2(m + 3) - 2c_4m)], \\ \psi_3 &= \frac{1}{24m^5} [-24m^3(c_2c_3(5m + 17) - 5c_5m) + 12c_3c_1^2m^2(10m^2 + 43m + 49) + 12c_1m^2\{c_2^2(10m^2 + 47m + 53) \\ &\quad - 2c_4m(5m + 13)\} - 4c_2c_1^3m(30m^3 + 163m^2 + 306m + 209) + c_1^5(24m^4 + 146m^3 + 355m^2 + 418m + 209)]\end{aligned}$$

Expanding Taylor series of $H(t)$ about 0 we have:

$$H(t) = H_0 + H_1t + \frac{H_2}{2!}t^2 + \frac{H_3}{3!}t^3 + O(e_n^4) \quad (2.9)$$

where $H_j = H^j(0)$ for $0 \leq j \leq 3$. Inserting the expressions (2.6)-(2.9) in the second substep of scheme (2.1), we have

$$\begin{aligned}z_n &= \mu + \frac{-(1+H_0)c_1e_n^2}{m} - \frac{(1+H_1+m-H_0(3+m)c_1^2)+2(-1+H_0)mc_2)e_n^3}{m^2} \\ &\quad + \frac{1}{2m^3} [(2+10H_1-H_2+4m+4H_1m+2m^2-H_0(13+11m+2m^2))c_1^3 \\ &\quad + 2m(-4-4H_1-3m+H_0(11+3m)c_1c_2-6(-1+H_0)m^2c_3)e_n^4] + z_5e_n^5 \\ &\quad + z_6e_n^6 + z_7e_n^7 + O(e_n^8).\end{aligned}$$

By selecting $H_0 = 1$ and $H_1 = 2$ we obtained

$$z_n = \mu + \frac{(c_1^3(9-H_2+m)-2mc_1c_2)}{2m^3}e_n^4 + z_5e_n^5 + z_6e_n^6 + z_7e_n^7 + O(e_n^8), \quad (2.10)$$

where

$$\begin{aligned}z_5 &= -\frac{1}{6m^4} \{c_1^4(125+H_3+84m+7m^2-3H_2(7+3m)+6m(-3H_2+4(7+m))c_1^2c_2+12c_2^2m^2+12c_2c_1m), \\ z_6 &= \frac{1}{24m^5} \{1507+1850m+677m^2+46m^3+4H_3(9+4m)-6H_2(59+53m+12m^2))c_1^5 \\ &\quad -4m(925+8H_3+594m+53m^2-3H_2(53+21m)c_1^3c_2+12m^2(83-9H_2+13m)c_1^2c_3-168m^3c_2c_3 \\ &\quad +12m^2c_1(115-12H_2+17m)c_2^2-6mc_4)\end{aligned}$$

and

$$\begin{aligned}z_7 &= -\{12c_1^2c_3m^2(36\beta+13m+11)+(37-168c_2c_3m^3+4c_1^3c_2m(96\beta^2+252\beta+53m^2+18(14\beta+5)m) \\ &\quad +12c_1m^2(c_2^2(48\beta+17m+19)-6c_4m)\}.\end{aligned}$$

Now, again by using the Taylor's series expansion for (2.10), we have

$$f(z_n) = f^{(m)}(\mu)e_n^{4m} \frac{2^{-m} \left(\frac{c_1^3(9-H_2+m)-2mc_1c_2}{m^3} \right)^m}{m!} - \frac{\left(2^{-m} \left(\frac{c_1^3(9-H_2+m)-2mc_1c_2}{m^3} \right)^{m-1} \rho_0 \right)}{3(m^3m!)} e_n \quad (2.11)$$

$$+ \sum_{j=0}^7 \overline{H}_j e_n^{j+1} + O(e_n^9),$$

where

$$\rho_0 = c_1^4(125 + H_3 + 84m + 7m^2 - 3H_2(7 + 3m))c_1^4 - 6m(-3H_2 + 4(7 + m))c_1^2c_2 + 12m^2c_2^2 + 12c_3c_1m^2).$$

With the help of expressions (2.4) and (2.11), we have

$$s = \frac{c_1^2(9 - H_2 + m) - 2mc_2}{2m^2} e_n^2 + \rho_1 e_n^3 + \rho_2 e_n^4 + \rho_3 e_n^5 + O(e_n^6), \quad (2.12)$$

where,

$$\rho_1 = -(1/(6m^3))\{c_1^3(98 + H^3 + 4m^2 + 54m - 6H^2(3 + m) - 12m(9 - H^2 + m)c_1c_2 + 12m^2c_3\},$$

$$\rho_2 = (1/(24m^4))899 + 1002m + 313m^2 + 18m^3 + 4H_3(8 + 3m) - 6H_2(43 + 33m + 6m^2)c_1^4 - 12m(167 + 2H_3 + 87m + 6m^2 - H_2(33 + 10m))c_1^2c_2 + 24m^2(26 - 3H_2 + 3m)c_1c_3 + 12m^2(c_2^2(35 - 4H_2 + 3m) - 6mc_4)$$

and

$$\begin{aligned} \rho_3 = & -(1/(60m^5))[-4257 - 7270m - 4455m^2 - 101m^3 - 48m^4 - 10H_3(37 + 30m + 6m^2) + 30H_2(60 + 75m + 31m^2 + 4m^3)c_1^5 + 10m(1454 + 60H_3 + 1548m + 21H_3m + 454m^2 + 24m^3 - 18H_2(25 + 18m + 3m^2))c_1^3c_2 - 30m^2(234 + 3H_3 + 118m + 8m^2 - 2H_2(24 + 7m))c_1^2c_3 - 60m^2c_1(141 + 2H_3 + 67m + 4m^2 - 2H_2(15 + 4m))c_2^2 + 2(-17 + 2H_2 - 2m)mc_4 - 120m^3(-25 + 3H_2 - 2m)c_2c_3 + 2mc_5\} + ((1/(720m^6))((102047 + 180H_2^2 + 204435m + 187055m^2 + 81525m^3 + 14738m^4 + 600m^5 + 40H_3(389498m + 214m^2 + 30m^3) - 45H_2(1223 + 2030m + 1353m^2 + 394m^3 + 40m^4)) - 30m(13629 + 22190m + 12915m^2 + 2746m^3 + 120m^4 + 16H_3(83 + 64m + 12m^2) - 6H_2(1015 + 1209m + 470m^2 + 56m^3)) + 120m^2(2063 + 2088m + 589m^2 + 30m^3 + H^3(88 + 30m) - 18H_2 + (36 + 25m + 4m^2)) + 80m^2(2323 + 2348m + 635m^2 + 30m^3 + 4H_3(289m) - 3H_2(259 + 173m + 26m^2)) - 2m(303 + 4H_3 + 149m + 10m^2 - 9H_2(7 + 2m)) - 720m^3((393 + 6H_3 + 178m + 10m^2 - H_2(87 + 22m))) + (-42 + 5H_2 - 5m)mc_5) + 20m^3((-473 - 8H^3 - 195m - 10m^2 + 12H^2(9 + 2m))c_2c_3 + 6m(65 - 8H_2 + 5m)c_2 + 3m((71 - 9H_2 + 5m)c_{10}mc_6). \end{aligned}$$

Since it is clear from (2.8) that u is of order e_n . Therefore, we can expand weight function $L(s, u)$ in the neighborhood of origin by Taylor's series expansion as follows:

$$L(s, u) = L_{00} + sL_{10} + uL_{01} + suL_{11} + \frac{s^2}{2!}L_{20}, \quad (2.13)$$

where $L_{i,j} = \frac{1}{i!j!} \frac{\partial^{i+j}}{\partial s^i \partial u^j} L(0,0)$. By using the expressions (2.4)-(2.13) in the proposed scheme

(2.1), we have

$$e_{n+1} = M_2 e_n^2 + M_3 e_n^3 + M_4 e_n^4 + M_5 e_n^5 + M_6 e_n^6 + M_7 e_n^7 + O(e_n^8), \quad (2.14)$$

where the coefficients $M_i (2 \leq i \leq 7)$ depends generally on m and the parameters $L_{i,j}$. For obtaining at least fifth-order convergence, we have to choose $L_{00} = 0, L_{10} = 1$ and get

$$e_{n+1} = \frac{((-2 + L_{01})c_1^2((-9 + H_2 - m)c_1^2 + 2mc_2))}{2m^4} e_n^5 + \overline{M}_6 e_n^6 + \overline{M}_7 e_n^7 + O(e_n^8)$$

where the coefficients $\overline{M}_i (6 \leq i \leq 7)$ depends generally on m and the parameters $L_{i,j}$. To obtain

eighth order of convergence we choose the following values of parameters:

$$H_2 = -2, H_3 = 36, L_{00} = 0, L_{10} = 1, L_{01} = 2, L_{20} = 2, L_{11} = 4 \quad (2.15)$$

which leads us to the following error equation:

$$e_{n+1} = \frac{1}{24m^7} [c_1(c_1^2(11+m) - 2mc_2)((677+108m+7m^2)c_1^4 - 24m(9+m)c_1^2c_2 + 12m^2c_2^2 + 12m^2c_1c_3)]e_n^8 + O(e_n^9) \quad (2.16)$$

The above asymptotic error constant (2.16) reveals that the proposed scheme (2.1) reaches at optimal eighth-order convergence by using only four functional evaluations (using. $f(x_n), f'(x_n), f(y_n)$ and $f(z_n)$) per iteration.

Special Cases of Weight Functions

From Theorem 1, several choices of weight functions can be obtained, we have considered the following:

Case 1: The polynomial form of the weight function satisfying conditions (2.2) can be represented as:

$$\begin{aligned} H(t) &= 1 + 2t - t^2 + 6t^3 \\ L(s, u) &= s + 2u + 4su + s^2 \end{aligned} \quad (3.1)$$

A particular iterative method related to (3.1) is given by:

SM-1:

$$\begin{aligned} y_n &= x_n - m \cdot \frac{f(x_n)}{f'(x_n)}, n \geq 0, \\ z_n &= y_n - m \cdot t \cdot (1 + 2t - t^2 + 6t^3) \cdot \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= z_n - m \cdot t \cdot (s + s^2 + 2u + 4su) \cdot \frac{f(x_n)}{f'(x_n)} \\ \text{where } t &= \sqrt[m]{\frac{f(y_n)}{f(x_n)}}, s = \sqrt[m]{\frac{f(z_n)}{f(y_n)}}, u = \sqrt[m]{\frac{f(z_n)}{f(x_n)}} \end{aligned} \quad (3.2)$$

Case 2: The second suggested form of the weight functions in which $k_f(t)$ is constructed using rational weight function satisfying conditions (2.2) is given by:

$$\begin{aligned} H(t) &= \frac{1 + 8t + 11t^2}{1 + 6t} \\ L(s, u) &= s + 2u + 4su + s^2 \end{aligned} \quad (3.3)$$

The corresponding iterative method (3.3) can be presented as:

SM-2:

$$\begin{aligned}
 y_n &= x_n - m \cdot \frac{f(x_n)}{f'(x_n)}, n \geq 0, \\
 z_n &= y_n - m \cdot t \cdot \left(\frac{1+8t+11t^2}{1+6t} \right) \frac{f(x_n)}{f'(x_n)}, \\
 x_{n+1} &= z_n - m \cdot t \cdot (s + s^2 + 2u + 4su) \cdot \frac{f(x_n)}{f'(x_n)}, \\
 \text{where } t &= \sqrt[m]{\frac{f(y_n)}{f(x_n)}}, s = \sqrt[m]{\frac{f(z_n)}{f(y_n)}}, u = \sqrt[m]{\frac{f(z_n)}{f(x_n)}}
 \end{aligned} \tag{3.4}$$

Case 3: The third suggested form of the weight function in which $K_f(t)$ is constructed using trigonometric weight satisfying conditions (2.2) is given by:

$$\begin{aligned}
 H(t) &= \frac{5+18t}{5+18t-11t^2} \\
 L(s, u) &= s + 2u + 4su + s^2
 \end{aligned} \tag{3.5}$$

The corresponding iterative method obtained using (3.5) is given by:

SM-3:

$$\begin{aligned}
 y_n &= x_n - m \cdot \frac{f(x_n)}{f'(x_n)}, n \geq 0, \\
 z_n &= y_n - m \cdot t \cdot \left(\frac{5+18t}{5+18t-11t^2} \right) \frac{f(x_n)}{f'(x_n)}, \\
 x_{n+1} &= z_n - m \cdot t \cdot (s + s^2 + 2u + 4su) \cdot \frac{f(x_n)}{f'(x_n)}, \\
 \text{where } t &= \sqrt[m]{\frac{f(y_n)}{f(x_n)}}, s = \sqrt[m]{\frac{f(z_n)}{f(y_n)}}, u = \sqrt[m]{\frac{f(z_n)}{f(x_n)}}.
 \end{aligned} \tag{3.6}$$

Numerical tests

In this section, we show the performance of the presented iterative family (2.1) by carrying out some numerical tests and comparing the results with existing method for multiple roots. All the numerical computations have been performed in Maple 16 programming package using 1000 significant digits of minimum number of precision. In that case μ is not exact, it is replaced by a more accurate value which has more number of significant digits than the assigned precision. The test functions along with their roots μ and multiplicity m are listed in Table 1 [16]. The proposed methods SM-1 (3.2), SM-2 (3.4) and SM-3(3.6) are compared with the methods of Geum et al. given in (1.3) and (1.5) denoted by GKM-1 and GKM-2 and with method of Bhel given in (1.7) denoted by BM and Zafar et. al method given in (1.9) denoted by ZM respectively. Tables 2-8 display the errors of approximations to the sought zeros ($|x_n - \mu|$) produced by different methods at the first three iterations, where $E(-i)$ denotes $E \times 10^{-i}$. The initial approximation x_0 for each test function and computational order of convergence (COC) is also included in these

tables, which is computed by the following expression [23]:

$$COC \approx \frac{\log|(x_{k+1} - \mu)/(x_k - \mu)|}{\log|(x_k - \mu)/(x_{k-1} - \mu)|}.$$

It is observed that, the performance of new method SM-2 is same as method of BM for function f_1 and better than method of ZM for function f_2 . The newly developed schemes SM-1, SM-2 and SM-3 are not only convergent but also their speed of convergence is better than methods of GKM-1 and GKM-2. On the other hand methods of ZM and BM show divergence for function f_3 . For f_4 , f_5 , f_6 and f_7 the newly developed schemes newly developed schemes SM-1, SM-2 and SM-3 are comparable with methods of ZM and BM. Hence, we conclude that the proposed family is comparable and robust among existing methods for multiple roots.

Table 1: Test functions

Test Functions	Exact root μ	Multiplicity m
$f_1(x) = (\cos(\frac{\pi x}{2}) + x^2 - \pi)^5$	2.034724896...	5
$f_2(x) = (e^x + x - 20)^2$	2.842438953...	2
$f_3(x) = (\ln x + \sqrt{(x^4 + 1)} - 2)^9$	1.222813963...	9
$f_4(x) = (\cos x - x)^3$	0.7390851332...	3
$f_5(x) = ((x-1)^3 - 1)^{50}$	2.0	50
$f_6(x) = (x^3 + 4x^2 - 10)^6$	1.365230013...	6
$f_7(x) = (8xe^{-x^2} - 2x - 3)^8$	-1.7903531791...	8

Table 2: Comparison of different methods for multiple roots

$f_1(x), x_0=2.5$							
	GKM-1	GKM-2	SM-1	SM-2	SM-3	ZM	BM
$ x_1 - \mu $	6.83(-4)	1.11(-3)	2.15(-4)	1.87(-4)	2.03(-4)	1.52(-4)	1.84(-4)
$ x_2 - \mu $	3.42(-14)	2.53(-18)	2.37(-29)	3.53(-30)	1.25(-29)	9.69(-31)	2.89(-30)
$ x_3 - \mu $	2.13(-55)	3.58(-106)	5.28(-299)	5.71(-236)	2.53(-231)	2.56(-240)	1.05(-236)
COC	4.00	6.00	8.00	8.00	8.00	8.00	8.00

Table 3: Comparison of different methods for multiple roots

$f_2(x), x_0=3.0$							
	GKM-1	GKM-2	SM-1	SM-2	SM-3	ZM	BM
$ x_1 - \mu $	1.18(-7)	5.27(-6)	2.33(-7)	1.21(-7)	1.90(-7)	1.40(-7)	1.16(-7)
$ x_2 - \mu $	2.62(-37)	1.15(-32)	1.30(-53)	2.21(-56)	1.99(-54)	1.30(-55)	1.57(-56)
$ x_3 - \mu $	3.07(-221)	1.25(-192)	1.19(-423)	2.67(-446)	2.87(-430)	7.37(-440)	1.73(-447)
COC	4.00	6.00	8.00	8.00	8.00	8.00	8.00

Table 4: Comparison of different methods for multiple roots

$f_3(x), x_0=3.0$							
	GKM-1	GKM-2	SM-1	SM-2	SM-3	ZM	BM
$ x_1 - \mu $	5.50(-1)	4.29(-2)	1.81(-2)	1.75(-2)	1.79(-2)	*	*
$ x_2 - \mu $	3.99(-7)	8.77(-10)	2.82(-15)	9.58(-16)	2.04(-15)	*	*
$ x_3 - \mu $	1.13(-27)	7.51(-56)	2.06(-117)	8.21(-122)	6.49(-119)	*	*
COC	4.00	6.00	8.00	8.00	8.00	*	*

“*” stands for divergence

Table 5: Comparison of different methods for multiple roots

$f_4(x), x_0=1.0$							
	GKM-1	GKM-2	SM-1	SM-2	SM-3	ZM	BM
$ x_1 - \mu $	2.77(-4)	2.55(-5)	6.78(-8)	5.45(-8)	6.29(-8)	4.90(-8)	5.15(-8)
$ x_2 - \mu $	3.28(-14)	6.83(-36)	7.95(-60)	8.55(-61)	3.83(-60)	4.06(-61)	4.91(-61)
$ x_3 - \mu $	5.86(-49)	2.51(-213)	2.82(-475)	3.11(-483)	7.18(-478)	8.99(-486)	3.36(-485)
COC	3.50	6.00	8.00	8.00	8.00	7.99	7.99

Table 6: Comparison of different methods for multiple roots

$f_5(x), x_0=2.1$							
	GKM-1	GKM-2	SM-1	SM-2	SM-3	ZM	BM
$ x_1 - \mu $	7.68(-5)	1.12(-5)	7.58(-7)	4.85(-7)	6.52(-7)	4.77(-7)	4.65(-7)
$ x_2 - \mu $	3.49(-17)	5.33(-29)	3.70(-47)	4.10(-49)	8.82(-48)	5.66(-49)	2.72(-49)
$ x_3 - \mu $	1.46(-66)	6.11(-169)	2.82(-369)	1.06(-385)	9.93(-375)	2.22(-384)	3.79(-387)
COC	3.99	6.00	8.00	8.00	8.00	7.99	7.99

Table 7: Comparison of different methods for multiple roots

$f_6(x), x_0=3.0$							
	GKM-1	GKM-2	SM-1	SM-2	SM-3	ZM	BM
$ x_1 - \mu $	5.44(-2)	1.01(-1)	5.40(-2)	5.30(-2)	5.36(-2)	4.36(-2)	5.39(-2)
$ x_2 - \mu $	7.40(-7)	5.37(-7)	1.10(-10)	8.60(-11)	8.60(-11)	1.36(-11)	4.92(-11)
$ x_3 - \mu $	3.54(-26)	1.86(-38)	5.28(-80)	576(-81)	5.76(-81)	1.80(-87)	3.14(-83)
COC	3.97	5.96	8.00	7.98	7.97	7.97	7.97

Table 8: Comparison of different methods for multiple roots

$f_7(x), x_0=-1.2$							
	GKM-1	GKM-2	SM-1	SM-2	SM-3	ZM	BM
$ x_1 - \mu $	2.65(-3)	2.15(-3)	4.38(-7)	4.24(-4)	4.32(-7)	3.41(-4)	4.26(-4)
$ x_2 - \mu $	7.24(-12)	9.63(-17)	4.44(-27)	1.11(-27)	3.11(-27)	3.58(-28)	1.14(-27)
$ x_3 - \mu $	4.05(-46)	7.81(-97)	4.97(-211)	2.55(-216)	2.28(-212)	5.27(-220)	3.06(-216)
COC	4.00	6.00	8.00	8.00	8.00	7.99	7.99

Dynamical analysis

Regarding the stability comparison, we use the routines presented in [5] for plotting the dynamical

planes corresponding to each method (SM-1, SM-2, S-M3, BM and ZM) for the non-linear functions $f_1, f_2, f_3, f_4, f_5, f_6, f_7$. For this, we define a mesh of 400×400 points, as each point of the mesh is an initial guess for the analyzed method on the specific nonlinear function. If the sequence of iteration method reaches (closer than 10^{-3}) the multiple root in less than 80 iterations, then this point is painted in orange color; if the iterate converges to another thing (strange fixed points, cycles, etc.) then the point is painted black. The multiple root is represented in the different figures by a white star.

We observe from Figures 1–7 that the only basin of attraction is that of the multiple root (that is, the set of initial points converging to it fills all the plotted region of the complex plane), plotted in orange in the figures; although in general, convergence to other roots, divergence or even convergence to other fixed points that are not roots of the non-linear function (known as strange fixed points), can appear. We see in the figures, that the orange region is more bigger and brighter for the proposed schemes SM1, SM2 and SM3 than the regions of methods BM and ZM for all examples, that confirms their stability and fast convergence speed.

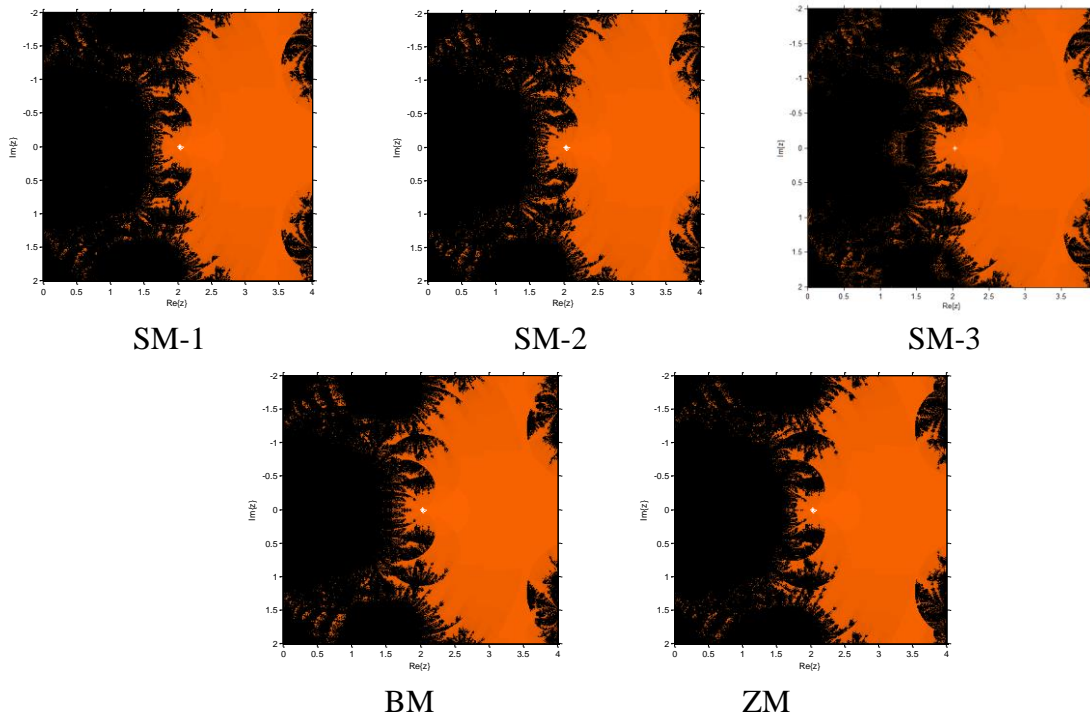


Fig. 1: Basins of attraction of different methods for f_l

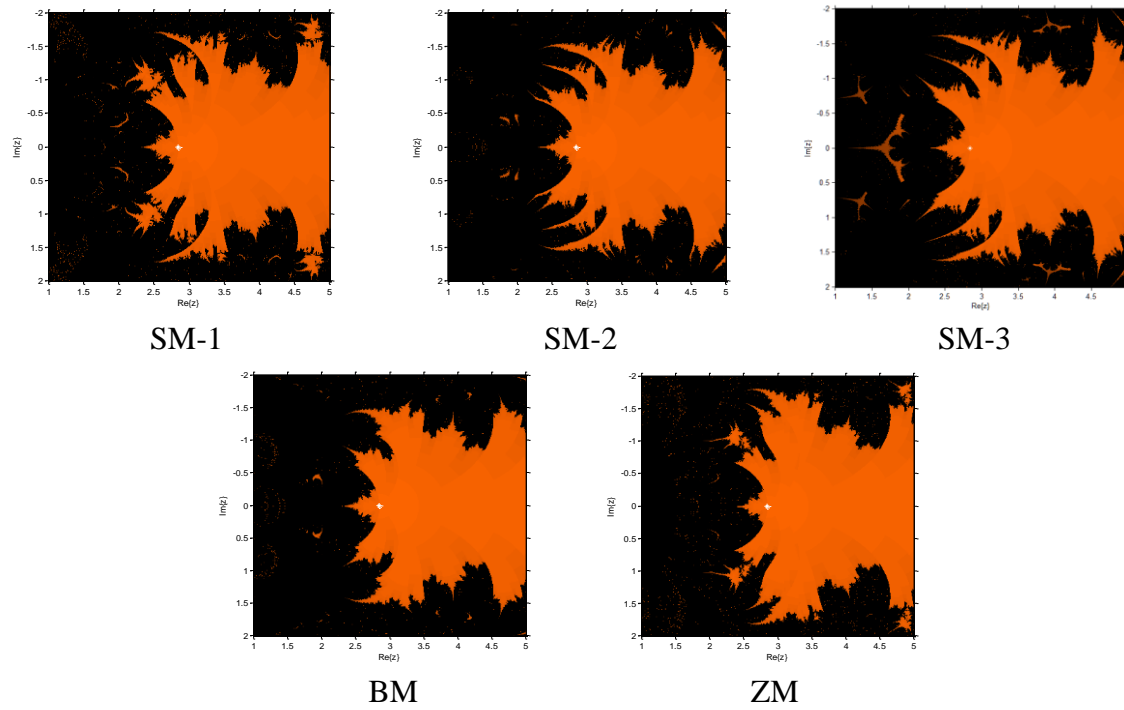


Fig. 2: Basins of attraction of different methods for f_2

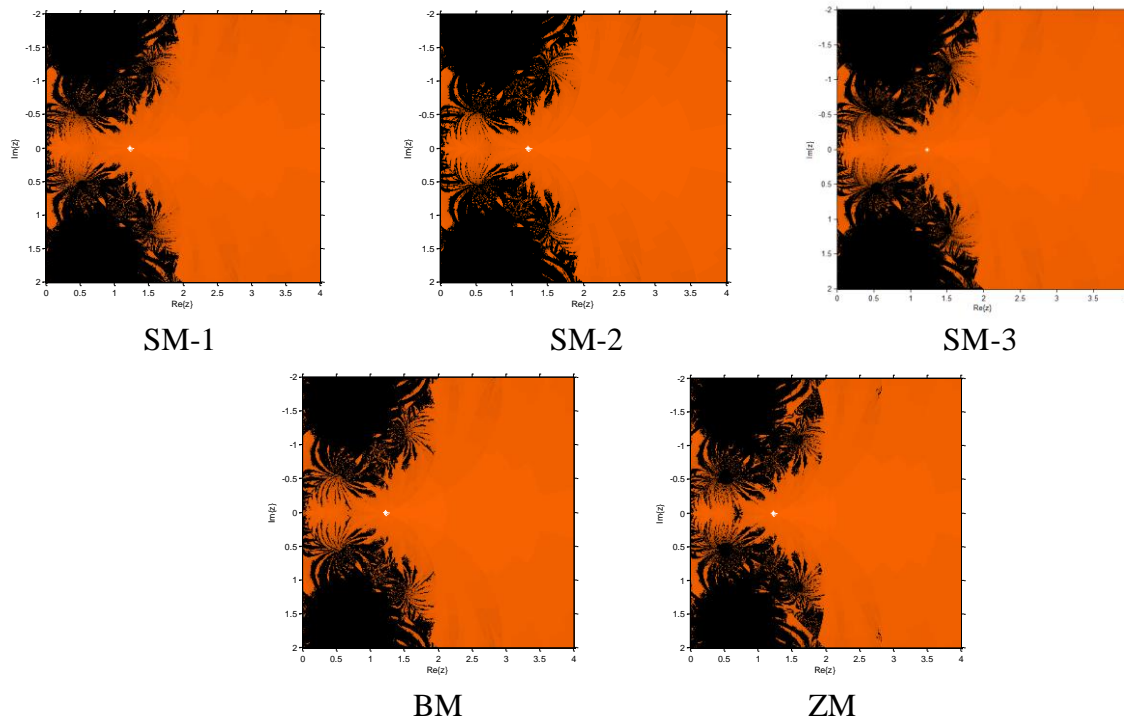


Fig. 3: Basins of attraction of different methods for f_3

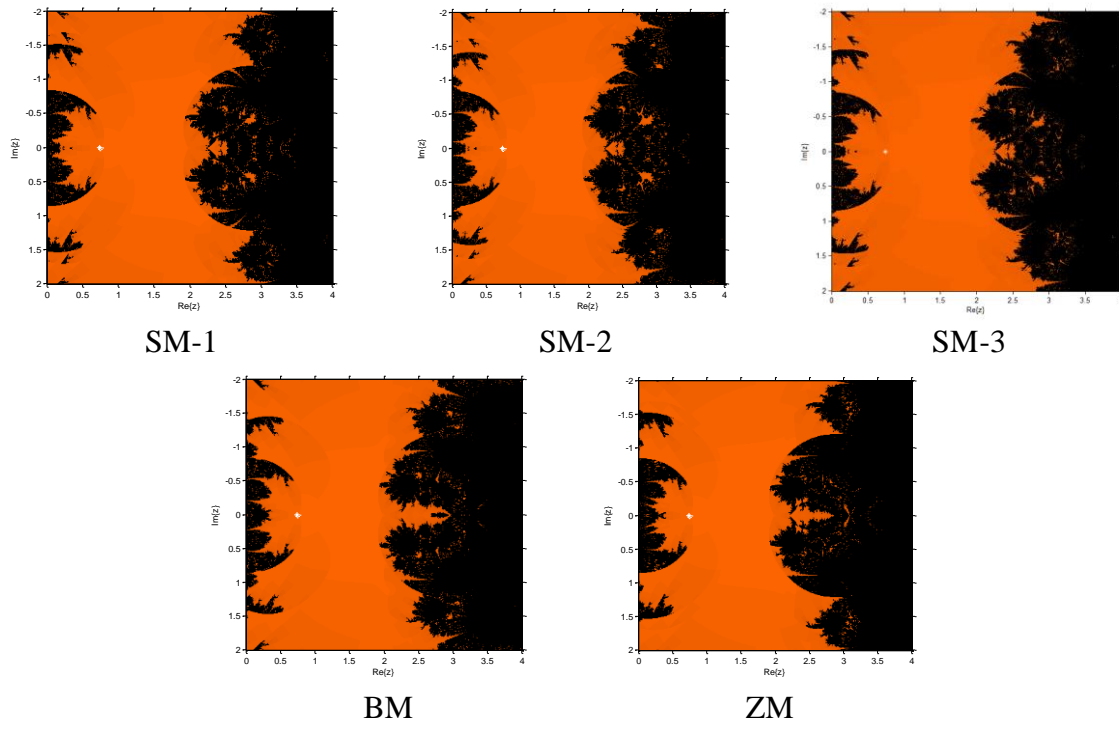


Fig. 4: Basins of attraction of different methods for f_4

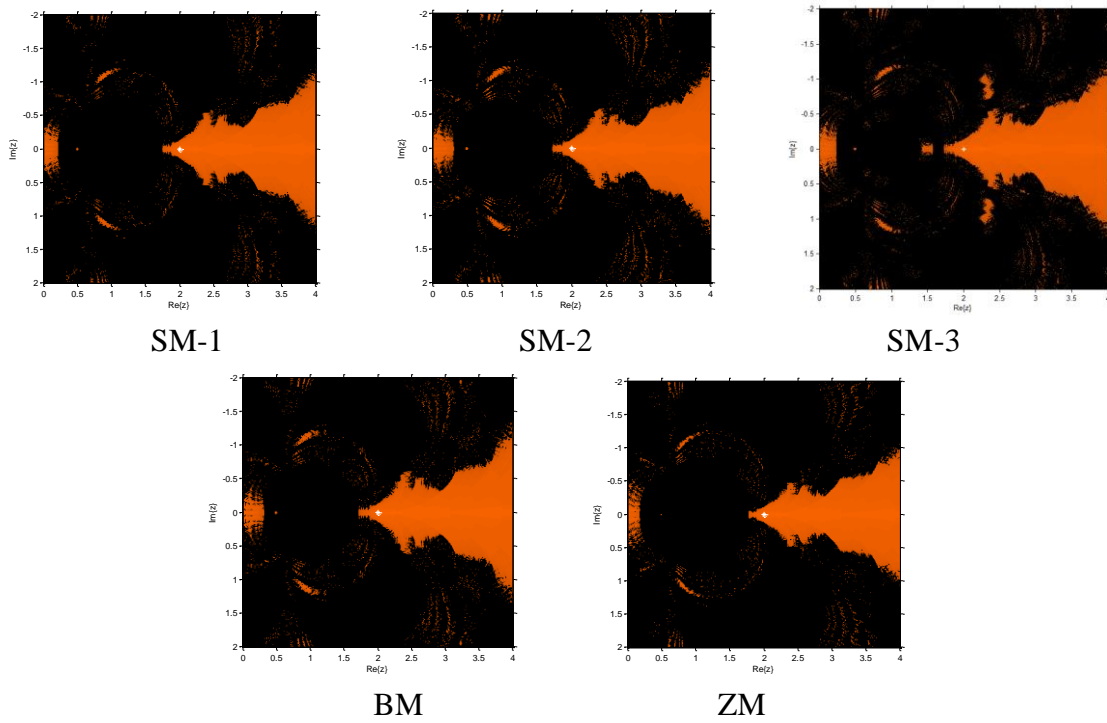


Fig. 5: Basins of attraction of different methods for f_5

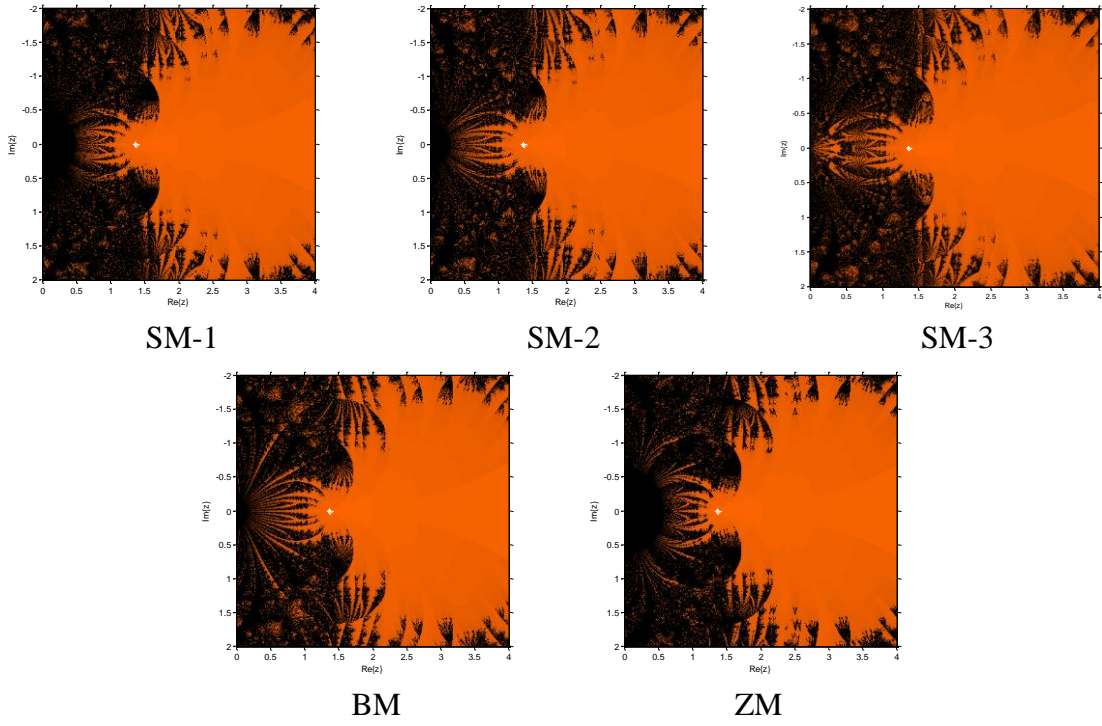


Fig. 6: Basins of attraction of different methods for f_6

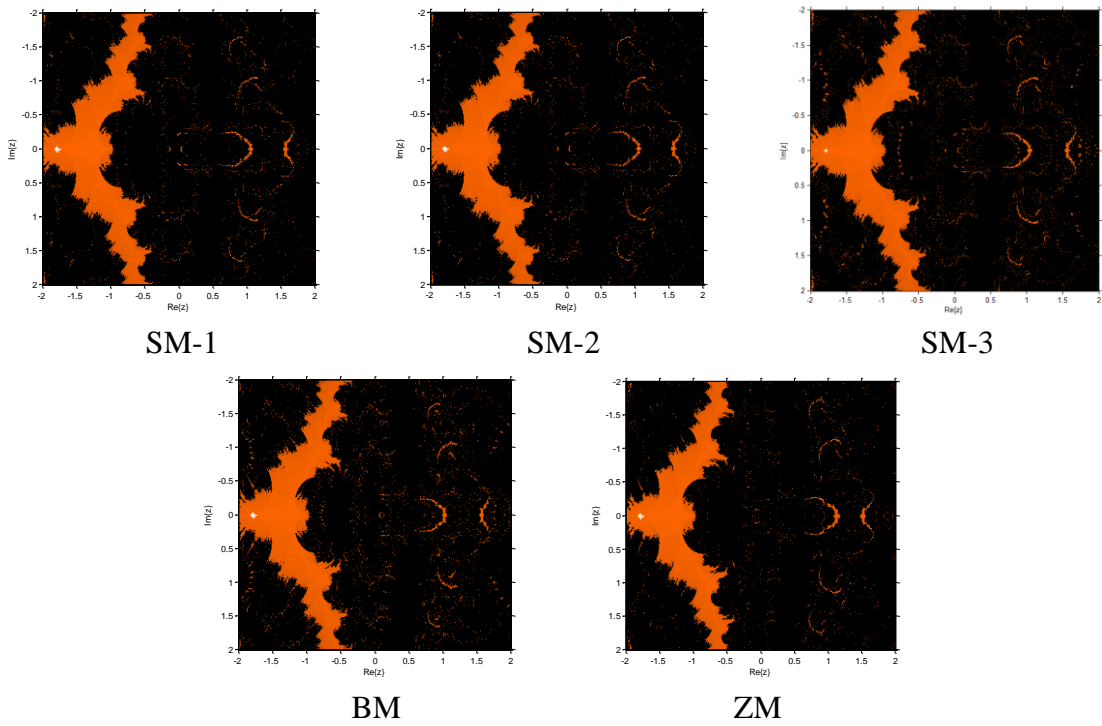


Fig. 7: Basins of attraction of different methods for f_7

Conclusion

In this paper, we present a new family of optimal eighth-order methods to find multiple roots of nonlinear equations. An extensive convergence analysis is done which verifies that the new family is optimal eighth order convergent. The proposed family requires four functional evaluations to obtain optimal eighth-order convergence with the efficiency index $8^{\frac{1}{4}} = 1.6817$ which is higher than the efficiency index of any of the methods for multiple roots and of the families of Geum et al. [7, 8]. Finally, numerical and dynamical tests confirm the theoretical results and show that the three members SM-1, SM-2 and SM-3 of the new family are better than existing methods for multiple roots. Hence, the proposed family is efficient among the domain of multiple root finding methods.

Conflict of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] A. Singh, P. Jaiswal, An efficient family of optimal fourth-order iterative methods for finding multiple roots of nonlinear equations, Proceedings of National Academy of Sciences, India, Section A 85(3), July (2015).
- [2] R. Behl, A. Cordero, S.S. Motsa, J.R. Torregrosa, An eighth-order family of optimal multiple root finders and its dynamics. Numer. Algor. (2017). doi:10.1007/s11075-017-0361-6.
- [3] R. Behl, A. Cordero, S.S. Motsa, J. R. Torregrosa, V. Kanwar, An optimal fourth-order family of methods for multiple roots and its dynamics, Numer. Algor. 71(4) (2016)775–796.
- [4] J. Biazar and B. Ghanbari, A new third-order family of nonlinear solvers for multiple roots, Comput. Math. Appl. 59(10) (2010)3315–3319.
- [5] Chicharro, F.I., Cordero, A., Torregrosa, J.R.: Drawing dynamical and parameters planes of iterative families and methods. Sci. World J., 2013, Article ID 780153, 9 pages, (2013)
- [6] C. Chun, H. J. Bae, and B. Neta, New families of nonlinear third-order solvers for finding multiple roots, Comput. Math. Appl. 57(9) (2009)1574–1582.
- [7] Y.H. Geum, Y.I. Kim, B. Neta, A class of two-point sixth-order multiple-zero finders of modified double-Newton type and their dynamics, Appl. Math. Comput. 270 (2015)387–400.
- [8] Y.H. Geum, Y.I. Kim, B. Neta, A sixth-order family of three-point modified Newton-like multiple-root finders and the dynamics behind their extraneous fixed points, Appl. Math. Comput. 283 (2016)120–140.
- [9] Jose L. Hueso, Eulalia Martinez, Carles Treuel, Determination of multiple roots of nonlinear equation and applications, J. Math. Chem. 53(3), (2015).
- [10] H.T. Kung, J.F. Traub, Optimal order of one-point and multipoint iteration, J. Assoc. Comput. Mach. 21, (1974)643–651.
- [11] Siyul Lee, Hyeongmin Choe, On fourth-order iterative methods for multiple roots of nonlinear equation with high efficiency, J.Comput. Anal. Appl. Vol. 18(1),pg 109, Jan (2015).
- [12] R.I. Lin, H.M. Ren, Z. Šmarda, Q.B.Wu, Y. Khan and J.L. Hu, New families of third-order iterative methods for finding multiple roots, J. Appl. Math. Vol 2014, Article ID:812072, 9 Pages.
- [13] S. Li, L. Cheng, B. Neta, Some fourth-order nonlinear solvers with closed formulae for multiple roots, Comput. Math. Appl. 59(2010)126–135.
- [14] Najumudin Ahmad, Vimal Pratap Singh, Some New Three step Iterative methods for solving nonlinear equation using Steffensen's and Halley method, Brit. J. Math. Comp. Sci. 19(2) (2016)1-9. DOI:10.9734/BJMCS/2016/19322.
- [15] B. Neta, Extension of Murakami's high-order non-linear solver to multiple roots, Int. J. Comput. Math. 87(5) (2010)1023–1031.
- [16] B. Neta, C. Chun, M. Scott, On the development of iterative methods for multiple roots, Appl. Math. Comput. 224 (2013)358–361.
- [17] A. M. Ostrowski, Solution of equations and systems of equations, Academic Press, New York (1960).

- [18] M.S. Petkovic, B. Neta, L.D. Petkovic, J. Dzunic, Multipoint methods for solving nonlinear equations, Academic Press (2013).
- [19] J.R. Sharma, R. Sharma, Modified Jarratt method for computing multiple roots, Appl. Math. Comput. 217 (2010)878–881.
- [20] E. Schröder, Über unendlich viele algorithmen zur auflösung der gleichungen, Math. Annal. 2 (1870)317–365.
- [21] J.F. Traub, Iterative Methods for the solution of equations. Prentice-Hall, Englewood Cliffs (1964).
- [22] F. Zafar, A. Cordero, R. Quratlain, J.R.Torregrosa, Optimal iterative methods for finding multiple roots of nonlinear equations using free parameters, J. Math. Chem. (2017). <https://doi.org/10.1007/s10910-017-0813-1>.
- [23] S. Weerakoon and T. G. I. Fernando, A variant of Newton’s method with accelerated third-order convergence, Appl. Math. Lett. 13(8) (2000)87–93.
- [24] X. Zhou, X. Chen, Y. Song, Constructing higher-order methods for obtaining the multiple roots of nonlinear equations, J. Comput. Appl. Math. 235(2011)4199–4206.

Selective cell-based smoothed finite element method using 10-node tetrahedral elements for large deformation of nearly incompressible solids

*Yuki Onishi¹

¹Department of Systems and Control Engineering, Tokyo Institute of Technology, Japan.

*Corresponding and Presenting author: yonishi@a.sc.e.titech.ac.jp

Abstract

A novel smoothed finite element method (S-FEM) using 10-node tetrahedral (T10) elements, SelectiveCS-FEM-T10, is proposed. In the proposed method, each T10 element is divided into T4 subelements and the strain smoothing is performed only within each T10 element, meaning no strain smoothing across elements. Also, the proposed method utilizes the selective reduced integration (SRI) technique for the stress integration. As a result, the proposed method avoids volume locking and pressure checkerboarding in nearly incompressible materials. A few examples of analysis reveal that the proposed method has good accuracy and robustness in large deformation analyses of nearly incompressible materials.

Keywords: Smoothed finite element method, Tetrahedral element, Large deformation, Volumetric locking, Pressure checkerboarding, Reaction force oscillation.

Introduction

Because of the difficulties in generating good-quality hexahedral meshes for complex shapes, finite element analyses with tetrahedral meshes are often demanded. However, it is well known that the standard tetrahedral elements are less accurate than hexahedral elements and make hard to obtain reliable results. The simplest tetrahedral element, the standard 4-node tetrahedral (T4) element, causes issues of shear locking, volume locking, and pressure checkerboarding. Although the standard 10-node tetrahedron (T10) element can resolve the issue of shear locking, it can not resolve the other issues. Moreover, it brings an additional issue of nodal reaction force oscillation. The issues of volume locking and pressure checkerboarding appear not only in rubber materials but also in materials with near incompressibility such as viscoelastic and elastoplastic materials. It is known that these issues can not be resolved at all by using finer meshes. From such a background, researches on accurate tetrahedral elements to analyze nearly incompressible materials are still being carried out actively.

The most widely used formulation as a highly accurate tetrahedral element is the mixed (hybrid) element [1, 2] based on the mixed variational principle. Various hybrid T4 or T10 elements have been proposed, but none of them has resolved all the issues above together yet. In addition, since hybrid elements require additional unknowns such as pressure, they all give rise to incompatibility with the dynamic explicit method.

On the other hand, the smoothed finite element methods (S-FEM) [3, 4] has recently attracted attention as the highly accurate tetrahedral formulations based on the pure displacement method without no additional unknowns. S-FEM is a type of strain smoothing method, and there are several formulations varied with the domains for strain smoothing: NS-FEM at nodes, ES-FEM at element edges, CS-FEM at elements and so on. We have proposed SelectiveES/NS-FEM-T4 [5, 6] combining ES-FEM-T4 and NS-FEM-T4 with the selective reduced integration (SRI)

method and also F-barES-FEM-T4 [7, 8, 9, 10] combining them with the F-bar method [11]. In particular, F-barES-FEM-T4 has been proved to be a formulation that can resolve all the issues above in static analysis. However, since our previous methods require strain smoothing across elements, it is difficult to implement them as user-defined elements of general-purpose FEM codes, which is a critical problem in practical engineering.

In this research, we propose a new S-FEM formulation with T10 elements, SelectiveCS-FEM-T10, which does not perform strain smoothing across elements. Following the method of Ostien et al. [12], a dummy node is introduced at the center of each T10 element, and the element is divided into twelve T4 subelements. By performing strain smoothing only across the subelements within each element, it is possible to implement the proposed S-FEM formulation as a user-defined T10 element of general-purpose FEM codes. In addition, SRI is applied in stress integration to combine the deviatoric and hydrostatic stresses derived from two different ways of strain smoothing. As a result, the proposed method avoids all the issues above in the analyses with nearly incompressible materials. In this paper, the outline of the formulation of the proposed method is described in Section 2, and then some results of example analyses are presented to confirm the effectiveness of the proposed method in Section 3, followed by the conclusion in Section 4.

Methods

The method proposed in this paper (SelectiveCS-FEM-T10) is a type of cell-based smoothing finite element method (CS-FEM). One dummy node and twelve T4 subelements are introduced in each T10 element and the strain smoothing is performed across the subelements within each element. In contrast to the node-based S-FEM (NS-FEM), edge-based S-FEM (ES-FEM) or face-based S-FEM (FS-FEM), there is no strain smoothing across elements. The deviatoric stress is evaluated at each subelement using the smoothed strain. On the other hand, we regard the set of 12 subelements as a patch to calculate pressure in element, which is similar to the concept of F-bar Patch method [13, 14]. The final nodal force is calculated by combining them in the manner of the selective reduced integration (SRI).

Addition of a dummy node

A schematic diagram of a T10 element defined by SelectiveCS-FEM-T10 is shown in Fig. 1. The exterior 10 nodes (4 corner nodes and 6 intermediate nodes) are the same as the standard T10 element, but SelectiveCS-FEM-T10 has one additional dummy node at the element center. The position of the dummy node x_{10} is defined by the following equation as the average position of all intermediate nodes.

$$x_{10} = \frac{1}{6} \sum_{P=4}^9 x_P, \quad (1)$$

where x_P represents the position vector of the node P . Since the position of the dummy node is a dependent variable, the number of unknowns does not increase. The total degrees of freedom of each element is 30: 10 nodes \times 3 dimensions. Note that the edges may be bent at the intermediate nodes although the edges in Fig. 1 are all straight for simplicity.

Table 1: List of the node numbers composing each subelement. Each number denotes an intra-element subelement/node number.

Subelement Number	Node Number
0	0, 4, 6, 7
1	1, 5, 4, 8
2	2, 6, 5, 9
3	3, 7, 9, 8
4	4, 6, 7, 10
5	5, 4, 8, 10
6	6, 5, 9, 10
7	7, 9, 8, 10
8	6, 4, 5, 10
9	7, 8, 4, 10
10	8, 9, 5, 10
11	9, 7, 6, 10

Subdivision of element into subelements

An elements are subdivided into 12 T4 subelements using the 11 nodes including the dummy one as shown in Fig. 1. Table 1 shows the list of intra-element node numbers composing each subelement. Subelement 0 to 3 are located at the four corners of the element whereas Subelement 4 to 11 are located in the remaining octahedron. There are 6 edges on each subelement, and there are 30 edges in an element without duplication.

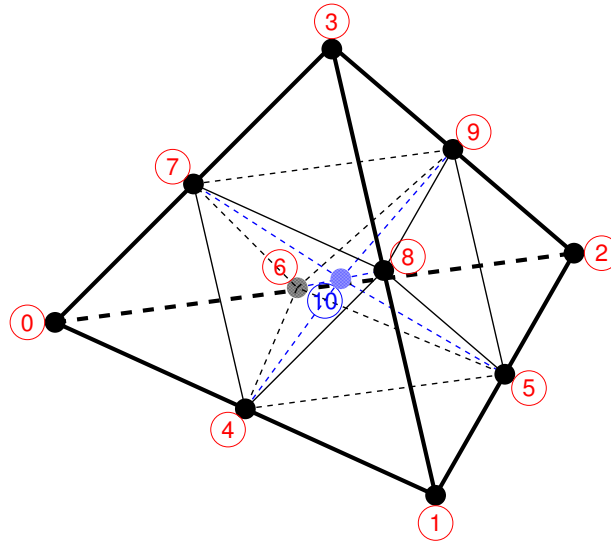


Figure 1: Schematic diagram of an element of SelectiveCS-FEM-T10. The numbers 0 to 10 denote the intra-element node numbers. Node 10 is a dummy node and is located at the average position of Node 4 to 9. An element is subdivided into 12 T4 subelements. There are 30 edges of the subelements in the element without duplication.

Shape functions of subelements and its derivatives

SelectiveCS-FEM-T10 is a T10 element but is formulated as a set of linear elements, not a quadratic elements. We consider each subelement as a standard T4 element and calculate a shape function $^{Sube}\mathbf{N}$ represented by volume coordinates for each subelement. In case of the subelement using the dummy node, the weight of the dummy node is distributed to the 6 intermediate nodes evenly. As a result, the term of the dummy node can be eliminated from the shape functions. The spatial derivative of the shape function in the initial state for each subelement $^{Sube}\mathbf{N}'^{ini}$ ($= d^{Sube}\mathbf{N}^{ini}/d\mathbf{x}$) is calculated in the same fashion as the standard T4 element. Since each subelement is considered as a standard T4 element, $^{Sube}\mathbf{N}'^{ini}$ is constant in each subelement, Now that $^{Sube}\mathbf{N}'^{ini}$ is fundamentally derived from the standard T4 element, and thus locking and pressure checkerboarding occur when we use $^{Sube}\mathbf{N}'^{ini}$ directly for strain evaluation. Therefore, SelectiveCS-FEM-T10 smoothes $^{Sube}\mathbf{N}'^{ini}$ s inside the element before strain evaluation.

Smoothed derivatives of shape functions

As SelectiveCS-FEM-T10 adopts SRI, we use two kinds of smoothed spatial derivatives of shape functions for deviatoric and hydrostatic stress components. The derivative of the shape function for the deviatoric stress component is defined on each subelement (12 in total), whereas that for the hydrostatic stress component is defined only on the element. As a result, the spatial order of the hydrostatic stress is reduced in comparison to the deviatoric stress. The following shows the derivation of the derivatives.

For the deviatoric stress, we perform a cycle of smoothing. First, in the same fashion as ES-FEM-T4, the smoothed derivative on each edge ($^{Edge}\widetilde{\mathbf{N}}'^{ini}$) is derived from $^{Sube}\mathbf{N}'^{ini}$ s as

$$^{Edge}\widetilde{\mathbf{N}}'_{P,j}{}^{ini} = \frac{1}{^{Edge}V_h^{ini}} \sum_{k \in ^{Edge}\mathbb{K}_h} ^{Sube}\mathbf{N}'_{P,j}{}^{ini} ^{Sube}V_k^{ini} / 6, \quad (2)$$

where $\mathbf{N}'_{P,j}$ is the derivative of the shape function on Node P in the j th direction ($= \partial N_P / \partial x_j$), $^{Edge}\mathbb{K}_h$ denotes the set of subelements adjacent to Edge h , $^{Sube}V_k^{ini}$ is the initial volume of Subelement k , and $^{Edge}V_h^{ini}$ is the initial corresponding volume of Edge h ($= \sum_{k \in ^{Edge}\mathbb{K}_h} ^{Sube}V_k^{ini} / 6$). Next, using the obtained $^{Edge}\widetilde{\mathbf{N}}'^{ini}$ s, the smoothed derivative in each subelement ($^{Sube}\widetilde{\mathbf{N}}'^{ini}$) is derived as

$$^{Sube}\widetilde{\mathbf{N}}'_{P,j}{}^{ini} = \sum_{h \in ^{Sube}\mathbb{H}_k} ^{Edge}\widetilde{\mathbf{N}}'_{P,j}{}^{ini} / 6, \quad (3)$$

where $^{Sube}\mathbb{H}_k$ denotes the set of edges adjacent to Subelement k .

For the hydrostatic stress, we perform a smoothing over all subelements. The smoothed derivative on the element ($^{Elem}\widetilde{\mathbf{N}}'^{ini}$) is derived from $^{Sube}\mathbf{N}'^{ini}$ s as the weighted average of all the subelements:

$$^{Elem}\widetilde{\mathbf{N}}'_{P,j}{}^{ini} = \frac{1}{^{Elem}V^{ini}} \sum_{k=0}^{11} ^{Sube}\mathbf{N}'_{P,j}{}^{ini} ^{Sube}V_k^{ini}, \quad (4)$$

where $^{Elem}V^{ini}$ is the total volume of the element ($= \sum_{k=0}^{11} ^{Sube}V_k^{ini}$).

Calculation of nodal internal force

Using the two kinds of smoothed derivatives of shape functions, contributions to the nodal internal force is calculated by dividing it into two parts.

For the contribution of deviatoric stress, the deformation gradient of each subelement in the trial state (${}^{\text{Sube}}\mathbf{F}^+$) is calculated as

$${}^{\text{Sube}}F_{ij}^+ = {}^{\text{Sube}}\tilde{N}_{P,j}^{\text{ini}} x_{P,i}^+, \quad (5)$$

where \square^+ denotes a trial state and $x_{P,i}$ is the j th coordinate of Node P . Putting ${}^{\text{Sube}}\mathbf{F}^+$ (and its history) into the material constitutive equation, the Cauchy stress of each subelement in the trial state (${}^{\text{Sube}}\mathbf{T}^+$) is obtained. The deviatoric component of ${}^{\text{Sube}}\mathbf{T}^+$ is then given by

$${}^{\text{Sube}}T_{ij}^{(\text{dev})+} = {}^{\text{Sube}}T_{ij}^+ - \delta_{ij} \text{trace}({}^{\text{Sube}}\mathbf{T}^+)/3, \quad (6)$$

where δ represents the Kronecker's delta. The contribution of the deviatoric stress to the nodal internal force $\{{}^{\text{Sube}}f^{\text{int}(\text{dev})+}\}$ is calculated with the following equation.

$${}^{\text{Sube}}f_{P:p}^{\text{int}(\text{dev})+} = {}^{\text{Sube}}\tilde{N}_{P,j}^{\text{ini}} {}^{\text{Sube}}F_{jl}^{+-1} {}^{\text{Sube}}T_{lp}^{(\text{dev})+} {}^{\text{Sube}}V^+, \quad (7)$$

where $f_{P:p}^{\text{int}}$ represents the internal force of Node P in the p th direction.

For the contribution of hydrostatic stress, the deformation gradient of the element in the trial state (${}^{\text{Elem}}\mathbf{F}^+$) is calculated as

$${}^{\text{Elem}}F_{ij}^+ = {}^{\text{Elem}}\tilde{N}_{P,j}^{\text{ini}} x_{P,i}^+, \quad (8)$$

Putting ${}^{\text{Elem}}\mathbf{F}^+$ (and its history) into the material constitutive equation, the Cauchy stress of the element in the trial state (${}^{\text{Elem}}\mathbf{T}^+$) is obtained. The hydrostatic component of ${}^{\text{Elem}}\mathbf{T}^+$ is then given by

$${}^{\text{Elem}}T_{ij}^{(\text{hyd})+} = \delta_{ij} \text{trace}({}^{\text{Elem}}\mathbf{T}^+)/3. \quad (9)$$

The contribution of the hydrostatic stress to $\{{}^{\text{Sube}}f^{\text{int}(\text{dev})+}\}$ is calculated by the following equation.

$${}^{\text{Elem}}f_{P:p}^{\text{int}(\text{hyd})+} = {}^{\text{Elem}}\tilde{N}_{P,j}^{\text{ini}} {}^{\text{Elem}}F_{jl}^{+-1} {}^{\text{Elem}}T_{lp}^{(\text{hyd})+} {}^{\text{Elem}}V^+. \quad (10)$$

Finally, the total contribution of the element to the nodal internal force is calculated as the sum of Eq. (7) and (10):

$${}^{\text{Elem}}f_{P:p}^{\text{int}+} = \sum_{k=0}^{11} \left({}^{\text{Sube}}f_{P:p}^{\text{int}(\text{dev})+} \right) + {}^{\text{Elem}}f_{P:p}^{\text{int}(\text{hyd})+}. \quad (11)$$

Calculation of tangent stiffness matrix

The tangent stiffness is obtained by calculating $\partial\{f^{\text{int}}\}/\partial\{x\}$ according to the definition. Details are omitted due to the limitation of space.

Characteristics in formulation

The characteristics seen in the formulation of SelectiveCS-FEM-T10 are summarized as follows.

- It is a pure displacement-based finite element method.
 \Rightarrow It is applicable to dynamic explicit analysis unlike hybrid elements.
- There is no need to smooth strains across elements.
 \Rightarrow It can be implemented into general-purpose FEM software as a user-defined T10 element.
- The shape functions are all linear (1st-order).
 \Rightarrow It has superior robustness in large deformation; meanwhile, its mesh convergence is slower than 2nd-order elements in small deformation.
- Selective reduction integration (SRI) is used.
 \Rightarrow It is difficult to deal with material constitutive models considering pressure dependence etc.

According to our numerical experiments, the proposed one cycle strain smoothing for deviatoric stress is the optimal procedure to achieve accuracy and stability. When we perform only the edge-based smoothing within the element for deviatoric stress, no smoothing is applied to the edges of the element outline and thus shear locking occurs. On the contrary, when we repeat the cycle of strain smoothing more than once, too much smoothing is conducted and thus zero-energy mode occurs.

Results

Bending of cantilever

A large deflection cantilever bending analysis of a nearly incompressible material is performed. The analysis domain is a cuboid of $10 \times 1 \times 1$ m, its left end face is perfectly constrained, and a concentrated load in the vertical downward direction is given to the tip corner point. The material is a neo-Hookean hyperelastic body with 6 GPa initial Young's modulus and 0.499 initial Poisson's ratio. An unstructured T10 mesh with 0.2 m mesh seed size is used. In addition to the analysis with SelectiveCS-FEM-T10, analyses with ABAQUS T10 elements (C3D10, C3D10M, C3D10H, C3D10MH, and C3D10HS) are also performed using the same mesh to compare their accuracy and stability.

The distribution of Mises stress and pressure when the concentrated load is 2×10^7 N is shown in Fig. 2. The deformations are almost the same in all methods, which confirms that SelectiveCS-FEM-T10 avoids volume locking. Only ABAQUS C3D10 suffers from moderate pressure checkerboarding and the other methods including SelectiveCS-FEM-T10 are free from pressure checkerboarding. The amount of deformation at the loaded node is somewhat larger in SelectiveCS-FEM-T10 in comparison with the ABAQUS T10 elements. This might be because SelectiveCS-FEM-T10 does not use 2nd-order shape functions unlike the ABAQUS T10 and behaves more softly in large deformation. Note that the superiority or inferiority of these deformation results is difficult to be determined because of stress singularity.

Barreling of cylinder

A large deformation cylinder barreling analysis of a nearly incompressible material is performed. The analysis domain is a 1/8 of a cylinder of 1 m radius and 2 m height, symmetric boundary

conditions are applied to the symmetric surfaces, and an enforced displacement to the vertical downward direction is applied to the top surface with constraints of in-plane displacements. The material is a neo-Hookean hyperelastic body with 6 GPa initial Young's modulus and 0.49 initial Poisson's ratio. An unstructured T10 mesh with 0.05 m mesh seed size is used. As in the previous example, analyses with SelectiveCS-FEM-T10 and five ABAQUS C3D10 elements are performed using the same mesh.

The distributions of Mises stress and pressure at 0.24 m enforced displacement (24% compression) are shown in Fig. 3. Although the stress distributions near the rim of the top surface are

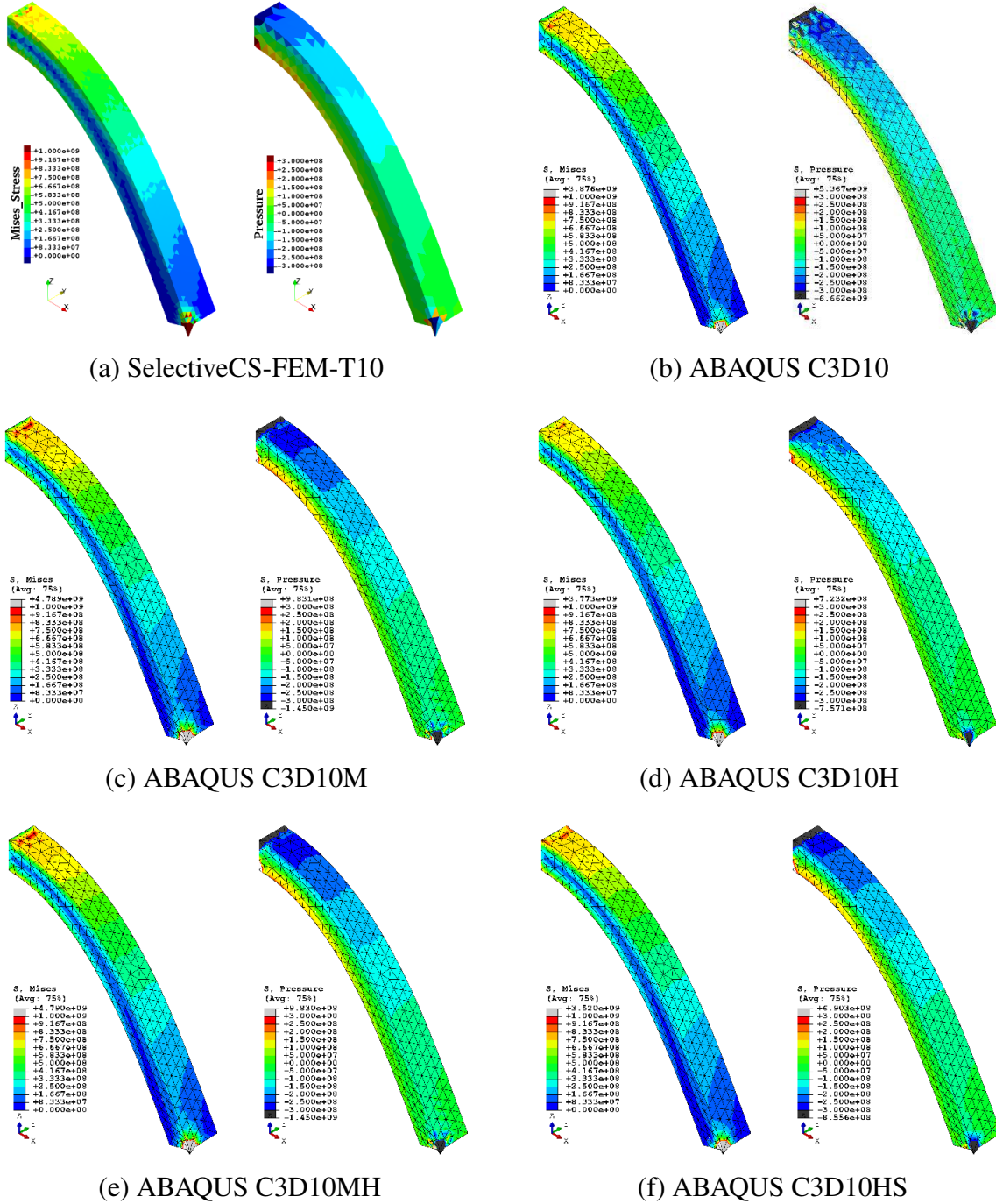


Figure 2: Comparison of Mises stress (left) and pressure (right) distributions in the cantilever bending analysis.

somewhat different each other due to stress singularity, their deformations and stress distributions have no much difference among all the methods. Looking at the deformation on the rim part carefully, we can see that the element edges are largely bent periodically at the intermediate nodes in the ABAQUS T10 elements. It is well known that the accuracy and stability of T10 elements drop greatly when the position of the intermediate node deviates largely from the midpoint of the corner nodes. In fact, all of the ABAQUS T10 elements get converge failure around 25% compression. On the other hand, SelectiveCS-FEM-T10 does not show such a bending at the intermediate nodes owing to the piecewise linear shape functions.

Fig. 4 shows the distribution of the nodal reaction forces on the top surface at the same time. Typical nodal reaction force oscillations are seen in the results of the non-modified ABAQUS elements (C3D10, C3D10H and C3D10HS), whereas SelectiveCS-FEM-T10 and the modified

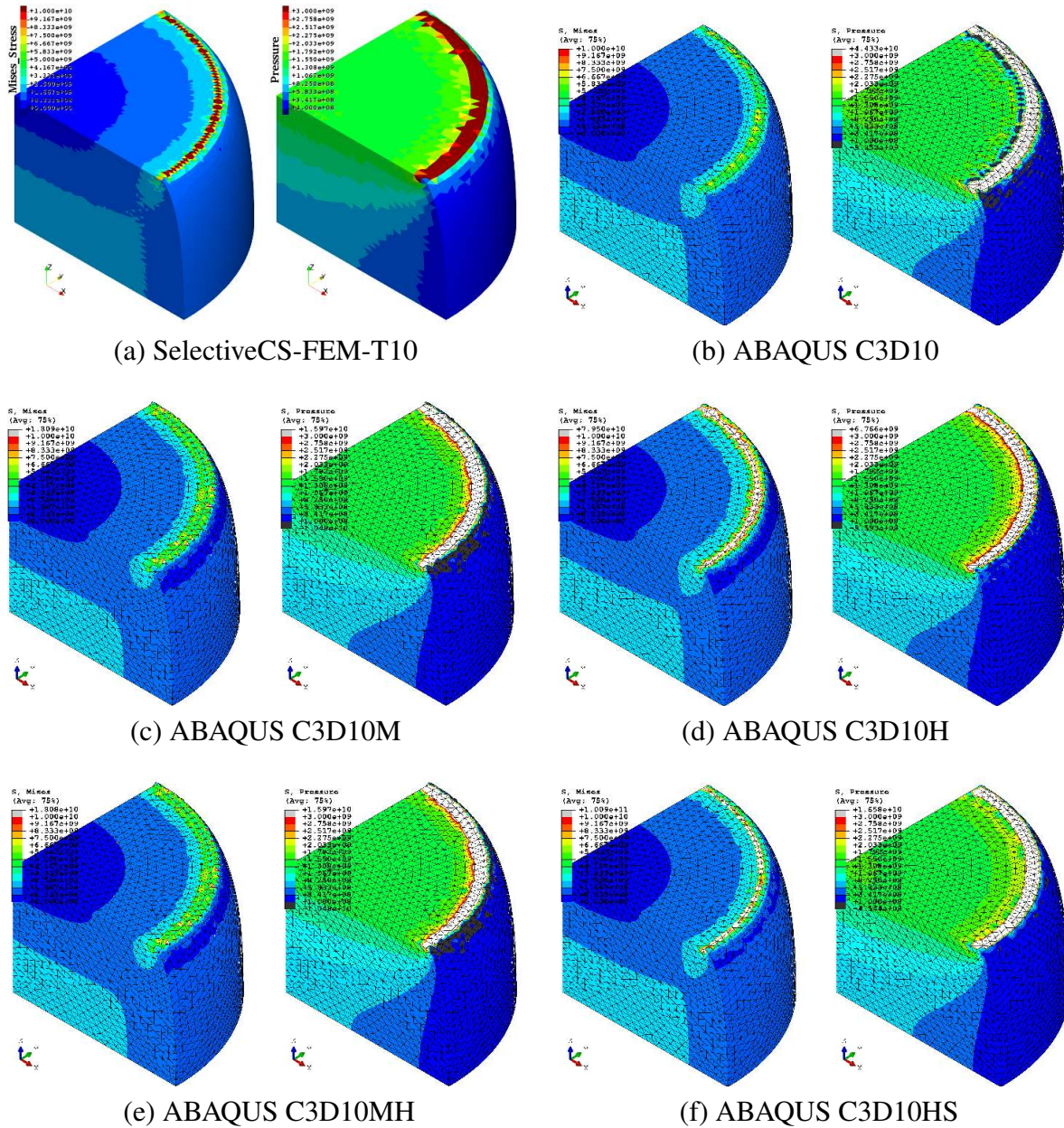


Figure 3: Comparison of Mises stress (left) and pressure (right) distributions in the cylinder barreling analysis at 24% compression states.

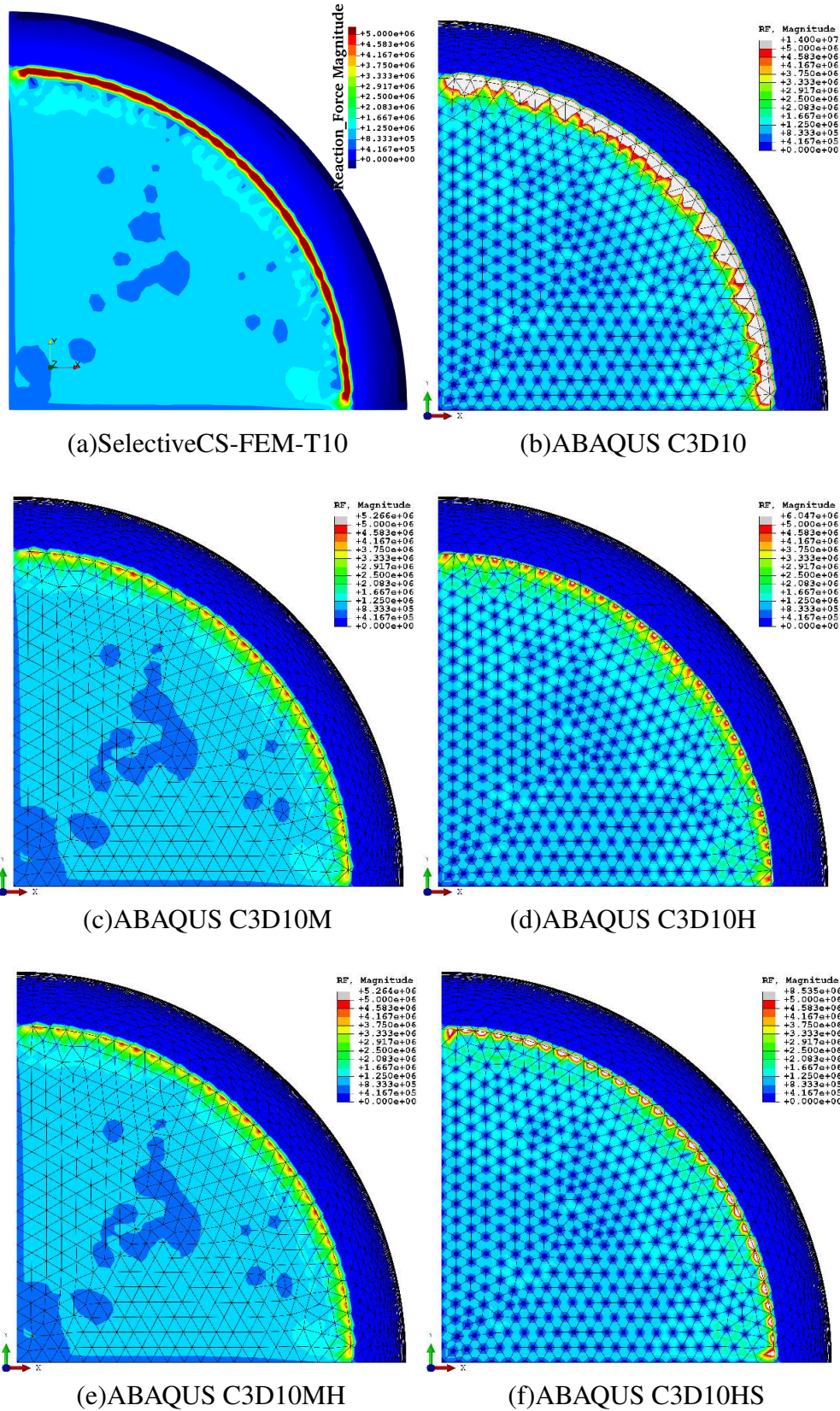


Figure 4: Comparison of nodal reaction force distributions of the upper face in the cylinder barreling analysis at 24% compression states. The proposed method (a) and the modified T10 elements of ABAQUS ((c) and (e)) represent valid distributions. In contrast, the non-modified T10 elements of ABAQUS ((b), (d) and (f)) represent oscillatory invalid distributions.

ABAQUS elements (C3D10M and C3D10MH) show valid distributions. Nodal reaction force oscillation is a severe issue especially when handling contacts, but it is confirmed that SelectiveCS-FEM-T10 avoids this issue.

As a demonstration, the distributions of Mises stress and pressure at 0.47 m enforced displacement (47% compression) are shown in Fig. 5. Only the result of SelectiveCS-FEM-T10 is shown in this figure because the ABAQUS T10 elements get convergence failure in earlier states, as mentioned above. SelectiveCS-FEM-T10 gets convergence failure at 48% compression in this case; however, it gives reasonable deformation and stress distribution until it reaches the convergence failure.

Conclusion

A novel smoothed finite element method (S-FEM) using 10-node tetrahedral (T10) element, SelectiveCS-FEM-T10, was proposed. By combining the selective reduced integration (SRI) and S-FEM with cyclic smoothing, the proposed method overcome various issues: shear/volumetric locking, pressure checkerboarding, and nodal reaction force oscillation. Unlike the conventional 4-node tetrahedral (T4) S-FEMs, the strain smoothings of the proposed method are only performed within each T10 element using T4 subelements. As a result, the proposed method can be implemented as a user-defined element of general-purpose FEM codes and also its computational time is almost equivalent to the conventional T10 elements. Moreover, like the conventional S-FEMs, the proposed method is applicable to the dynamic explicit analysis because it is a pure displacement-based finite element method.

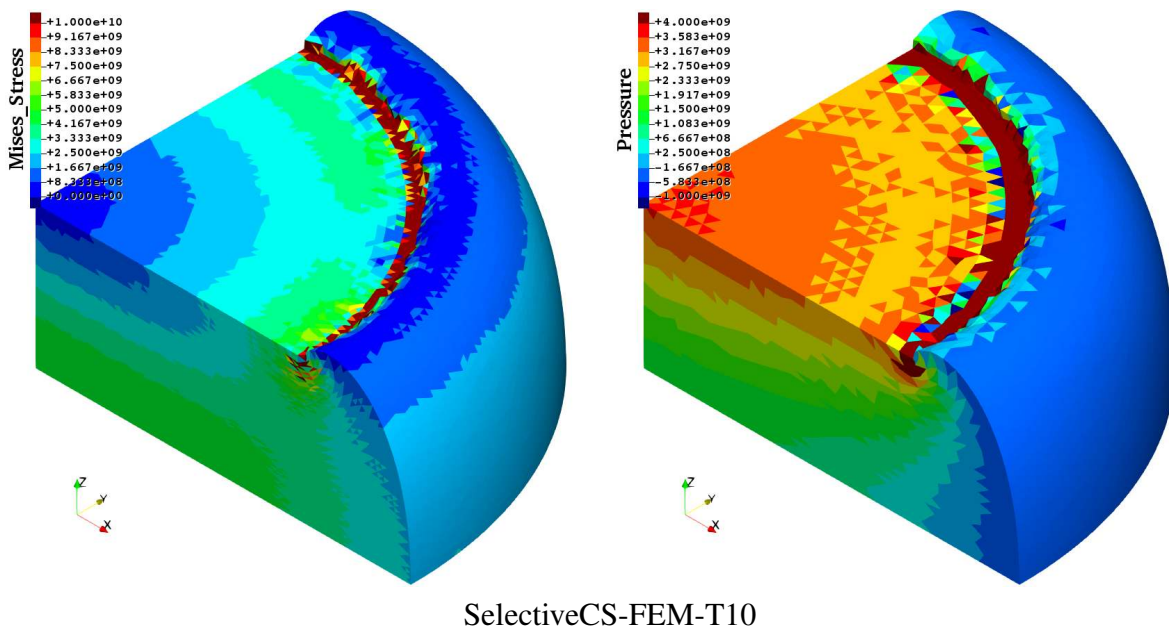


Figure 5: Mises stress (left) and pressure (right) distributions of SelectiveCS-FEM-T10 in the cylinder barreling analysis at 47% compression states. Every ABAQUS T10 elements get convergence failure around 25% compression and thus their results are not shown here.

References

- [1] D. S. S. Corp., *ABAQUS 6.13 Theory Guide*. Providence, RI, USA: Dassault Systèmes Simulia Corp., 2013.
- [2] MSC Software Corp., *MARC 2013.1 Volume A: Theory and User Information*. Santa Ana, CA, USA: MSC Software Corp., 2013.
- [3] G. R. Liu and T. Nguyen-Thoi, *Smoothed Finite Element Methods*. Boca Raton, FL, USA: CRC Press, 2010.
- [4] W. Zeng and G. R. Liu, “Smoothed finite element methods (S-FEM): An overview and recent developments,” *Archives of Computational Methods in Engineering*, pp. 1–39, 2016.
- [5] Y. Onishi and K. Amaya, “A locking-free selective smoothed finite element method using tetrahedral and triangular elements with adaptive mesh rezoning for large deformation problems,” *International Journal for Numerical Methods in Engineering*, vol. 99, no. 5, pp. 354–371, 2014.
- [6] Y. Onishi and K. Amaya, “Performance evaluation of the selective smoothed finite element methods using tetrahedral elements with deviatoric/hydrostatic split in large deformation analysis,” *Theoretical and Applied Mechanics Japan*, vol. 63, pp. 55–65, 2015.
- [7] Y. Onishi, R. Iida, and K. Amaya, “F-bar aided edge-based smoothed finite element method using tetrahedral elements for finite deformation analysis of nearly incompressible solids,” *International Journal for Numerical Methods in Engineering*, vol. 109, no. 11, pp. 1582–1606, 2017.
- [8] Y. Onishi, “F-bar aided edge-based smoothed finite element method with 4-node tetrahedral elements for static large deformation elastoplastic problems,” *International Journal of Computational Methods*, vol. 15, no. 1, 2018.
- [9] Y. Onishi, R. Iida, and K. Amaya, “Accurate viscoelastic large deformation analysis using F-bar aided edge-based smoothed finite element method for 4-node tetrahedral meshes (F-BarES-FEM-T4),” *International Journal of Computational Methods*, vol. 17, no. 7, p. 1845003, 2018.
- [10] R. Iida, Y. Onishi, and K. Amaya, “A stabilization method of F-barES-FEM-T4 for dynamic explicit analysis of nearly incompressible materials,” *International Journal of Computational Methods*, in Print.
- [11] E. de Souza Neto, D. Peric, M. Dutko, and D. Owen, “Design of simple low order finite elements for large strain analysis of nearly incompressible solids,” *International Journal of Solids and Structures*, vol. 33, no. 20-22, pp. 3277–3296, 1996.
- [12] J. T. Ostien, J. W. Foulk, A. Mota, and M. G. Veilleux, “A 10-node composite tetrahedral finite element for solid mechanics,” *International Journal for Numerical Methods in Engineering*, vol. 107, no. 13, pp. 1145–1170, 2016.
- [13] E. A. de Souza Neto, F. M. A. Pires, and D. R. J. Owen, “F-bar-based linear triangles and tetrahedra for finite strain analysis of nearly incompressible solids. Part I: formulation and benchmarking,” *International Journal for Numerical Methods in Engineering*, vol. 62, no. 3, pp. 353–383, 2005.
- [14] E. A. de Souza Neto, D. Peric, and D. R. J. Owen, *Computational Methods for Plasticity: Theory and Applications*. Hoboken, NJ, USA: Wiley, 2008.

An Evaluation of Hegemonic Epochs – A Time Series Analysis

†M. Caramihai¹, *Irina Severin²

¹Faculty of Automatic Control & Computer Sciences, ²Faculty of Engineering / Management of Technological Systems, University POLITEHNICA Bucharest, ROMANIA.

*Presenting author: m.caramihai@yahoo.com

†Corresponding author: m.caramihai@yahoo.com

Abstract

The paper aims to present the time evolution of the number of hegemonic states based on a temporal series analysis. The basic idea is that a hegemonic state is the state that culminates by a distinct, universal civilization. As a consequence, the analysis has as basic input data the number and the duration of general recognized civilization. There are identified a general evolution trend and a periodic component. Finally, a prediction is made concerning the number of hegemonic states during the next five-hundred years period.

Keywords: hegemonic state, universal civilization, time series analysis

1. Introduction

Stability of the world depends on the equilibrium established between forces driving it. Those forces are mainly issued from states which imposed their leadership either cultural, economic or military to the rest of the world – the hegemonic states.

For this reason, many researchers in history and sociology were interested in studying the characteristics of hegemonic states, the mechanisms which drive their formation, their influence zones and their evolution or involution in time.

This paper aimed to study by mathematical means the frequency of hegemonic states apparition in order to make predictions concerning the structure of the geo-political world map.

The results obtained until now focused on the number of hegemonic states existing in a given time period, but future research could offer predictions concerning their position.

The history and the sociology didn't elaborate a "definite standard" on what is and what characterize a hegemonic state. This paper is based mainly on the theory and data elaborated by a well-recognized historian – Arnold Toynbee [1], [2]. One of the reasons is that his theories have raised a resuscitated interest in the last years, being used in the writings of very contemporary politicians [3].

2. The hegemonic state and the universal civilization

Accordingly to the theories of Toynbee, a hegemonic state represents a civilization which reached an universality state.

A civilization is actually defined as a cultural entity: language, religion, life style defines a civilization. The universal civilizations of the human history were usually defined by their

religion. This doesn't mean that civilizations could be very exact delimited or that into a given civilization couldn't subsist groups of people with a religion different of the dominant one.

A civilization is the most extended cultural entity: states, ethnic groups, nationalities, religious groups – all of them have distinct cultures at different levels of heterogeneity. Though a civilization is a group of people with the same cultural concept – defined both by common objectives as language, history, religion, customs, institutions and by their own subjective auto-identification.

People have different levels of identity: as an example, a Roman resident could define itself by various degrees of identity as Roman, Italian, Catholic, Christian, Occidental, European.

This point of view became less obvious in our world, when levels of identity are less discriminatory, but were extremely powerful in former historical phases. Anyway, Toynbee used the identity universal civilization – hegemonic state and defined consequently the following hegemonic states with their respective starting/ ending periods (Table 1).

Table 1: Universal civilizations

N°	Civilization	Starting year	Ending year
1	Summerian	-2298	-1905
2	Babylonian	-650	-529
3	Indic	-322	-185
4	Synic	-221	172
5	Helenic	-31	378
6	Egyptian (1)	-2070	-1660
7	Egyptian (2)	-1580	-1175
8	Orthodox (Russia)	1478	1881
9	Extreme oriental	1597	1868
10	Occident (1)	1797	1814
11	Occident (2)	1526	1918
12	Andine	1430	1533
13	Syriac (1)	-525	-332
14	Syriac (2)	640	969
15	Extreme Oriental (1)	1280	1351
16	Extreme Oriental (2)	1644	1853
17	Central American	1521	1821
18	Orthodox	1372	1768
19	Hindus (1)	1572	1707
20	Hindus (2)	1818	1947
21	Mynoic	-1750	-1400
22	Maya	300	690

3. A mathematical approach

3.1. Conceptual aspects

A temporal series (dynamic series) consists of a lot of observations resulting from measurements made in successive periods of time [4], [5]. Commonly, it is noted $\{x_t, t \in T\}$,

where x_t data represents consecutive measurements, taken at quasi-equal intervals (hours, days, weeks, etc.). A temporal series illustrates the dynamics of a particular process over time, x_t representing the measured value at time $t \in T$.

Since time is continuous, in time series analysis it is fragmented in equidistant periods: hours, days, months, years, etc., hence it has a discrete character.

The evolutionary time series include three basic components:

- i. Trend or tendency (T);
- ii. Periodic component or cycle (P);
- iii. The random or stochastic component (A).

Under these conditions, it can be considered that a temporal series $Y(t)$ consists of the sum of the three basic components:

$$Y(t) = T + P + A \text{ (summative model)}$$

or from their product:

$$Y(t) = T \times P \times A. \text{ (multiplicative model)}$$

where: the tendency indicates the ascending or descending change of the evolution of the series, the periodic component includes oscillations (cycles) that are repeated at regular time intervals, with regular or irregular amplitudes and the random component expresses the residue or deviations of the series values from the theoretical values corresponding to the trend.

Time series analysis involves separating the three components and interpreting them.

3.2. Data analysis

The first approach concerned the time distribution of the universal civilization (Figure 1). Numbers in the figure identify the civilizations presented in Table 1. As the figure shows the time density of hegemonic states has increased in the last 2000 years.

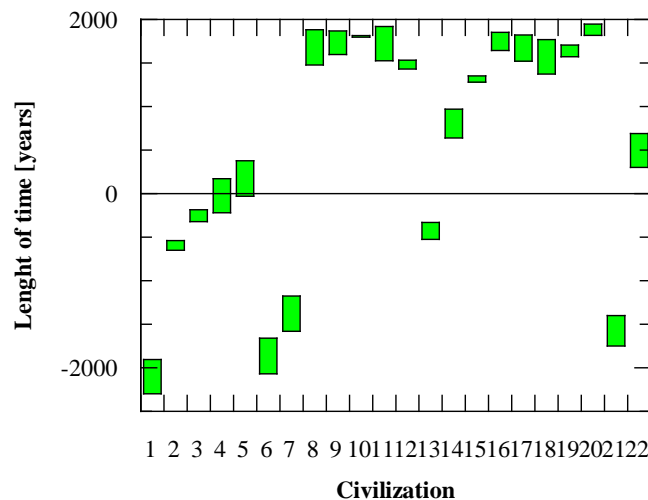


Figure 1: Time density of universal civilization

The representation as time series of the data included in Table 1 resulted in Table 2.

Table 2: Number of hegemonic states/ time	
Time interval	Number of hegemonic states (NHS)
-3000 -2500	0
-2500-2000	2
-2000-1500	3
-1500-1000	2
-1000-500	2
-500-0	4
0-500	3
500-1000	2
1000-1500	4
1500-2000	9
2000-2500	3 (estimated value)

The sample period used for the study is one of 500 years – the main reason being the use of the same period in the Toynbee approach.

Figure 2 shows from another point of view the evolution of the density in time of the number of hegemonic states. A “peak” of this evolution appears in the last five hundred years – a historical period in which many states identified themselves as nations and implicitly as civilizations.

Moreover, in Figure 2 could be observed a slight increasing trend of the number of universal civilization and the existence of a certain periodicity of the evolution.

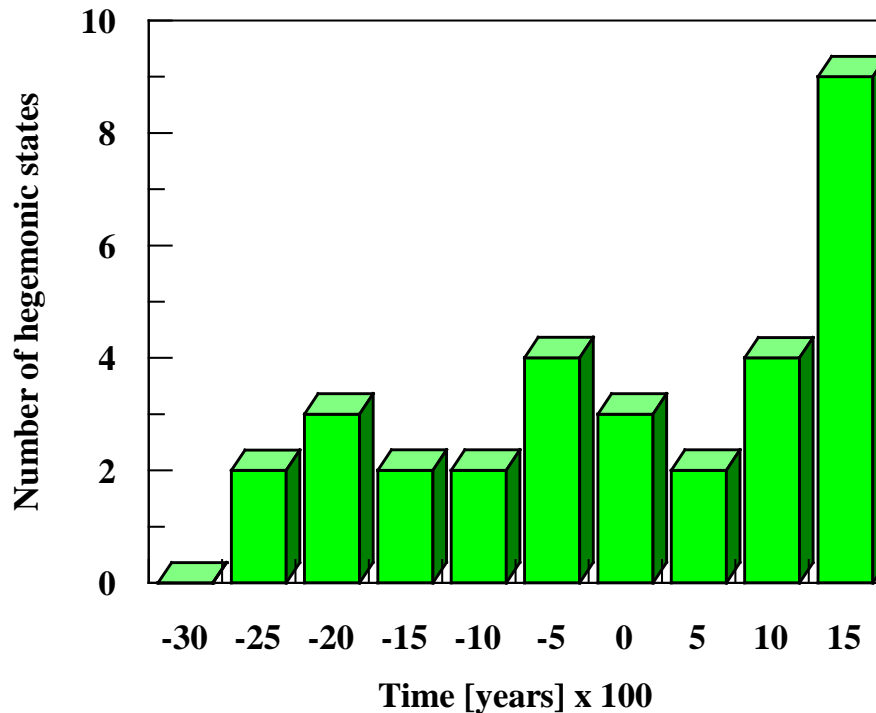


Figure 2: Time - density of hegemonic states

Another representation of the former data (i.e. as temporal series) is shown in Figure 3, in which the evolution periodicity (Figure 5) of the number of hegemonic states (NHS) and the increasing trend (Figure 4) are more evident.

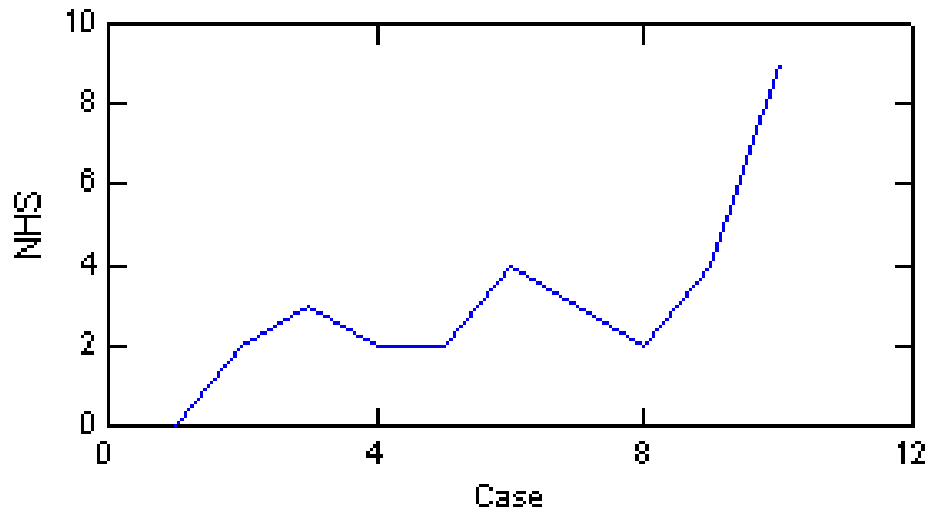


Figure 3: A time series of hegemonic states evolution

The correlation coefficient of the time-series model is 0.733, which is acceptable, considering that, as mentioned above, it is difficult to clearly define temporal “boundaries” of a civilization.

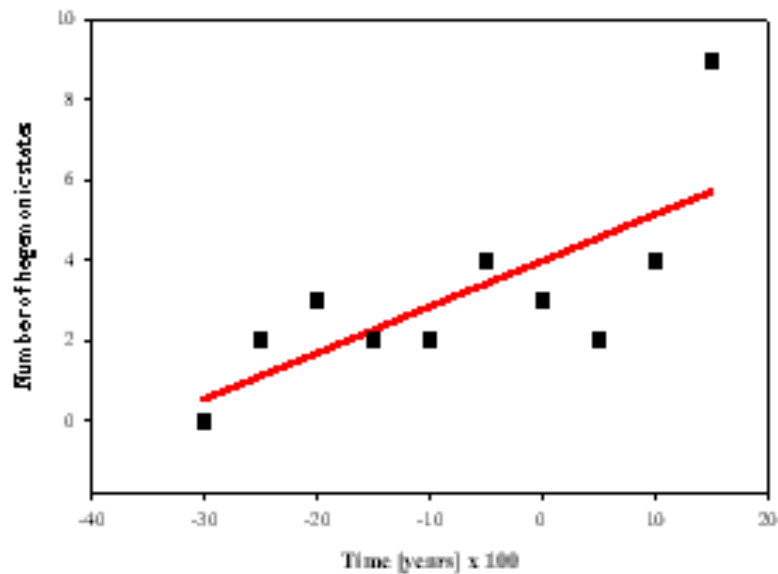


Figure 4: Increasing trend of NHS

The slope of the linear trend is 0.11, but it is corrected by evolution periodicity which imposes for the immediate future a decreasing of the number of universal civilizations.

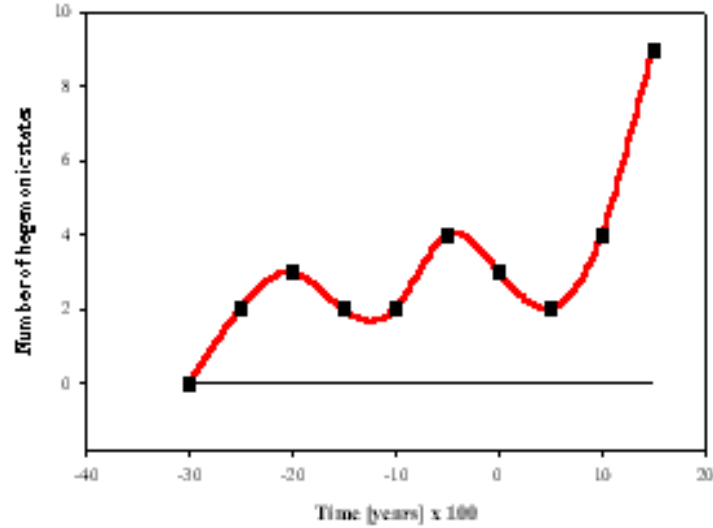


Figure 5: Evolution periodicity of NHS

If the trend is eliminated, the evolution of NHS results as is shown in Figure 6.

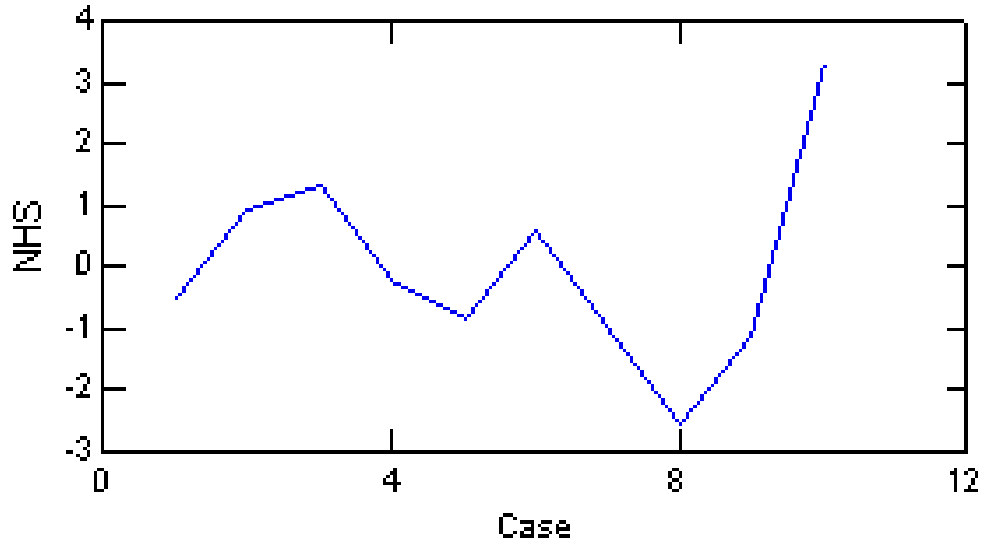


Figure 6.

The time series modeling the evolution of NHS without the increasing trend was identified as being:

$$f = \sin\left(\frac{2\pi x}{T} + k\right)$$

where x is the time-variable and T and k are constants with following values: T=29.17 and k = -5.68

Based on this time series, the prediction for the next 500 years-period is the existence of 3 hegemonic states.

Conclusion

The present paper aimed to model by time series means the evolution of the number of hegemonic states in order to predict a future evolution of the geo-political structure.

The assumptions on which the study is based are those made of a historian and sustained by many politicians, so as their theory is modeled by statistical means. Interpretation of the results could be made only following the starting theory.

References

- [1] Toynbee A.J.(1934) A Study of History. Oxford University Press
- [2] Somervell D.C.(1946) A study of History-Abridgment. Oxford University Press
- [3] Huntington S.P. (1997). The Clash of Civilizations and the Remaking of World Order. Simon & Schuster
- [4] Shumway, Robert H. (2000) Time series analysis and its applications, Springer
- [5] Pal, Avishek (2017) Practical Time Series Analysis, Packt Publishing Ltd,

An optimal iteration function for multiple zeros with eighth-order convergence

†Ramandeep Behl¹ and Ali Saleh Alshomrani²

^{1,2}Department of Mathematics, King Abdulaziz University, Jeddah 21577, Saudi Arabia.

†Corresponding author: ramanbehl87@yahoo.in

Abstract

Over the past several years, many scholars have attempted to construct higher-order schemes for locating multiple solutions of a univariate function having known multiplicity $m \geq 1$. But till date, we have a very limited literature (only four research articles) of eighth-order convergence iteration functions for multiple zeros. The primary contribution of this study is to propose an optimal eighth-order scheme for multiple zeros having simple and compact body structure with faster convergence. An extensive convergence analysis is also present with the main theorem which clearly show the eighth-order convergence of propose iteration scheme. Finally, numerical tests on some real-life problems, such as a Van der Waals equation of state and the conversion problem from the chemical engineering, among others are presented, which confirm the theoretical results to great extent of this study.

Keywords: Nonlinear equations, King-Traub conjecture, multiple roots, optimal iterative methods, efficiency indexce.

Introduction

Finding the multiple zeros of the involved function $f(x) = 0$ (where $f : \mathbb{D} \subset \mathbb{C} \rightarrow \mathbb{C}$ is a holomorphic function in the enclosed region \mathbb{D} containing the required zero) is one of the most challenging, of great significance and difficult tasks in the area of computational mathematics. It is quite tough to obtain exact solution in analytic way of such problems or we can say that it is almost fictitious. So, we have to satisfy ourselves by obtaining approximated and efficient solution up to any specific degree of accuracy by the means of iterative procedure.

This is one of the main reason that researchers are putting their great efforts to resort an iteration function since the past few decades. Additionally, this accuracy is also depend on some other facts like: the considered iterative function, structure of the considered problem, initial guess and programming software namely, Maple, MATLAB, Fortran, Mathematica, etc. Further, the people or researchers using these iterative methods have to struggle with many problems, some of them are like: choice of initial guess/approximation, slower convergence, non-convergence, divergence, oscillation problem close to the initial guess, failure etc. (for the details please see Ostrowski 1960 [17], Traub 1964 [25], Ortega and Rheinboldt 1970 [18], Burden and Faires 2001 [?], Petkovic et al. 2012 [19]).

In addition, we dont have a single iteration function which is applicable to every problem until now. This is the main reason that we have an excessive amount of literature on the iteration functions for scalar equations. Here, we concern about the multiple zeros of the involved univariate function in this study. Unfortunately, we have a small amount of literature belongs to higher-order iteration function in the case of scalar equations that can handle multiple roots. The tough calculation work and more time consumption are the main reason behind of this. Moreover, it is more challenging task to construct iterative procedure for multiple zeros as compared to simple.

Eighth-order multi-point methods have faster convergence and better efficiency index as compare to fourth-order [6, 16, 15, 27, 21, 22, 23, 28, 14, 4, 5, 24] and sixth-order [12, 13] iteration functions. Our mean to say that we can save computational time and cost by using them and obtain the approximate solution in a small number of iterations as compared to them. However, we have only four research articles [3, 26, 2, 8] till date that talk about the eighth-order convergence for multiple zeros with known multiplicity $m \geq 1$, according to our best knowledge. But, we know that there is always a scope in the research to obtain better approximation techniques with simple and compact body structure.

While keep all these things in our mind, we not only present an eighth-order iteration scheme having optimal convergence for obtaining the multiple solutions of scalar equation which is better than the existing ones. But, lower residual errors, lower error among two consecutive iterations and more stable computational order of convergence belong to our methods when we compared them to the existing ones of identical order of convergence. Moreover, we present a main theorem which demonstrate the eighth-order convergence when multiplicity of zeros is known in advance. Finally, we give a practical exhibition of our newly propose methods to the real life problems.

Construction of higher-order scheme

In this section, we present the main contribution of this study. Our mean to say that we present an eighth-order scheme for multiple zeros having simple and compact body structure. Therefore, we consider the new scheme in the following way:

$$\begin{aligned} y_n &= x_n - m \frac{f(x_n)}{f'(x_n)}, \\ z_n &= y_n - \mu H(\nu) \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= z_n - \kappa \mu \left(G(\mu) + \frac{m\kappa}{1-4\mu} \right) \frac{f(x_n)}{f'(x_n)}, \end{aligned} \quad (1)$$

where $\alpha, \beta \in \mathbb{R}$ are two free disposable parameters and two weight functions $H : \mathbb{C} \rightarrow \mathbb{C}$ and $G : \mathbb{C} \rightarrow \mathbb{C}$ are analytic functions in the neighborhoods of (1) and (0) with $\nu = \frac{1+\alpha\mu}{1+\beta\mu}$, $\mu = \left(\frac{f(y_n)}{f(x_n)} \right)^{\frac{1}{m}}$, $\kappa = \left(\frac{f(z_n)}{f(y_n)} \right)^{\frac{1}{m}}$.

In Theorem 1, we illustrate that the constructed scheme (1) attain maximum eighth-order of convergence for all $\alpha, \beta \in \mathbb{R}$ ($\alpha \neq \beta$), without using any extra functional evaluation. It is interesting to observe that the weight functions H and G play significant role in the construction of scheme (for details please see Theorem 1).

Theorem 1 *Let us consider $x = \xi$ (say) be a multiple zero with multiplicity $m \geq 1$ of the involved function f . In addition, we assume that $f : \mathbb{D} \subset \mathbb{C} \rightarrow \mathbb{C}$ be an analytic function in the region \mathbb{D} enclosing a multiple zero ξ . Then, the scheme defined by (1) has an eighth-order convergence, when it satisfies the following values*

$$\begin{aligned} H(1) &= m, \quad H'(1) = \frac{2m}{\alpha - \beta} (\alpha \neq \beta), \quad G(0) = m, \quad G'(0) = 2m, \quad G''(0) = H''(1)(\alpha - \beta)^2 + (2 - 4\beta)m, \\ G'''(0) &= (\alpha - \beta)^2 \left(H'''(1)(\alpha - \beta) - 6(\beta - 1)H''(1) \right) + 12m(\beta^2 - 2\beta - 2). \end{aligned} \quad (2)$$

Proof: Let us consider that $e_n = x_n - \xi$ and $c_k = \frac{m!}{(m-1+k)!} \frac{f^{m-1+k}(\xi)}{f^m(\xi)}$, $k = 2, 3, 4, \dots, 8$ are the

error in n th iteration and asymptotic error constant numbers, respectively. Now, we expand the Taylor's series expansions of the functions $f(x_n)$ and $f'(x_n)$ about $x = \xi$, which are given by

$$f(x_n) = \frac{f^{(m)}(\xi)}{m!} e_n^m \left(1 + c_1 e_n + c_2 e_n^2 + c_3 e_n^3 + c_4 e_n^4 + c_5 e_n^5 + c_6 e_n^6 + c_7 e_n^7 + c_8 e_n^8 + O(e_n^9) \right) \quad (3)$$

and

$$f'(x_n) = \frac{f^m(\xi)}{m!} e_n^{m-1} \left(m + (m+1)c_1 e_n + (m+2)c_2 e_n^2 + (m+3)c_3 e_n^3 + (m+4)c_4 e_n^4 + (m+5)c_5 e_n^5 + (m+6)c_6 e_n^6 + (m+7)c_7 e_n^7 + (m+8)c_8 e_n^8 + O(e_n^9) \right), \quad (4)$$

respectively.

By using the expressions (3) and (4) in the first substep of scheme (1), we have

$$y_n - \xi = \frac{c_1}{m} e_n^2 + \frac{1}{m^2} (2mc_2 - (m+1)c_1^2) e_n^3 + \sum_{i=0}^4 \theta_i e_n^{i+4} + O(e_n^9), \quad (5)$$

where $\theta_i = \theta_i(m, c_1, c_2, \dots, c_8)$ are given in terms of m, c_2, c_3, \dots, c_8 for example $\theta_0 = \frac{1}{m^3} [3m^2 c_3 + (m+1)^2 c_1^3 - m(3m+4)c_1 c_2]$ and $\theta_1 = \frac{1}{m^4} [2c_2 c_1^2 m(2m^2 + 5m + 3) - 2c_3 c_1 m^2(2m+3) - 2m^2(c_2^2(m+2) - 2c_4 m) - c_1^4(m+1)^3]$, etc.

With the help of expression (5) and Taylor Series expansion, we further obtain

$$f(y_n) = f^{(m)}(\xi) e_n^{2m} \left[\frac{\left(\frac{c_1}{m}\right)^m}{m!} + \frac{(2mc_2 - (m+1)c_1^2) \left(\frac{c_1}{m}\right)^m e_n}{m! c_1} + \left(\frac{c_1}{m}\right)^{1+m} \frac{1}{2m! c_1^3} \{(3 + 3m + 3m^2 + m^3) c_1^4 - 2m(2 + 3m + 2m^2) c_1^2 c_2 + 4(m-1)m^2 c_2^2 + 6m^2 c_1 c_3\} e_n^2 + \sum_{i=0}^4 \bar{\theta}_i e_n^{i+3} + O(e_n^8) \right]. \quad (6)$$

From the expressions (3) and (6), we have

$$\mu = \frac{c_1 e_n}{m} + \frac{2mc_2 - (m+2)c_1^2}{m^2} e_n^2 + \sum_{i=0}^4 \bar{\theta}_i e_n^{i+3} + O(e_n^8), \quad (7)$$

which further leads us

$$\nu = \frac{\alpha\mu + 1}{\beta\mu + 1} = 1 + (\alpha - \beta) \sum_{k=1}^8 \gamma_k e_n^k + O(e_n^9), \quad (8)$$

where $\gamma_k = \gamma_k(m, \alpha, \beta, c_1, c_2, \dots, c_8)$ are given in terms of $m, \alpha, \beta, c_2, c_3, \dots, c_8$ for example $\gamma_1 = \frac{c_1}{m}$, $\gamma_2 = \frac{1}{m^2} [2c_2 m - c_1^2(\beta + m + 2)]$, $\gamma_3 = \frac{1}{2m^3} [(2\beta^2 + 8\beta + 2m^2 + (4\beta + 7)m + 7) c_1^3 + 6c_3 m^2 - 2c_2 c_1 m(4\beta + 3m + 7)]$, etc.

Now, let us consider $\nu = 1 + \Omega$. Then, from the expression (8) that the remainder $\Omega = \nu - 1$ is infinitesimal with the order e_n . Therefore, we can expand weight function $H(\nu)$ in the neighborhood of (1) by Taylor's series expansion up to third-order terms in the following way:

$$H(\nu) = H(1) + H'(1)\Omega + \frac{1}{2!} H''(1)\Omega^2 + \frac{1}{3!} H'''(1)\Omega^3. \quad (9)$$

By using expressions (3)–(9) in the second substep of scheme (1), we obtain

$$z_n - \xi = -\frac{c_1(H(1) - m)}{m^2}e_n^2 + \sum_{i=0}^5 A_i e_n^{i+3} + O(e_n^9), \quad (10)$$

where $A_i = A_i(m, c_1, c_2, \dots, c_8, \alpha, \beta, H(1), H'(1), H''(1), H'''(1))$ are given in terms of $m, c_1, c_2, c_3, \dots, c_8, \alpha, \beta, H(1), H''(1), H'''(1)$. For example, first coefficient explicitly written as $A_0 = \frac{1}{m^3} \left[2c_2m(m - H(1)) - c_1^2(m^2 + m - H(1)(m + 3) + (\alpha - \beta)H'(1)) \right]$ and we can also write other ones in the similar way.

It is straightforward to say from the expression (10) that we can easily obtain at least third-order convergence, when we consider

$$H(1) = m. \quad (11)$$

With the help of expression (11) and $A_0 = 0$, we obtain

$$\frac{c_1^2(H'(1)(\beta - \alpha) + 2m)}{m^3} = 0, \quad (12)$$

which further yield

$$H'(1) = \frac{2m}{\alpha - \beta}, \quad \alpha \neq \beta. \quad (13)$$

In this way, we reach optimal fourth-order convergence. Now, by inserting the expressions (11) and (13) in (10), we have

$$z_n - \xi = \left[\frac{(m^2 - H''(1)(\alpha - \beta)^2 + (4\beta + 9)m)c_1^3 - 2c_1c_2m^2}{2m^4} \right] e_n^4 + \sum_{i=2}^5 A_i e_n^{i+3} + O(e_n^9). \quad (14)$$

Again, with the help of Taylor series expansion and expression (14), we obtain

$$f(z_n) = f^{(m)}(\xi)e_n^{4m} \left[\frac{2^{-m} \left(\frac{c_1^3(-H''(1)(\alpha - \beta)^2 + m^2 + (4\beta + 9)m) - 2c_1c_2m^2}{m^4} \right)^m}{m!} + \sum_{i=1}^5 \bar{A}_i e_n^i + O(e_n^6) \right]. \quad (15)$$

From the expressions (6) and (15), we further have

$$\kappa = \frac{c_1^2(m^2 - H''(1)(\alpha - \beta)^2 + (4\beta + 9)m) - 2c_2m^2}{2m^3} e_n^2 + \sum_{i=1}^5 \bar{A}_i e_n^{i+2} + O(e_n^8). \quad (16)$$

It is clear from the expression (16) that the κ is of order e_n^2 . Therefore, we can expand weight function $G(\mu)$ in the neighborhood of origin (0) by Taylor's series expansion up to third-order terms in the following way:

$$G(\mu) = G(0) + G'(0)\mu + \frac{1}{2!}G''(0)\mu^2 + \frac{1}{3!}G'''(0)\mu^3. \quad (17)$$

Insert the expressions (3) – (17) in the last substep of scheme (1), we obtain

$$e_{n+1} = \frac{c_1(G(0) - m) \left[c_1^2 \left(m^2 - H''(1)(\alpha - \beta)^2 + (4\beta + 9)m \right) - 2c_2m^2 \right]}{2m^5} e_n^4 + \sum_{i=1}^4 L_i e_n^{i+4} + O(e_n^9), \quad (18)$$

where $L_i = L_i(\alpha, \beta, m, c_1, c_2, \dots, c_8, H''(1), H'''(1), G'(0), G''(0), G'''(0))$.

It is noteworthy that we can obtain at least fifth-order convergence if we choose

$$G(0) = m. \quad (19)$$

By using the value of $G(0) = m$ and $L_1 = 0$, we have

$$-\frac{c_1^2(G'(0) - 2m) \left[c_1^2 \left(m^2 - H''(1)(\alpha - \beta)^2 + (4\beta + 9)m \right) - 2c_2m^2 \right]}{2m^6} = 0, \quad (20)$$

which further yield

$$G'(0) = 2m. \quad (21)$$

Again, by inserting the value of $G(0)$ and $G'(0)$ in $L_2 = 0$, we yield

$$-\frac{c_1^3 \left[c_1^2 \left(m^2 - H''(1)(\alpha - \beta)^2 + (4\beta + 9)m \right) - 2c_2m^2 \right] \left(G''(0) - H''(1)(\alpha - \beta)^2 + (4\beta - 2)m \right)}{4m^7} = 0, \quad (22)$$

which further have

$$G''(0) = H''(1)(\alpha - \beta)^2 + (2 - 4\beta)m. \quad (23)$$

By using the expressions (19), (21) and (23) in $L_3 = 0$, leads us

$$-\frac{c_1^4 \left(c_1^2 \left(-H''(1)(\alpha - \beta)^2 + m^2 + (4\beta + 9)m \right) - 2c_2m^2 \right)}{12m^8} \times \left(G'''(0) + (\alpha - \beta)^2(6(\beta - 1)H''(1) + H'''(1)(\beta - \alpha)) - 12m(\beta^2 - 2\beta - 2) \right) = 0, \quad (24)$$

which further provide

$$G'''(0) = (\alpha - \beta)^2 \left(H'''(1)(\alpha - \beta) - 6(\beta - 1)H''(1) \right) + 12m(\beta^2 - 2\beta - 2). \quad (25)$$

In order to obtain final asymptotic error constant term, we insert the expressions (19), (21), (23) and (25) in (18). Then, we have

$$e_{n+1} = \frac{c_1 \left(c_1^2 (m^2 - H''(1)(\alpha - \beta)^2 + (4\beta + 9)m) - 2c_2m^2 \right)}{24m^9} \left[c_1^4 \left\{ (\alpha - \beta)^2 (3(6\beta^2 - 8\beta + 15)H''(1) - 2(3\beta - 2)(\alpha - \beta)H'''(1)) - m(24\beta^3 - 48\beta^2 + 180\beta + 3H''(1)(\alpha - \beta)^2 + 433) + 6(2\beta + 1)m^2 + 7m^3 \right\} - 6c_2c_1^2m(4m^2 - H''(1)(\alpha - \beta)^2 + (4\beta + 2)m) + 12c_3c_1m^3 + 12c_2^2m^3 \right] e_n^8 + O(e_n^9). \quad (26)$$

The expression (26) demonstrate that our scheme (1) reaches maximum eighth-order conver-

gence for all α and β (provided $\alpha \neq \beta$) by using only four functional evaluations per full iteration. Hence, it is an optimal scheme in the sense of Kung-Traub conjecture, completing the proof. \square

Some special cases

In this section, we discuss some special cases of our proposed scheme (1) based on different weight functions $H(\nu)$ and $G(\mu)$. Therefore, we have depicted some special cases of the scheme (1) in Table 1. We can also easily obtain several new eighth-order iterative methods for multiple zeros by choosing different kind of weight functions provided they should satisfy the conditions of Theorem 1.

Table 1: Some special cases of the proposed scheme (1).

Cases	$H(\nu)$	$G(\mu)$
Case-1	$\frac{m(\alpha-\beta+2\nu-2)}{\alpha-\beta}$	$m \left[1 + 2\mu + (1 - 2\beta)\mu^2 + 2(\beta^2 - 2\beta - 2)\mu^3 \right]$
Case-2	$\frac{m(\alpha-\beta+2\nu-2)}{\alpha-\beta}$	$\frac{m(2\beta^2\mu+\beta(2-4\mu^2)-(3\mu+1)^2)}{2\beta^2\mu+\beta(2-4\mu)-4\mu-1}$
Case-3	$a_1 + \frac{a_2}{\nu}$ where, $a_1 = -\frac{2m}{\alpha-\beta}$, $a_2 = \frac{m(\alpha-\beta+2)}{\alpha-\beta}$	$m \left[1 + 2\mu + (1 - 2\alpha)\mu^2 + 2(\alpha^2 - 2\alpha - 2)\mu^3 \right]$
Case-4	$a_1 + \frac{a_2}{\nu}$ where, $a_1 = -\frac{2m}{\alpha-\beta}$, $a_2 = \frac{m(\alpha-\beta+2)}{\alpha-\beta}$	$\frac{m(2\alpha^2\mu+\alpha(2-4\mu^2)-(3\mu+1)^2)}{2\alpha^2\mu+\alpha(2-4\mu)-4\mu-1}$
Case-5	$\frac{b_1}{\nu} + \frac{b_2}{1+\nu}$ where, $b_1 = \frac{m(-\alpha+\beta-4)}{\alpha-\beta}$, $b_2 = \frac{4m(\alpha-\beta+2)}{\alpha-\beta}$	$\frac{m}{4} (4 + 8\mu - 2b_3\mu^2 + b_4\mu^3)$ $b_3 = \alpha^2 - 2\alpha(\beta - 3) + \beta^2 - 2\beta - 2,$ $b_4 = 3\alpha^3 - 5\alpha^2(\beta - 2) + \alpha(\beta^2 + 4\beta - 24) + \beta^3 - 6\beta^2 + 8\beta - 16$

Let us remark that the order of the proposed scheme (1) does not depend on the values of α and β (provided $\alpha \neq \beta$). So, these elements can be considered as free parameters in order to analyze the computational results.

Numerical experiments

In this section, we illustrate the efficiency and convergence behavior of our iteration functions for particular cases. Therefore, we use case-1 for $(\alpha = 0, \beta = -2)$, $(\alpha = \frac{1}{2}, \beta = -\frac{3}{2})$, $(\alpha = \frac{1}{4}, \beta = -\frac{7}{4})$ and case-2 for $(\alpha = 0, \beta = -2)$ in expression (1), known by *PM1*, *PM2*, *PM3* and *PM4*, respectively. In this regards, we choose four real life problems having multiple and simple zeros and two standard academic problems with multiple zeros. The details are outline in the examples (1)–(6).

For better comparison of our iterative methods, we consider several existing methods of order six and eight (optimal). Firstly, we compare our methods with a non optimal family of sixth-order iteration functions given by Geum et al. [13], out of them we choose the case 5YD, which is given by

$$\begin{aligned}
 y_n &= x_n - m \frac{f(x_n)}{f'(x_n)}, \quad m \geq 1, \\
 w_n &= x_n - m \left[\frac{(u_n - 2)(2u_n - 1)}{(u_n - 1)(5u_n - 2)} \right] \frac{f(x_n)}{f'(x_n)}, \\
 x_{n+1} &= x_n - m \left[\frac{(u_n - 2)(2u_n - 1)}{(5u_n - 2)(u_n + v_n - 1)} \right] \frac{f(x_n)}{f'(x_n)},
 \end{aligned} \tag{27}$$

where $u_n = \left(\frac{f(y_n)}{f(x_n)}\right)^{\frac{1}{m}}$ and $v_n = \left(\frac{f(w_n)}{f(x_n)}\right)^{\frac{1}{m}}$, is denoted by *GM*.

In addition, we demonstrate comparison of them with an optimal eighth-order iteration function proposed by Behl et al. [8], which is given by (this was one of the best scheme claimed by them):

$$\begin{aligned} y_n &= x_n - m \frac{f(x_n)}{f'(x_n)}, \\ z_n &= y_n - m u_n \frac{f'(x_n)}{f'(x_n)} \left[\frac{1 + \beta u_n}{(\beta - 2)u_n + 1} \right], \\ x_{n+1} &= z_n - u_n v_n \frac{f(x_n)}{f'(x_n)} \left[\frac{1}{2} m \left\{ (2v_n + 1) (4(\beta^2 - 6\beta + 6)u_n^3 + (10 - 4\beta)u_n^2 + 4u_n + 1) + 1 \right\} \right] \end{aligned} \quad (28)$$

where $u_n = \left(\frac{f(y_n)}{f(x_n)}\right)^{\frac{1}{m}}$ and $v_n = \left(\frac{f(z_n)}{f(y_n)}\right)^{\frac{1}{m}}$, is known by *BM*.

Moreover, we compare them with optimal eighth-order iterative methods constructed by Zafar et al. [26]. We choose the following schemes out of them

$$\begin{aligned} y_n &= x_n - m \frac{f(x_n)}{f'(x_n)}, \\ z_n &= y_n - m u_n (6u_n^3 - u_n^2 + 2u_n + 1) \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= w_n - m u_n v_n (1 + 2u_n)(1 + v_n) \left(\frac{2w_n + 1}{A_2 P_0} \right) \frac{f(x_n)}{f'(x_n)} \end{aligned} \quad (29)$$

and

$$\begin{aligned} y_n &= x_n - m \frac{f(x_n)}{f'(x_n)}, \\ z_n &= y_n - m u_n \left(\frac{1 - 5u_n^2 + 8u_n^3}{1 - 2u_n} \right) \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= w_n - m u_n v_n (1 + 2u_n)(1 + v_n) \left(\frac{3w_n + 1}{A_2 P_0 (1 + w_n)} \right) \frac{f(x_n)}{f'(x_n)}, \end{aligned} \quad (30)$$

where $u_n = \left(\frac{f(y_n)}{f(x_n)}\right)^{\frac{1}{m}}$, $v_n = \left(\frac{f(z_n)}{f(y_n)}\right)^{\frac{1}{m}}$, $w_n = \left(\frac{f(z_n)}{f(x_n)}\right)^{\frac{1}{m}}$, with $A_2 = P_0 = 1$ (both schemes (29) and (30)) are known as *FM1* and *FM2*, respectively.

Finally, we also contrast them with another optimal family of eighth-order methods presented by Behl et al. [2], out of them we choose the following methods

$$\begin{aligned} y_n &= x_n - m \frac{f(x_n)}{f'(x_n)}, \\ z_n &= x_n - m u_n (1 + 2u_n) \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= z_n - \frac{u_n w_n}{1 - w_n} \left(\frac{m (u_n (8v_n + 6) + 9u_n^2 + 2v_n + 1)}{4u_n + 1} \right) \frac{f(x_n)}{f'(x_n)} \end{aligned} \quad (31)$$

and

$$\begin{aligned} y_n &= x_n - m \frac{f(x_n)}{f'(x_n)}, \\ w_n &= y_n - m u_n (1 + 2u_n) \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= z_n - \frac{u_n w_n}{1 - w_n} (4u_n^3 - u_n^2 - 2u_n - 2v_n - 1) \frac{f(x_n)}{f'(x_n)}, \end{aligned} \quad (32)$$

where $u_n = \left(\frac{f(y_n)}{f(x_n)}\right)^{\frac{1}{m}}$, $v_n = \left(\frac{f(z_n)}{f(y_n)}\right)^{\frac{1}{m}}$, $w_n = \left(\frac{f(z_n)}{f(x_n)}\right)^{\frac{1}{m}}$, are called by *RM1* and *RM2*, respectively.

In Tables 2 – 3, we display the number of iteration indexes (n), error in the consecutive iterations $|x_{n+1} - x_n|$, computational order of convergence (ρ) (we used the formula given by Cordero and Torregrosa [10] in order to calculate ρ) and absolute residual error of the corresponding function ($|f(x_n)|$). We make our calculations with several number of significant digits (minimum 5000 significant digits) to minimize the round off error.

As we mentioned in the above paragraph we calculate the values of all the constants and functional residuals up to several number of significant digits. However, due to the limited paper space, we display the value of errors in the consecutive iterations $|x_{n+1} - x_n|$ and absolute residual errors in the function $|f(x_n)|$ up to 2 significant digits with exponent power which are depicted in Tables 2 – 3. Moreover, computational order of convergence is provided up to 5 significant digits. Finally, we mentioned the values of approximated zeros up to 25 significant digits for each of the examples.

All computations have been performed using the programming package *Mathematica* 11 with multiple precision arithmetic. Further, the meaning of $a(\pm b)$ is shorthand for $a \times 10^{(\pm b)}$ in the Tables 2–3.

Example 1 Fractional conversion in a chemical reactor:

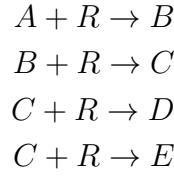
Let us consider the following expression (for the details of this problem please see [20])

$$f_1(x) = \frac{x}{1-x} - 5 \log \left[\frac{0.4(1-x)}{0.4-0.5x} \right] + 4.45977, \quad (33)$$

In the above expression x represents the fractional conversion of species A in a chemical reactor. Since, there will be no physical meaning of above fractional conversion if x is less than zero or greater than one. In this sense, x is bounded in the region $0 \leq x \leq 1$. In addition, our required zero to this problem is $\xi = 0.7573962462537538794596413$. Moreover, it is interesting to note that the above expression will be undefined in the region $0.8 \leq x \leq 1$ which is very close to our desired zero. Furthermore, there are some other properties to this function which make the solution more difficult. The derivative of the above expression will be very close to zero in the region $0 \leq x \leq 0.5$ and there is an infeasible solution for $x = 1.098$.

Example 2 Continuous stirred tank reactor (CSTR)

Let us consider the isothermal continuous stirred tank reactor (CSTR). Components A and R are fed to the reactor at rates of Q and $q - Q$, respectively. Then, we obtain the following reaction scheme in the reactor (for the details see [9]):



The problem was analyzed by Douglas [11] in order to design simple feedback control systems. He presented the following expression for the transfer function of the reactor

$$K_C \frac{2.98(x + 2.25)}{(x + 1.45)(x + 2.85)^2(x + 4.35)} = -1,$$

where K_C is the gain of the proportional controller. The control system is stable for values of K_C that yields roots of the transfer function having negative real part. If we choose $K_C = 0$ we get the poles of the open-loop transfer function as roots of the nonlinear equation:

$$f_2(x) = x^4 + 11.50x^3 + 47.49x^2 + 83.06325x + 51.23266875. \quad (34)$$

No doubts, the above function f_2 has four zeros $\xi = -1.45, -2.85, -2.85, -4.35$. However, our required zero is $\xi = -4.35$ for expression (34).

Example 3 Van der Waals equation of state

$$\left(P + \frac{a_1 n^2}{V^2}\right)(V - na_2) = nRT,$$

explains the behavior of a real gas by introducing in the ideal gas equations two parameters, α_1 and α_2 , specific for each gas. The determination of the volume V of the gas in terms of the remaining parameters requires the solution of a nonlinear equation in V

$$PV^3 - (na_2P + nRT)V^2 + \alpha_1 n^2 V - \alpha_1 \alpha_2 n^2 = 0.$$

Given the constants α_1 and α_2 of a particular gas, one can find values for n , P and T , such that this equation has a three simple roots. By using the particular values, we obtain the following nonlinear function

$$f_3(x) = x^3 - 5.22x^2 + 9.0825x - 5.2675.$$

have three zeros and out of them one is a multiple zero $\xi = 1.75$ of multiplicity of order two and other one simple zero $\xi = 1.72$. However, our desired root is $\xi = 1.75$.

Example 4 Multifactor problem

An undesirable RF breakdown which may happen in the high power microwave devices working under the vacuum condition is known as is multifactor [1]. For example, multifactor appears inside a parallel plate waveguide. There exists an electric field with an electric potential difference which creates the electron movement between these two plates. An interesting case in the study of the electron trajectories is when the electron reaches a plate with root of multiplicity 2. The trajectory of an electron in the air gap between two parallel plates is as follows

$$y(t) = y_0 + (v_0 + e \frac{E_0}{m\omega} \sin(\omega t_0 + \alpha))(t - t_0) + e \frac{E_0}{m\omega^2} (\cos(\omega t + \alpha) - \cos(\omega t_0 + \alpha)) \quad (35)$$

where m and e are the mass and charge of the electron at rest, $E_0 \sin(\omega t + \alpha)$ is the RF electric field between plates and y_0 and v_0 are the position and velocity of the electron at time t_0 . We consider the following particular case of (35), where the parameters have been normalized:

$$f_4(x) = x + \cos(x) - \frac{\pi}{2} \quad (36)$$

with the zero $\xi = \frac{\pi}{2}$ of multiplicity 3.

Example 5 Let us consider a polynomial equation similar to [18], which is given by

$$f_5(x) = ((x - 1)^3 - 1)^{100}. \quad (37)$$

The above function has one multiple zero at $\xi = 2$ of multiplicity $m = 100$.

Example 6 Let us consider the following standard nonlinear test function from Behl et al. [6]

$$f_6(x) = \left(1 - \sqrt{1 - x^2} + x + \cos\left(\frac{\pi x}{2}\right)\right)^3. \quad (38)$$

The above function has a multiple zero at $\xi = -0.7285840464448267167123331$ of multiplicity 3.

Conclusion

In this study, we propose a new eighth-order iteration function having optimal eight-order convergence for multiple zeros of a univariate function with faster convergence, simple and compact body structure. The construction of the present scheme is based on the weight function approach that play an important role in the establishment of eighth-order convergence. In addition, we presented an extensive convergence analysis with the main theorem which clearly show the eighth-order convergence. Each member of our scheme is optimal in the sense of the classical Kung-Traub conjecture. The computational efficiency index is defined as $E = p^{1/\theta}$, where p is the order of convergence and θ is the number of functional evaluations per iteration. Thus, the efficiency index of the present methods is $E = \sqrt[4]{8} \approx 1.682$ which is better than the classical Newton's method $E = \sqrt[3]{2} \approx 1.414$.

Moreover, we can easily obtain several new methods by considering different weight functions in our scheme (1). Lower residual errors, lower error among two consecutive iteration and stable computational order of convergence belongs to our methods when we compared them to the existing ones of same order on problem like chemical conversion, continuous stirred tank reactor, Van der Waals equation of state, multi factor problem, etc. Finally, on accounts of the results obtained, it can be concluded that our proposed methods are highly efficient and perform better than the existing methods.

References

- [1] Anza, S., Vicente, C., Gimeno, B., Boria, V.E. and Armendáriz, J. (2007) Long-term multipactor discharge in multicarrier systems. *Physics of Plasmas* **14**(8) 82–112.
- [2] Behl, R., Alshomrani, A.S. and Motsa, S.S. (2018) An optimal scheme for multiple roots of nonlinear equations with eighth-order convergence. *J. Math. Chem.* **56**(7), 2069–2084.
- [3] Behl, R., Cordero, A., Motsa, S.S. and Torregrosa, J.R. (2018) An eighth-order family of optimal multiple root finders and its dynamics. *Numer. Algor.* **77**(4), 1249–1272.
- [4] Behl, R., Cordero, A., Motsa, S.S. and Torregrosa, J.R. (2015) On developing fourth-order optimal families of methods for multiple roots and their dynamics. *Appl. Math. Comput.* **265**, 520–532.

Table 2: Difference between two consecutive iterations ($|x_{n+1} - x_n|$) of different iteration functions.

$f(x)$	n	GM	BM	$FM1$	$FM2$	$RM1$	$RM2$	$PM1$	$PM2$	$PM3$	$PM4$
$f_1(x)$	1	1.8(-10)	5.1(-12)	5.1(-11)	7.7(-11)	8.0(-12)	1.4(-11)	8.2(-13)	9.4(-13)	2.9(-12)	1.3(-14)
	2	1.7(-53)	1.2(-81)	1.6(-72)	5.9(-71)	1.4(-79)	9.4(-78)	8.1(-89)	5.8(-88)	9.7(-84)	4.3(-105)
	3	1.3(-311)	1.5(-638)	1.5(-564)	7.3(-552)	1.2(-621)	4.7(-607)	7.3(-697)	1.3(-689)	1.4(-655)	7.4(-829)
	ρ	6.0000	8.0000	8.0000	8.0000	8.0000	8.0000	8.0000	8.0000	8.0000	8.0000
$f_2(x)$	1	9.5(-3)	2.0(-2)	2.0(-2)	2.0(-2)	2.7(-4)	2.7(-4)	2.0(-2)	2.0(-2)	2.0(-2)	2.0(-2)
	2	8.1(-16)	4.2(-18)	5.2(-18)	5.2(-18)	9.1(-14)	9.1(-14)	4.2(-18)	4.2(-18)	4.2(-18)	4.2(-18)
	3	3.9(-94)	3.1(-143)	1.9(-142)	1.7(-142)	3.4(-42)	3.4(-42)	3.0(-143)	3.0(-143)	3.0(-143)	3.0(-143)
	ρ	5.9929	7.9858	7.9846	7.9847	3.0005	3.0005	7.9862	7.9861	7.9861	7.9862
$f_3(x)$	1	3.9(-4)	2.6(-4)	3.9(-4)	4.1(-4)	2.6(-4)	2.7(-4)	7.2(-5)	2.9(-5)	5.1(-5)	3.3(-5)
	2	1.0(-14)	3.6(-19)	5.2(-17)	9.8(-17)	1.4(-19)	1.1(-18)	9.4(-24)	1.1(-27)	3.9(-25)	2.4(-27)
	3	3.9(-78)	6.1(-138)	5.9(-120)	1.2(-117)	1.0(-141)	6.1(-134)	8.0(-175)	3.3(-207)	5.0(-186)	7.5(-207)
	ρ	5.9975	7.9977	7.9945	7.9941	8.0026	7.9971	7.9994	7.9996	7.9995	7.9995
$f_4(x)$	1	2.5(-6)	4.3(-6)	4.3(-6)	4.3(-6)	1.4(-10)	1.4(-10)	1.3(-17)	1.3(-17)	1.3(-17)	1.3(-17)
	2	1.5(-18)	1.4(-30)	1.4(-30)	1.4(-30)	3.8(-52)	3.8(-52)	1.4(-30)	1.4(-30)	1.4(-30)	1.4(-30)
	3	3.7(-55)	5.9(-153)	5.9(-153)	5.9(-153)	5.3(-260)	5.3(-260)	5.9(-153)	5.9(-153)	5.9(-153)	5.9(-153)
	ρ	3.0000	5.0000	5.0000	5.000	5.0000	5.0000	5.0000	5.0000	5.0000	5.0000
$f_5(x)$	1	2.0(-7)	9.5(-8)	4.8(-7)	6.5(-7)	6.3(-8)	1.9(-7)	5.1(-8)	2.3(-8)	2.7(-8)	1.5(-8)
	2	1.8(-41)	1.6(-55)	5.7(-49)	8.4(-48)	4.2(-57)	8.0(-53)	8.0(-58)	2.6(-59)	4.6(-60)	1.7(-15)
	3	1.0(-245)	1.3(-437)	2.2(-384)	6.6(-375)	5.9(-169)	9.6(-416)	2.9(-456)	3.2(-454)	3.2(-474)	1.9(-118)
	ρ	6.0000	8.0000	8.0000	8.0000	2.2745	8.0000	8.0000	8.0000	8.0000	14.862
$f_6(x)$	1	3.5(-6)	1.7(-7)	2.4(-7)	2.4(-7)	9.3(-8)	9.7(-8)	1.2(-7)	1.2(-7)	1.2(-7)	1.1(-7)
	2	1.2(-32)	4.4(-53)	2.0(-51)	2.5(-51)	3.0(-55)	5.8(-55)	1.2(-54)	1.2(-54)	1.2(-54)	2.6(-55)
	3	1.8(-191)	9.4(-418)	5.3(-404)	3.6(-403)	3.1(-435)	1.0(-432)	8.7(-431)	8.7(-431)	1.3(-430)	2.8(-436)
	ρ	6.0000	8.0000	8.0000	8.0000	8.0000	8.0000	8.0000	8.0000	8.0000	8.0000

(* means the corresponding fails to work. ** means computational order of convergence is not calculated in the case divergence.)

Table 3: Comparison based on residual error (i.e. $|f(x_n)|$) of different iteration functions.

$f(x)$	n	GM	BM	$FM1$	$FM2$	$RM1$	$RM2$	$PM1$	$PM2$	$PM3$	$PM4$
$f_1(x)$	1	1.4(-8)	4.1(-10)	4.1(-9)	6.1(-9)	6.4(-10)	1.1(-9)	6.6(-11)	7.5(-11)	2.3(-10)	1.0(-12)
	2	1.4(-51)	9.9(-80)	1.3(-70)	4.7(-69)	1.1(-77)	7.5(-76)	6.5(-87)	4.7(-86)	7.8(-82)	3.4(-103)
	3	1.0(-309)	1.2(-636)	1.2(-562)	5.8(-550)	9.7(-620)	3.8(-605)	5.8(-695)	1.0(-687)	1.1(-653)	5.9(-827)
$f_2(x)$	1	1.9(-4)	8.0(-4)	8.5(-4)	8.5(-4)	1.5(-7)	1.5(-7)	8.0(-4)	8.0(-4)	8.0(-4)	8.0(-4)
	2	1.4(-30)	3.7(-35)	5.7(-35)	5.6(-35)	1.7(-26)	1.7(-26)	3.7(-35)	3.7(-35)	3.7(-35)	3.7(-35)
	3	3.2(-187)	2.0(-285)	7.3(-284)	6.3(-284)	2.5(-83)	2.5(-83)	1.9(-285)	1.9(-285)	1.9(-285)	1.9(-285)
$f_3(x)$	1	4.6(-9)	2.0(-9)	4.6(-9)	5.1(-9)	2.0(-9)	2.3(-9)	1.6(-10)	2.5(-11)	7.7(-11)	3.2(-11)
	2	3.2(-30)	4.0(-39)	8.0(-35)	2.9(-34)	5.9(-40)	3.4(-38)	2.6(-48)	3.3(-56)	4.6(-51)	1.7(-55)
	3	4.6(-157)	1.1(-276)	1.1(-240)	4.3(-236)	3.1(-284)	1.2(-268)	1.9(-350)	3.3(-415)	7.6(-373)	1.7(-410)
$f_4(x)$	1	2.6(-18)	1.3(-17)	1.3(-17)	1.3(-17)	4.7(-31)	4.7(-31)	4.7(-31)	4.7(-31)	4.7(-31)	4.7(-31)
	2	6.2(-55)	5.0(-91)	5.0(-91)	5.0(-91)	9.1(-156)	9.1(-156)	5.0(-91)	5.0(-91)	5.0(-91)	5.0(-91)
	3	8.4(-165)	3.5(-458)	3.5(-458)	3.5(-458)	2.4(-779)	2.4(-779)	3.5(-458)	3.5(-458)	3.5(-458)	3.5(-458)
$f_5(x)$	1	1.1(-622)	2.2(-655)	4.4(-585)	5.3(-572)	3.9(-673)	3.1(-626)	4.5(-682)	1.3(-709)	2.5(-709)	3.7(-736)
	2	9.7(-4027)	1.4(-5431)	1.2(-4777)	8.7(-4661)	11.(-5590)	9.1(-5163)	1.6(-5662)	6.7(-5376)	2.3(-5886)	5.3(-1429)
	3	5.4(-24451)	4.1(-43641)	2.7(-38318)	5.1(-37371)	1.1(-16775)	7.8(-41455)	3.3(-45506)	6.1(-41287)	1.6(-47302)	5.9(-11726)
$f_6(x)$	1	1.1(-6)	1.3(-20)	3.5(-20)	3.7(-20)	2.1(-21)	2.3(-21)	4.8(-21)	4.8(-21)	4.5(-21)	3.5(-21)
	2	4.3(-96)	2.2(-157)	2.1(-152)	4.2(-152)	6.7(-164)	5.1(-163)	4.3(-162)	4.3(-162)	4.6(-162)	4.7(-164)
	3	1.5(-572)	2.1(-1251)	3.8(-1210)	1.2(-1207)	7.5(-1304)	2.5(-1296)	1.7(-1290)	1.7(-1290)	6.2(-1290)	5.4(-1307)

- [5] Behl, R., Cordero, A., Motsa, S.S. and Torregrosa, J.R. (2016) V. Kanwar, An optimal fourth-order family of methods for multiple roots and its dynamics. *Numer. Algor.* **71** (4), 775–796.
- [6] Behl, R., Cordero, A., Motsa, S.S., Torregrosa, J.R. and Kanwar, V. (2016) An optimal fourth-order family of methods for multiple roots and its dynamics. *Numer. Algor.* **71** (4), 775–796.
- [7] Burden, R.L. and Faires, J.D. (2001) Numerical Analysis. *PWS Publishing Company, Boston*.
- [8] Behl, R., Zafar, F., Alshomrani, A.S., Junjuaz, M. and Yasmin, N. (2018) An optimal eighth-order scheme for multiple zeros of univariate functions. *Int. J. Comput. Meth.* <https://doi.org/10.1142/S0219876218430028>.
- [9] Constantinides, A. and Mostoufi, N. (1999) Numerical Methods for Chemical Engineers with MATLAB Applications, *Prentice Hall PTR, New Jersey*.
- [10] Cordero, A. and Torregrosa, J.R. (2007) Variants of Newton's method using fifth-order quadrature formulas. *Appl. Math. Comput.* **190** (1), 686–698.
- [11] Douglas, J.M. (1972) Process Dynamics and Control, vol. 2 *Prentice Hall, Englewood Cliffs*.
- [12] Geum, Y.H., Kim, Y.I. and Neta, B. (2015) A class of two-point sixth-order multiple-zero finders of modified double-Newton type and their dynamics. *Appl. Math. Comput.* **270**, 387–400.
- [13] Geum, Y.H., Kim, Y.I. and Neta, B. (2016) A sixth-order family of three-point modified Newton-like multiple-root finders and the dynamics behind their extraneous fixed points. *Appl. Math. Comput.* **283**, 120–140.
- [14] Hueso, J.L., Martinez, E. and Teruel, C. (2015) Determination of multiple roots of nonlinear equations and applications. *J. Math. Chem.* **53**, 880–892.
- [15] Li, S.G., Cheng, L.Z. and Neta, B. (2010) Some fourth-order nonlinear solvers with closed formulae for multiple roots. *Comput. Math. Appl.* **59**, 126–135.
- [16] Neta, B. (2010) Extension of Murakami's high-order non-linear solver to multiple roots. *Int. J. Comput. Math.* **87**(5), 1023–1031.
- [17] Ostrowski, A.M. (1964) Solutions of equations and system of equations. *Academic press, New York*.
- [18] Ortega, J.M. and Rheinboldt, W.C. (1970) iterative solution of nonlinear equations in several variables. *Academic Press, New-York*.
- [19] Petković, M.S., Neta, B., Petković, L.D. and Džunić, J.(2012) Multipoint methods for solving nonlinear equations. *Academic Press*.
- [20] Shacham, M. (1986) Numerical solution of constrained nonlinear algebraic equations. *Int. J. Numer. Method Eng.* **23**, 1455–1481.
- [21] Sharifi, M. and Babajee, D.K.R. (2012) F. Soleymani, Finding the solution of nonlinear equations by a class of optimal methods. *Comput. Math. Appl.* **63**, 764–774.
- [22] Soleymani, F. and Babajee, D.K.R. (2013) Computing multiple zeros using a class of quartically convergent methods. *Alex. Eng. J.* **52**, 531–541.
- [23] Soleymani, F., Babajee, D.K.R. and Lofti, T. (2013) On a numerical technique for finding multiple zeros and its dynami. *J. Egypt. Math. Soc.* **21**, 346–353.
- [24] Thukral, R. (2013) Introduction to higher-order iterative methods for finding multiple roots of nonlinear equations. *J. Math.* **Article ID 404635**, **3** pages, <http://dx.doi.org/10.1155/2013/404635>.
- [25] Traub, J.F. (1964) Iterative methods for the solution of equations. *Prentice-Hall, Englewood Cliffs*.
- [26] Zafar, F., Cordero, A., Motsa, S.S. and Torregrosa, J.R. (2017) Optimal iterative methods for finding multiple roots of nonlinear equations using free parameters. *J. Math. Chem.* DOI 10.1007/s10910-017-0813-1.
- [27] Zhou, X., Chen, X. and Song, Y. (2011) Constructing higher-order methods for obtaining the multiple roots of nonlinear equations. *J. Comput. Math. Appl.* **235**, 4199–4206.
- [28] Zhou, X., Chen, X. and Song, Y. (2013) Families of third and fourth order methods for multiple roots of nonlinear equations. *Appl. Math. Comput.* **219**, 6030–6038.

Identification of Neumann boundary condition assuring the destruction of target region of biological tissue

†Lukasz Turchan and Ewa Majchrzak

Institute of Computational Mechanics and Engineering, Silesian University of Technology, Poland.

*Presenting and corresponding author: lukasz.turchan@polsl.pl

Abstract

An axially symmetrical domain of biological tissue exposed to an external heat source is analyzed. The temperature field is described using the dual-phase lag equation supplemented by Neumann boundary conditions and initial conditions. At first, the direct problem is solved using the implicit scheme of finite difference method. Based on the determined temperatures, the Arrhenius integral is calculated. Next, the inverse problem related to the identification of Neumann boundary condition assuring the destruction of target region of biological tissue is considered. The inverse problem is solved using the gradient method. In the final part of the paper the results of computations and conclusions are presented.

Keywords: Bioheat transfer, Dual-phase lag model, Arrhenius integral, Inverse problem

Introduction

Controlled or uncontrolled heating of biological tissues can lead to their destruction. One of the mathematical methods for assessing the degree of tissue destruction is the so-called Arrhenius integral [1]-[3]. The estimation of its value requires the knowledge of temperature history at the selected set of points (measuring points) from the domain considered. To determine the temperature distribution, the different mathematical models can be used e.g. Pennes equation [4]-[9], Cattaneo-Vernotte model [10]-[12], dual-phase lag model [13]-[18] or generalized dual phase lag model [19]-[22]. These equations should be supplemented by appropriate boundary and initial conditions. The solution obtained using the selected model and the suitable numerical method allows one to determine the temperature distribution in the biological tissue and then the values of the Arrhenius integral.

In the inverse problem considered here, the knowledge of the Arrhenius integral at the set of measuring points of the domain is assumed and on this basis the parameters of the Neumann boundary condition are identified. This procedure can be helpful in planning artificial hyperthermia treatment, because it allows to predict the amount of necessary heat delivered to the tissue that ensures the destruction of the target region.

Direct problem

An axially symmetrical domain of biological tissue exposed to an external heat source is considered. Thermal processes can be described by dual-phase lag model [14], [18], [21]

$$(r, z) \in \Omega: \quad c\rho \left[\frac{\partial T(r, z, t)}{\partial t} + \tau_q \frac{\partial^2 T(r, z, t)}{\partial t^2} \right] = \lambda \left[\nabla^2 T(r, z, t) + \tau_T \frac{\partial}{\partial t} \nabla^2 T(r, z, t) \right] + w c_b \rho_b [T_b - T(r, z, t)] + Q_m \quad (1)$$

where c , c_b are the specific heat of tissue and blood, respectively, ρ , ρ_b are the mass density of tissue and blood, λ is the thermal conductivity of tissue, τ_q is the relaxation time, τ_T is the thermalization time, w is the perfusion coefficient, T_b is the arterial blood temperature, T is the tissue temperature, Q_m is the metabolic heat source, r , z , t denote the spatial coordinates and time.

On the upper surface of domain, the Neumann condition is assumed

$$q_b(r, 0, t) = \begin{cases} q_0 \frac{t}{t_e} \left(1 - \frac{t}{t_e}\right) \exp\left(-\frac{r^2}{r_D^2}\right), & t \leq t_e \\ 0, & t > 0 \end{cases} \quad (2)$$

where q_0 is the constant value and t_e is the exposure time, while $r_D \leq R$ where R is the radius of cylinder.

On the remaining boundaries the no-heat flux conditions can be accepted.

The initial conditions are also known

$$t = 0: \quad T(r, z, 0) = T_0, \quad \left. \frac{\partial T(r, z, t)}{\partial t} \right|_{t=0} = u(r, z) \quad (3)$$

where T_0 is the constant initial temperature of tissue and $u(r, z)$ is the initial heating rate.

The thermal damage parameter can be evaluated according to the Arrhenius integral [1]-[3], [23]

$$A(r, z, t^f) = P \int_0^{t^f} \exp\left(-\frac{E}{R_g T(r, z, t)}\right) dt \quad (4)$$

where P [1/s] is the pre-exponential factor, E [J/mole] is the activation energy, R_g [J/(mole K)] is the universal gas constant, $T(r, z, t)$ [K] is the tissue temperature and $[0, t^f]$ is the time interval under consideration.

A value of damage integral $A(r, z, t^f) = 1$ corresponds to a 63% probability of cell death at a specific point, while $A(r, z, t^f) = 4.6$ corresponds to 99% probability of cell death at this point.

Inverse problem

The inverse problem formulated here concerns the estimation of the boundary heat flux (2), more specifically, the values q_0 and t_e , which ensure the destruction of target region of biological tissue. Thus, the following criterion is formulated

$$S(q_0, t_e) = \sum_{f=1}^F \sum_{i=1}^M \left[A(r_i, z_i, t^f, q_0, t_e) - A_m(r_i, z_i, t^f) \right]^2 \quad (5)$$

where $A_m(r_i, z_i, t^f)$ is the ‘measured’ Arrhenius integral. $A(r_i, z_i, t^f, q_0, t_e)$ is the calculated Arrhenius integral obtained from the direct problem solution with the current estimation of the unknown parameters q_0 and t_e , while M is the number of points and F is the number of time steps.

In the case of typical gradient method application [24], [25] the criterion (5) is differentiated with respect to the unknown parameters q_0 , t_e and next the necessary condition of optimum is used. Finally, one obtains the following system of equations

$$\begin{aligned}\frac{\partial S(q_0, t_e)}{\partial q_0} &= 2 \sum_{f=1}^F \sum_{i=1}^M (A_i^f - A_{m,i}^f) \frac{\partial A(r_i, z_i, t^f, q_0, t_e)}{\partial q_0} = 0 \\ \frac{\partial S(q_0, t_e)}{\partial t_e} &= 2 \sum_{f=1}^F \sum_{i=1}^M (A_i^f - A_{m,i}^f) \frac{\partial A(r_i, z_i, t^f, q_0, t_e)}{\partial t_e} = 0\end{aligned}\quad (6)$$

where (c.f. equation (4))

$$R_{1,i}^f = \frac{\partial A(r_i, z_i, t^f, q_0, t_e)}{\partial q_0} = P \int_0^{t^f} \frac{E}{R_g T^2(r_i, z_i, t)} \exp\left(-\frac{E}{R_g T(r_i, z_i, t)}\right) \frac{\partial T(r_i, z_i, t)}{\partial q_0} dt \quad (7)$$

$$R_{2,i}^f = \frac{\partial A(r_i, z_i, t^f, q_0, t_e)}{\partial t_e} = P \int_0^{t^f} \frac{E}{R_g T^2(r_i, z_i, t)} \exp\left(-\frac{E}{R_g T(r_i, z_i, t)}\right) \frac{\partial T(r, z, t)}{\partial t_e} dt \quad (8)$$

and $A_i^f = A(r_i, z_i, t^f, q_0, t_e)$, $A_{m,i}^f = A(r_i, z_i, t^f)$.

Function A_i^f is expanded into a Taylor series for the known values of q_0^k and t_e^k , this means

$$A_i^f = (A_i^f)^k + (R_{1,i}^f)^k (q_0^{k+1} - q_0^k) + (R_{2,i}^f)^k (t_e^{k+1} - t_e^k) \quad (9)$$

where k is the number of iteration, q_0^k and t_e^k for $k = 0$ are the arbitrary assumed values of q_0 and t_e , while for $k > 0$ q_0^k and t_e^k result from the previous iteration.

Introducing formula (9) to equations (6) one obtains

$$\begin{aligned}\sum_{f=1}^F \sum_{i=1}^M \left[(A_i^f)^k + (R_{1,i}^f)^k (q_0^{k+1} - q_0^k) + (R_{2,i}^f)^k (t_e^{k+1} - t_e^k) - A_{m,i}^f \right] (R_{1,i}^f)^k &= 0 \\ \sum_{f=1}^F \sum_{i=1}^M \left[(A_i^f)^k + (R_{1,i}^f)^k (q_0^{k+1} - q_0^k) + (R_{2,i}^f)^k (t_e^{k+1} - t_e^k) - A_{m,i}^f \right] (R_{2,i}^f)^k &= 0\end{aligned}\quad (10)$$

it means

$$\begin{aligned}(q_0^{k+1} - q_0^k) \sum_{f=1}^F \sum_{i=1}^M \left[(R_{1,i}^f)^k \right]^2 + (t_e^{k+1} - t_e^k) \sum_{f=1}^F \sum_{i=1}^M (R_{1,i}^f)^k (R_{2,i}^f)^k &= \sum_{f=1}^F \sum_{i=1}^M \left[A_{m,i}^f - (A_i^f)^k \right] (R_{1,i}^f)^k \\ (q_0^{k+1} - q_0^k) \sum_{f=1}^F \sum_{i=1}^M (R_{1,i}^f)^k (R_{2,i}^f)^k + (t_e^{k+1} - t_e^k) \sum_{f=1}^F \sum_{i=1}^M \left[(R_{2,i}^f)^k \right]^2 &= \sum_{f=1}^F \sum_{i=1}^M \left[A_{m,i}^f - (A_i^f)^k \right] (R_{2,i}^f)^k\end{aligned}\quad (11)$$

or in the matrix form

$$\begin{bmatrix} \sum_{f=1}^F \sum_{i=1}^M \left[(R_{1,i}^f)^k \right]^2 & \sum_{f=1}^F \sum_{i=1}^M (R_{1,i}^f)^k (R_{2,i}^f)^k \\ \sum_{f=1}^F \sum_{i=1}^M (R_{1,i}^f)^k (R_{2,i}^f)^k & \sum_{f=1}^F \sum_{i=1}^M \left[(R_{2,i}^f)^k \right]^2 \end{bmatrix} \begin{bmatrix} \Delta q_0^k \\ \Delta t_e^k \end{bmatrix} = \begin{bmatrix} \sum_{f=1}^F \sum_{i=1}^M \left[A_{m,i}^f - (A_i^f)^k \right] (R_{1,i}^f)^k \\ \sum_{f=1}^F \sum_{i=1}^M \left[A_{m,i}^f - (A_i^f)^k \right] (R_{2,i}^f)^k \end{bmatrix} \quad (12)$$

After solving the system of equations (12), the new values of identified parameters are determined using the formulas

$$\begin{aligned} q_0^{k+1} &= q_0^k + \Delta q_0^k \\ t_e^{k+1} &= t_e^k + \Delta t_e^k \end{aligned} \quad (13)$$

The iterative process is continued until the assumed number K of iterations is achieved.

Sensitivity analysis

To solve the inverse problem, the sensitivity functions, it means the partial derivatives of Arrhenius integral and tissue temperature with respect to the parameters q_0 and t_e (c.f. equations (7), (8)) should be determined. For this purpose the direct approach of sensitivity analysis [26]-[29] can be used. Thus, the governing equations are differentiated with respect to the parameter p_s , $s = 1, 2$, where $p_1 = q_0$, $p_2 = t_e$. The differentiation of equation (1) gives

$$\begin{aligned} & c\rho \left\{ \frac{\partial}{\partial p_s} \left[\frac{\partial T(r, z, t)}{\partial t} \right] + \tau_q \frac{\partial}{\partial p_s} \left[\frac{\partial^2 T(r, z, t)}{\partial t^2} \right] \right\} = \\ & \lambda \left\{ \frac{\partial}{\partial p_s} \left[\nabla^2 T(r, z, t) \right] + \tau_T \frac{\partial}{\partial p_s} \left[\frac{\partial}{\partial t} \nabla^2 T(r, z, t) \right] \right\} - w c_b \rho_b \frac{\partial T(r, z, t)}{\partial p_s} \end{aligned} \quad (14)$$

or

$$\begin{aligned} & c\rho \left[\frac{\partial U_s(r, z, t)}{\partial t} + \tau_q \frac{\partial^2 U_s(r, z, t)}{\partial t^2} \right] = \\ & \lambda \left\{ \nabla^2 U_s(r, z, t) + \tau_T \frac{\partial}{\partial t} \left[\nabla^2 U_s(r, z, t) \right] \right\} - w c_b \rho_b U_s(r, z, t) \end{aligned} \quad (15)$$

where

$$U_s(r, z, t) = \frac{\partial T(r, z, t)}{\partial p_s} \quad (16)$$

are the sensitivity functions.

The boundary condition (2) and initial conditions (3) are also differentiated. Thus

$$\frac{\partial q_b(r, 0, t)}{\partial q_0} = \begin{cases} \frac{t}{t_e} \left(1 - \frac{t}{t_e} \right) \exp\left(-\frac{r^2}{r_D^2}\right), & t \leq t_e \\ 0, & t > 0 \end{cases} \quad (17)$$

$$\frac{\partial q_b(r, 0, t)}{\partial t_e} = \begin{cases} q_0 \frac{t}{t_e^2} \left(2 \frac{t}{t_e} - 1 \right) \exp\left(-\frac{r^2}{r_D^2}\right), & t \leq t_e \\ 0, & t > 0 \end{cases} \quad (18)$$

and

$$t = 0: \quad U_s(r, z, 0) = 0, \quad \left. \frac{\partial U_s(r, z, t)}{\partial t} \right|_{t=0} = 0 \quad (19)$$

It should be noted that in the dual-phase lag model the Neumann condition should be formulated in a different way than in the macroscopic Fourier model, namely [16], [17]

$$(r, z) \in \Gamma: \quad -\lambda \left[\mathbf{n} \cdot \nabla T(r, z, t) + \tau_T \frac{\partial [\mathbf{n} \cdot \nabla T(r, z, t)]}{\partial t} \right] = q_b(r, z, t) + \tau_q \frac{\partial q_b(r, z, t)}{\partial t} \quad (20)$$

where \mathbf{n} is the normal outward vector.

Thus, the differentiation of boundary condition (20) gives

$$\begin{aligned} & -\lambda \left\{ \mathbf{n} \cdot \frac{\partial}{\partial p_s} [\nabla T(r, z, t)] + \tau_T \frac{\partial}{\partial p_s} \left[\frac{\partial [\mathbf{n} \cdot \nabla T(r, z, t)]}{\partial t} \right] \right\} = \\ & \frac{\partial q_b(r, z, t)}{\partial p_s} + \tau_q \frac{\partial}{\partial p_s} \left[\frac{\partial q_b(r, z, t)}{\partial t} \right] \end{aligned} \quad (21)$$

or

$$-\lambda \left\{ \mathbf{n} \cdot \nabla U_s(r, z, t) + \tau_T \frac{\partial}{\partial t} [\mathbf{n} \cdot U_s(r, z, t)] \right\} = \frac{\partial q_b(r, z, t)}{\partial p_s} + \tau_q \frac{\partial}{\partial t} \left[\frac{\partial q_b(r, z, t)}{\partial p_s} \right] \quad (22)$$

Finally, the Arrhenius integral (4) is differentiated with respect to the parameter p_s

$$R_s^f(r, z, t^f) = \frac{\partial A(r, z, t^f, p_s)}{\partial p_s} = P \int_0^{t^f} \frac{E}{R_g T^2(r, z, t)} \exp\left(-\frac{E}{R_g T(r, z, t)}\right) U_s(r, z, t) dt \quad (23)$$

Method of solution

A structure of equations (1) and (15) is similar, therefore they can be written in the form

$$c\rho \left[\frac{\partial Z_s(r, z, t)}{\partial t} + \tau_q \frac{\partial^2 Z_s(r, z, t)}{\partial t^2} \right] = \lambda \left\{ \nabla^2 Z_s(r, z, t) + \tau_T \frac{\partial}{\partial t} \left[\nabla^2 Z_s(r, z, t) \right] \right\} - wc_b \rho_b Z_s(r, z, t) + W_s(r, z, t) \quad (24)$$

where $Z_0(r, z, t) = T(r, z, t)$, $Z_s(r, z, t) = U_s(r, z, t)$, $s = 1, 2$ and

$$W_s(r, z, t) = \begin{cases} wc_b \rho_b T_b + Q_m, & s = 0 \\ 0, & s = 1, 2 \end{cases} \quad (25)$$

In a similar way the boundary conditions (20) and (22) can be expressed as

$$(r, z) \in \Gamma: -\lambda \left\{ \mathbf{n} \cdot \nabla Z_s(r, z, t) + \tau_T \frac{\partial}{\partial t} [\mathbf{n} \cdot \nabla Z_s(r, z, t)] \right\} = V_s(r, z, t) \quad (26)$$

where

$$V_s(r, z, t) = \begin{cases} q_b(r, z, t) + \tau_q \frac{\partial q_b(r, z, t)}{\partial t}, & s = 0 \\ \frac{\partial q_b(r, z, t)}{\partial p_s} + \tau_q \frac{\partial}{\partial t} \left[\frac{\partial q_b(r, z, t)}{\partial p_s} \right], & s = 1, 2 \end{cases} \quad (27)$$

The equations (24) – (27) are supplemented by the initial conditions (3) and (19).

The basic problem and additional ones connected with the sensitivity functions are solved using the implicit scheme of finite difference method.

The following approximation of equation (24) is proposed

$$c\rho \left(\frac{Z_{i,j}^f - Z_{i,j}^{f-1}}{\Delta t} + \tau_q \frac{Z_{i,j}^f - 2Z_{i,j}^{f-1} + Z_{i,j}^{f-2}}{(\Delta t)^2} \right) = \lambda \left(\nabla^2 Z_{i,j}^f + \tau_T \frac{\nabla^2 Z_{i,j}^f - \nabla^2 Z_{i,j}^{f-1}}{\Delta t} \right) - wc_b \rho_b Z_{i,j}^f + W_{i,j}^f \quad (28)$$

where

$$\begin{aligned} \nabla^2 Z_{i,j}^f &= \frac{Z_{i,j-1}^f - 2Z_{i,j}^f + Z_{i,j+1}^f}{h^2} + \frac{1}{r_{i,j}} \frac{Z_{i,j+1}^f - Z_{i,j-1}^f}{2h} + \frac{Z_{i-1,j}^f - 2Z_{i,j}^f + Z_{i+1,j}^f}{h^2} \\ \nabla^2 Z_{i,j}^{f-1} &= \frac{Z_{i,j-1}^{f-1} - 2Z_{i,j}^{f-1} + Z_{i,j+1}^{f-1}}{h^2} + \frac{1}{r_{i,j}} \frac{Z_{i,j+1}^{f-1} - Z_{i,j-1}^{f-1}}{2h} + \frac{Z_{i-1,j}^{f-1} - 2Z_{i,j}^{f-1} + Z_{i+1,j}^{f-1}}{h^2} \end{aligned} \quad (29)$$

In the above equations, index s is omitted for simplification.

After mathematical manipulations one has

$$\begin{aligned}
 Z_{i,j}^f = & \frac{\lambda \Delta t (\Delta t + \tau_T)}{M} \left(Z_{i-1,j}^f + Z_{i+1,j}^f + \frac{2r_{i,j} - h}{2r_{i,j}} Z_{i,j-1}^f + \frac{2r_{i,j} + h}{2r_{i,j}} Z_{i,j+1}^f \right) + \\
 & \frac{c\rho h^2 (\Delta t + 2\tau_q) + 4\lambda \Delta t \tau_T}{M} Z_{i,j}^{f-1} - \frac{\lambda \Delta t \tau_T}{M} \left(Z_{i-1,j}^{f-1} + Z_{i+1,j}^{f-1} + \frac{2r_{i,j} - h}{2r_{i,j}} Z_{i,j-1}^{f-1} + \frac{2r_{i,j} + h}{2r_{i,j}} Z_{i,j+1}^{f-1} \right) - \\
 & \frac{c\rho h^2 \tau_q}{M} Z_{i,j}^{f-2} + \frac{h^2 (\Delta t)^2}{M} W_{i,j}^f
 \end{aligned} \quad (30)$$

where

$$M = h^2 \left[c\rho (\Delta t + \tau_q) + w c_b \rho_b (\Delta t)^2 \right] + 4\lambda \Delta t (\Delta t + \tau_T) \quad (31)$$

The approximation of the boundary conditions (26) is the following

$$-\lambda \left\{ \mathbf{n} \cdot \nabla Z_s(r, z, t^f) + \frac{\tau_T}{\Delta t} \left[\mathbf{n} \cdot Z_s(r, z, t^f) - \mathbf{n} \cdot Z_s(r, z, t^{f-1}) \right] \right\} = V_s(r, z, t^f) \quad (32)$$

Thus, one obtains the formulas

- for $j=1, 2, \dots, n-1$

$$Z_{0,j}^f = Z_{1,j}^f - \frac{\tau_T}{\Delta t + \tau_T} (Z_{0,j}^{f-1} - Z_{1,j}^{f-1}) + \frac{h \Delta t}{\lambda (\Delta t + \tau_T)} V_{0,j}^f \quad (33)$$

$$Z_{n,j}^f = Z_{n-1,j}^f + \frac{\tau_T}{\Delta t + \tau_T} (Z_{n,j}^{f-1} - Z_{n-1,j}^{f-1}) - \frac{h \Delta t}{\lambda (\Delta t + \tau_T)} V_{n,j}^f \quad (34)$$

- for $i=1, 2, \dots, n-1$

$$Z_{i,0}^f = Z_{i,1}^f - \frac{\tau_T}{\Delta t + \tau_T} (Z_{i,1}^{f-1} - Z_{i,0}^{f-1}) - \frac{r_{i,0} h \Delta t}{\lambda (\Delta t + \tau_T)} V_{i,0}^f \quad (35)$$

$$Z_{i,n}^f = Z_{i,n-1}^f + \frac{\tau_T}{\Delta t + \tau_T} (Z_{i,n}^{f-1} - Z_{i,n-1}^{f-1}) - \frac{r_{i,n} h \Delta t}{\lambda (\Delta t + \tau_T)} V_{i,n}^f \quad (36)$$

It should be noted that the implicit scheme of finite difference method is unconditionally stable [30].

Results of computations

As axially symmetrical domain of biological tissue is considered ($R = Z = 0.015$ mm). The following values of parameters are assumed [31], [32]: specific heat of tissue $c = 4000$ [J/(kg K)], specific heat of blood $c_b = 3770$ [J/(kg K)], tissue density $\rho = 1000$ [kg/m³], blood density $\rho_b = 1060$ [kg/m³], thermal conductivity of tissue $\lambda = 0.5$ [W/(m K)], blood temperature $T_b = 37$ [°C], blood perfusion rate $w = 0.0005$ [1/s], relaxation and thermalization times $\tau_q = \tau_T = 0.5$ [s], metabolic heat source $Q_m = 250$ [W/m³]. Initial temperature of tissue is equal to $T_0 = 37$ [°C] and initial heating rate equals $u = 0$. Parameters for Arrhenius's integral [33]: $P = 1.98 \cdot 10^{106}$ [1/s] and $E = 6.67 \cdot 10^5$ [J/mol].

At first, the direct problem is solved under the assumption that $t_e = 200$ s and $q_0 = 15\,000$ [W/m²] (c.f. equation (2)) and $r_D = R / 4$. The spatial grid step equals $h = 0.0003$ [m] and the time step $\Delta t = 0.1$ [s].

In Figure 1 the temperature histories at the points $P_1(0.3\text{mm}; 0.3\text{mm})$ and $P_2(2.7\text{mm}; 2.7\text{mm})$ are shown. As can be seen maximum temperature at point P_1 (the most heated node of domain interior) is almost 52 °C. The maximum temperature appears after 120 s, while the exposure time is 100 s. This is due to the relaxation time and the thermalization time that occur in the considered model.

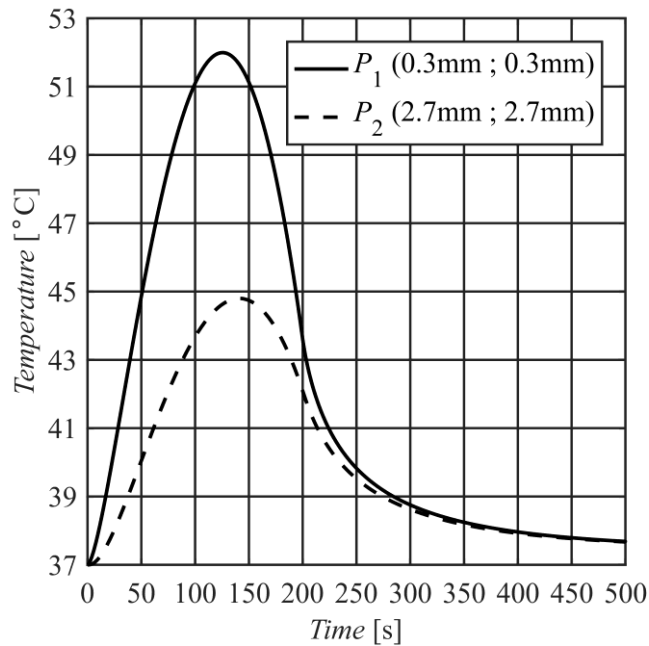


Figure 1. Temperature history at selected points - direct problem

The distribution of Arrhenius integral in the domain considered is presented in Figure 2. It should be noted that Arrhenius integral at the point P_1 is above 4.6, while at the point P_2 is lower than 1.

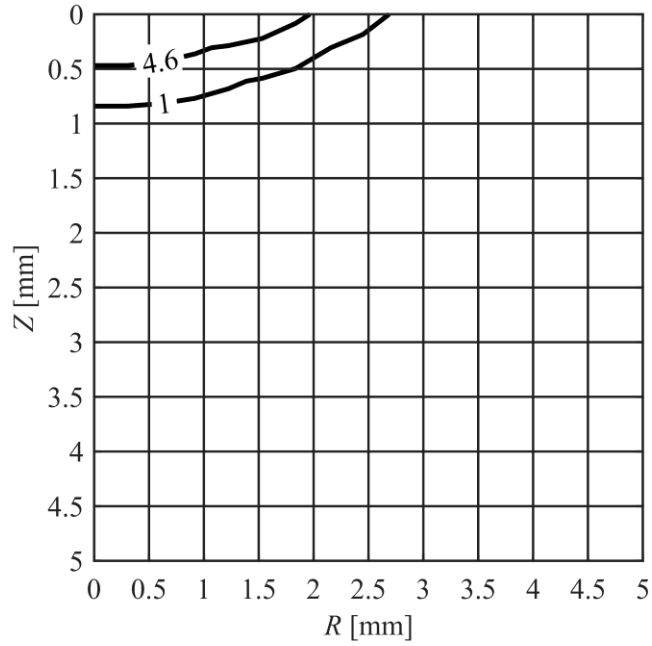


Figure 2. Arrhenius integral distribution

As the measuring points, the fifty points uniformly distributed in the range R : 0 – 2.7 mm and Z : 0 – 2.7 mm were selected. At those points the history of Arrhenius integral (A_m) was saved during solving the direct problem.

Next, three variants of starting values have been assumed:

- L_{V1} : $q_0 = 50\,000\text{ W/m}^2$, $t_e = 500\text{ s}$
- L_{V2} : $q_0 = 13\,000\text{ W/m}^2$, $t_e = 150\text{ s}$
- L_{V3} : $q_0 = 50\,000\text{ W/m}^2$, $t_e = 110\text{ s}$

In Figures 3 – 5 the convergence of the algorithm is presented. It can be seen that for all three variants the concurrence is obtained after different numbers of iterations. When the starting values are close to the values used in direct problem, the algorithm needs fewer numbers of iterations, what can be seen in Figure 4.

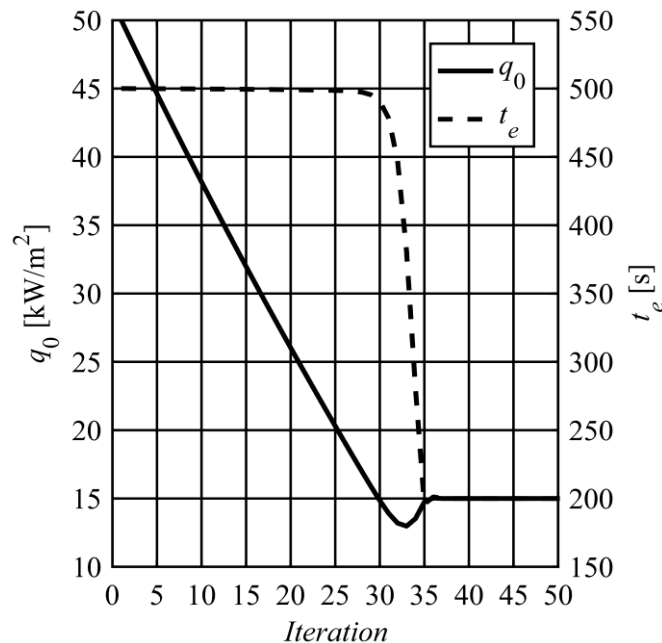


Figure 3. Convergence of the algorithm for the starting values L_{V1}

For the starting values L_{V1} and L_{V3} ($q_0 = 50 \text{ kW/m}^2$) the value of q_0 straightforward goes to the desired value, while the value of t_e for several iterations remains constant, and then quickly changes to the desired value (c.f. Figures 1 and 3).

In Figure 4 the values of both variables in first two iterations increase significantly, while after tenth iteration they are almost concurrent.

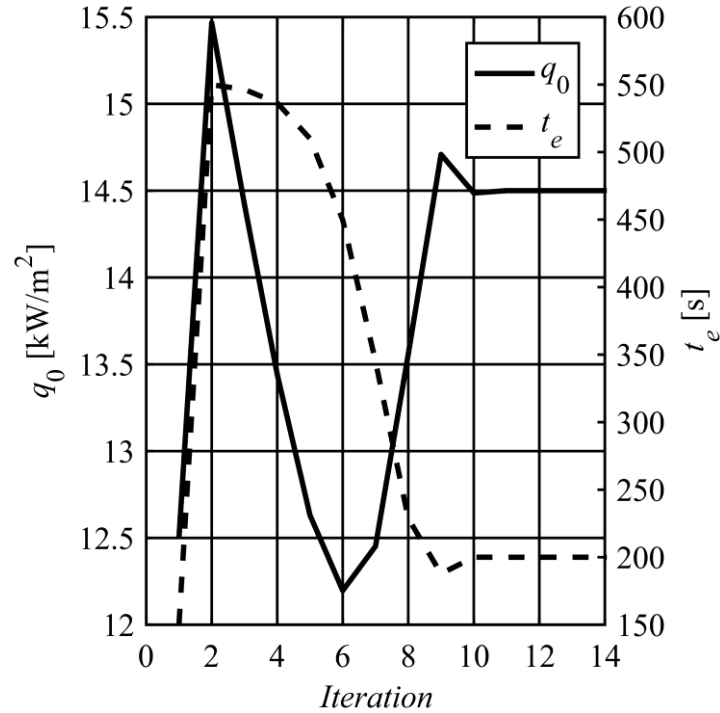


Figure 4. Convergence of the algorithm for the starting values L_{V2}

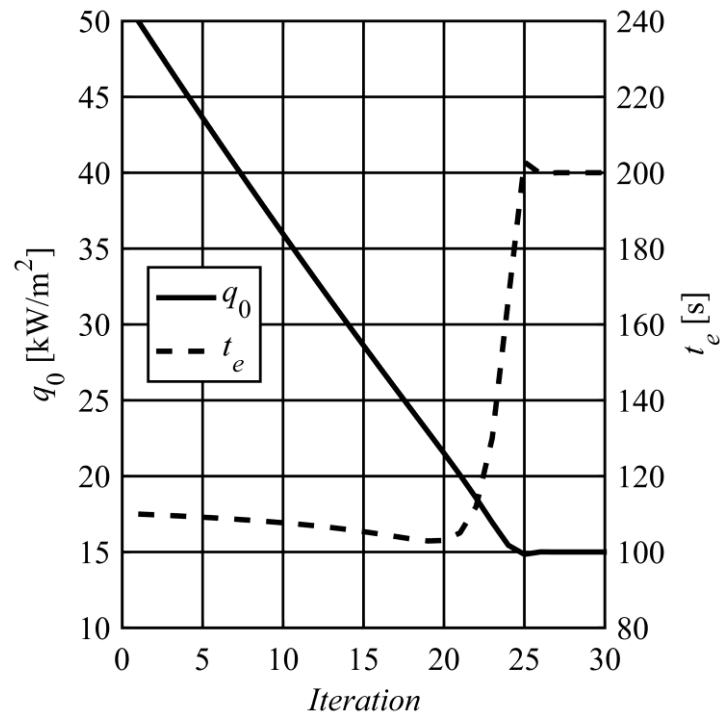


Figure 5. Convergence of the algorithm for the starting values L_{V3}

Conclusions

Thermal processes proceeding in axially symmetrical domain of heated tissue are considered. The inverse problem concerns the estimation of the boundary condition parameters. The mathematical model of the required sensitivity functions is also discussed.

The inverse problem was solved on the basis of the results of direct problem solution. The convergence of algorithm is presented for three different starting sets of values.

The presented approach can be efficiently used for determining the parameters of the artificial hyperthermia treatment.

Acknowledgements

The paper and research are co-financed within the project 2015/19/B/ST8/01101 sponsored by National Science Centre (Poland) and statutory subsidy of the Faculty of Mechanical Engineering of the Silesian University of Technology 10/990/BK_17/0064.

References

- [1] Niemz, M.H. (2007) *Laser-Tissue Interaction: Fundamentals and Applications*, Comptes Rendusdel' Académiedes Sciences-Series I-Mathematics, Springer-Verlag, Berlin, Heidelberg, Germany.
- [2] Sapareto, S.A., Dewey, W.C. (1984) Thermal dose determination in cancer therapy, *International Journal of Radiation Oncology Biology Physics*, 10, 6, 787-800.
- [3] Jasiński, M. (2015) Modeling of thermal damage in laser irradiated tissue, *Journal of Applied Mathematics and Computational Mechanics*, 14, 4, 67-78.
- [4] Pennes, H.H. (1948) Analysis of tissue and arterial blood temperatures in the resting human forearm, *Journal of Applied Physiology*, 1, 93-122.
- [5] Jamil, M., Ng, E.Y.K. (2013) Ranking of parameters in bioheat transfer using Taguchi analysis, *International Journal of Thermal Sciences*, 63, 15-21.
- [6] Ciesielski, M., Mochnacki, B. (2014) Application of the control volume method using the voronoi polygons for numerical modeling of bio-heat transfer processes, *Journal of Theoretical And Applied Mechanics*, 52, 4, 927-935.
- [7] Mochnacki, B., Piasecka-Belkhat, A. (2013) Numerical modeling of skin tissue heating using the interval finite difference method, *MCB: Molecular & Cellular Biomechanics*, 10, 3, 233-244.
- [8] Majchrzak, E., Mochnacki, B., Dziwoński, M., Jasiński, M. (2011) Numerical modelling of hyperthermia and hypothermia processes, *Advanced Materials Research*, 268-270, 257-262.
- [9] Majchrzak, E., Mochnacki, B., Jasiński, M. (2003) Numerical modelling of bioheat transfer in multi-layer skin tissue domain subjected to a flash fire, *Computational Fluid and Solid Mechanics*, 1, 2, 1766-1770.
- [10] Kaminski, W. (1990) Hyperbolic heat conduction equation for materials with a nonhomogeneous inner structure, *Journal of Heat Transfer*, 112, 555-560.
- [11] Mitra, K., Kumar, S., Vedavarz, A., Moallemi, M.K. (1995) Experimental evidence of hyperbolic heat conduction in processed meat, *Journal of Heat Transfer*, 117, 568-573.
- [12] Antaki, P.J. (2005) New interpretation of non-Fourier heat conduction in processed meat, *Journal of Heat Transfer*, 127, 189-193.
- [13] Zhou, J., Chen, J.K., Zhang, Y. (2009) Dual phase lag effects on thermal damage to biological tissues caused by laser irradiation, *Computers in Biology and Medicine*, 39, 286-293.
- [14] Zhou, J., Zhang, Y., Chen, J.K. (2009) An axisymmetric dual-phase lag bio-heat model for laser heating of living tissues, *International Journal of Thermal Sciences*, 48, 8, 1477-1485.
- [15] Liu, K.C., Chen, H.T. (2010) Investigation for the dual phase lag behavior of bioheat transfer, *International Journal of Thermal Sciences*, 49, 1138-1146.
- [16] Majchrzak, E. (2010) Numerical solution of dual phase lag model of bioheat transfer using the general boundary element method, *CMES: Computer Modeling in Engineering & Sciences*, 69, 1, 43-60.
- [17] Majchrzak, E., Turchan, L. (2015) The general boundary element method for 3D dual-phase lag model of bioheat transfer, *Engineering Analysis with Boundary Elements*, 50, 76-82.
- [18] Mochnacki, B., Majchrzak, E. (2017) Numerical model of thermal interactions between cylindrical cryoprobe and biological tissue using the dual-phase lag equation, *International Journal of Heat and Mass Transfer*, 108, 1-10.
- [19] Zhang, Y. (2009) Generalized dual-phase lag bioheat equations based on nonequilibrium heat transfer in living biological tissues, *International Journal of Heat and Mass Transfer*, 52, 4829-4834.

- [20] Afrin, N., Zhou, J., Zhang, Y., Tzou, D.Y., Chen, J.K. (2012) Numerical simulation of thermal damage to living biological tissues induced by laser irradiation based on a generalized dual phase lag model, *Numerical Heat Transfer, Part A: Applications*, 61, 7, 483-501.
- [21] Majchrzak, E., Turchan, L., Dziatkiewicz, J. (2015) Modeling of skin tissue heating using the generalized dual-phase lag equation, *Archives of Mechanics*, 67, 6, 417-437.
- [22] Jasiński, M., Majchrzak, E., Turchan, L. (2016) Numerical analysis of the interactions between laser and soft tissues using dual-phase lag model, *Applied Mathematical Modeling*, 40, 2, 750-762.
- [23] Paruch, M., Turchan, L., Mathematical Modelling of the Destruction Degree of Cancer Under the Influence of a RF Hyperthermia, *AIP Conference Proceedings 1922*, Lublin, Poland 2017, <https://doi.org/10.1063/1.5019064>
- [24] Kurpisz, K., Nowak, A.J. (1995) *Inverse thermal problems*, Computational Mechanics Publications, Southampton, Boston 1995.
- [25] Paruch, M. (2017) Identification of the cancer ablation parameters during RF hyperthermia using gradient, evolutionary and hybrid algorithms, *International Journal of Numerical Methods for Heat & Fluid Flow*, 27, 3, 674-697.
- [26] Kleiber, M. (1997) *Parameter Sensitivity in Non-linear Mechanics*, J. Willey & Sons, London.
- [27] Dems, K., Rousselet, B. (1999) Sensitivity analysis for transient heat conduction in a solid body – Part I, *Structural Optimization*, 17, 36-45.
- [28] Dziewoński, M., Mochnacki, B., Szopa, R. (2011) Sensitivity of biological tissue freezing process on the changes of cryoprobe cooling rate, *Book Series: Mechanika Kaunas University of Technology*, 82-87.
- [29] Kaluza, G., Majchrzak, E., Turchan, L. (2017) Sensitivity analysis of temperature field in the heated soft tissue with respect to the perturbations of porosity, *Applied Mathematical Modelling*, 49, 498-513.
- [30] Majchrzak, E., Mochnacki, B. (2017) Implicit scheme of the finite difference method for 1d dual-phase lag equation, *Journal Of Applied Mathematics And Computational Mechanics*, 16, 3, 37-46.
- [31] Stańczyk, M., Telega, J.J. (2002) Modelling of heat transfer in biomechanics, a review. Part 1. Soft tissues, *Acta of Bioengineering and Biomechanics*, 4, 1, 31-61.
- [32] Huang, H.W., Chen, Z.P., Roemer, R.B. (1996) A counter current vascular network model of heat transfer in tissues, *Journal of Biomechanical Engineering*, 118, 120-129.
- [33] Junmeng, C., Fang, H., Weiming, Y., Fusheng, Y. (2006) A new formula approximating the Arrhenius integral to perform the nonisothermal kinetics, *Chemical Engineering Journal*, 124, 15-18.

A novel approach for regulation of the diffusive effects of limiters in viscous-compressible-flow computations using a boundary-layer sensor

†*Anoop Kumar Dass¹, and Paragmoni Kalita²

¹Professor, Department of Mechanical Engineering, I.I.T. Guwahati, Assam, India-781039

²Assistant Professor, Department of Mechanical Engineering, Tezpur University, Assam, India-784028

*Presenting author: anoop@iitg.ernet.in

†Corresponding author: anoop@iitg.ernet.in

Abstract

Accurate computations of high-speed-viscous flows demand the use of higher-order-accurate schemes for computing the inviscid-flux vectors. However, the constraint of monotonicity preservation calls for the need of limiters in the solution or flux reconstructions used for obtaining higher-order accuracy. The necessity for use of limiters is strong in the inviscid-flow regions owing to the presence of discontinuities like shocks, contact surfaces, slip lines etc. In contrast, the flow field inside the boundary layer is smoother compared with that outside the shear layer in viscous-compressible flows. It is a well-known fact that all the limiters inherently possess diffusive effects like extremum clipping. These diffusive effects of limiters add up with the physical dissipation present inside the boundary layer and spoil the solution accuracy. To overcome this problem, this paper proposes a novel approach to control the limiters used for MUSCL reconstruction using a recently introduced boundary layer sensor. The limiters are switched off inside the boundary layer using the sensor. This approach results in controlling the diffusive effects of the limiters in the higher-order-accurate computation of viscous-compressible flows. The higher accuracy of the new methodology is demonstrated by a number of carefully selected test problems using van Albada limiter.

Keywords: MUSCL, reconstruction, limiter, boundary layer sensor

Introduction

The design of accurate, robust and efficient schemes for computing high-speed flows has been an area of prime importance in the field of computational gasdynamics [1]. The dynamics of inviscid-compressible flows is governed by the Euler equations of gasdynamics. Even for high-speed viscous applications governed by the compressible Navier-Stokes equations, the convective fluxes are numerically computed by using the same flux formulas developed for the Euler equations. The design of a stable and accurate numerical scheme for the convective fluxes is a highly challenging task because the highly nonlinear behaviour of these equations admits discontinuous solutions in the forms of shocks, contact discontinuities, slip surfaces, and also expansion waves with sonic points [2]. Since the unsteady Euler equations are hyperbolic in nature admitting wave-like solution, simple central differencing of the fluxes leads to numerical instabilities necessitating the development of upwind schemes which comprise one-sided differencing that respects the direction of signal propagation. Roe's Flux-Difference Splitting (FDS) [3], van Leer's Flux-Vector Splitting (FVS) [4], Liou and Steffen's Advection Upstream Splitting Method (AUSM) [5], Nishikawa and Kitamura's Rotated Riemann Solvers [6], Residual Distribution (RD) schemes [7-8], multidimensional Riemann solvers [9-12] are examples of some popular upwind methods. It can be shown that an upwind scheme is equivalent to central-space discretization plus an "implicit" numerical

diffusion term. In other words, addition of numerical diffusion plays the role of imparting stability to a purely central space discretization of the convective-flux vectors. Contrary to the upwind schemes, the central schemes choose a symmetric stencil irrespective direction of signal propagation. For numerical stability some artificial diffusion is added to the central-space discretization of the flux. The Lax-Friedrichs [13], Lax-Wendroff [14], Local-Lax Friedrichs [15], Jameson-Schmidt-Turkel (JST) [16] are to mention a few noteworthy central schemes.

While numerical diffusion is essential for stability, excessive diffusion spoils the solution accuracy by smearing the discontinuities and shear layers. Especially for viscous-flow computations excessive numerical diffusion causes the smearing of boundary layers, under prediction of skin friction and wall-heat fluxes [5, 17, 18] and over-prediction of separation-bubble sizes [19]. For example, though van Leer's FVS scheme is robust and accurate for the computation of the Euler equations, it is less accurate than AUSM for viscous-flow computations owing to the more diffusive nature of the former [5]. It may be noted that the numerical diffusion required for stability is high in zones of shocks or sharp gradients, while in smooth-flow regions its requirement is less. These requirements call for the regulation of numerical diffusion based on the smoothness of flow field.

Godunov showed that a monotonous conservative linear scheme can be at best first-order accurate [20]. However, because of excessive numerical diffusion the first-order accurate computations are not preferred for viscous-flow computations. Naturally for accurate computation of viscous flows one needs to go for higher-order-accurate schemes. Nevertheless, the higher-order-accurate reconstructions of the fluxes require the use of nonlinear limiters in order to avoid spurious numerical oscillations [21 (van Leer, 1979)]. The limiters aid in the attainment of monotonous solutions, but they also induce diffusive effects like extremum clipping [22]. Kalita and Dass [18] presented an improved version of the Diffusion-Regulated Local Lax Friedrichs (DRLLF) scheme [23] for viscous computations by scaling down its numerical diffusion inside the boundary layer using a new boundary-layer sensor with different limiters in the higher-order reconstructions.

In viscous flows solution gradients may exist inside the boundary layers. However, these gradients are mild owing to the diffusive effects of physical viscosity. Therefore, a scope exists to control the diffusive effects of limiters inside shear layers in the computation of viscous-compressible flows. To the best of our knowledge, efforts to suppress the action of limiters inside the boundary layers are not reported in the available literature. This work presents a novel approach of controlling the limiters only inside the boundary layers in computations of high-speed-viscous flows on a finite volume framework. This approach switches off the limiters inside the viscous-shear layers using a recently introduced boundary layer sensor [18]. Outside the boundary layer the original-solution-reconstruction approach with full limiting is followed. For the demonstration, the performance of higher-order-accurate AUSM scheme with MUSCL reconstruction [21] using the van Albada limiter [24 (van Albada)] is shown for a number of viscous supersonic and hypersonic test cases. The numerical experiments reveal that the new approach of controlling the diffusive effects of limiters produces more accurate results compared with the higher-order versions of the same scheme with full limiting over the entire flow field.

This paper is organized in four sections. The next section presents the numerical methodology for computing the boundary-layer sensor and an algorithm for controlling the limiters using the same. The improved performance of the new approach is demonstrated in the section on "Numerical Simulations, Results and Discussion" with a number of standard-test cases, before making the concluding remarks in the "Conclusions" section.

The numerical methodology and the algorithm to control limiters

We first introduce the MUSCL reconstruction for a one-dimensional (1D) formulation. The same methodology can be easily extended to multiple dimensions. A 1D-computational domain is shown in Fig. 1. The left and right states for any variable U across any cell-interface $I + \frac{1}{2}$ between the cells I and $I + 1$ are computed by using the MUSCL approach as [25]

$$U_L = \bar{U}_I + \frac{1}{4} \left[(1 + \hat{k}) \Phi_{I+\frac{1}{2}}^- \Delta_{I+\frac{1}{2}} \bar{U} + (1 - \hat{k}) \Phi_{I-\frac{1}{2}}^+ \Delta_{I-\frac{1}{2}} \bar{U} \right] \quad (1)$$

$$U_R = \bar{U}_{I+1} - \frac{1}{4} \left[(1 + \hat{k}) \Phi_{I+\frac{1}{2}}^+ \Delta_{I+\frac{1}{2}} \bar{U} + (1 - \hat{k}) \Phi_{I+\frac{3}{2}}^- \Delta_{I+\frac{3}{2}} \bar{U} \right] \quad (2)$$

where \bar{U} is the cell-averaged value stored at a cell centre, $\Phi_{I+\frac{1}{2}}^\pm$ is a slope limiter, subscripts L and R represent the left and right states of the variable U across the cell-interface, and \hat{k} is an integer that determines the stencil size. The expression $\Delta_{I+\frac{1}{2}} \bar{U}$ is evaluated as

$$\Delta_{I+\frac{1}{2}} \bar{U} = \bar{U}_{I+1} - \bar{U}_I \quad (3)$$

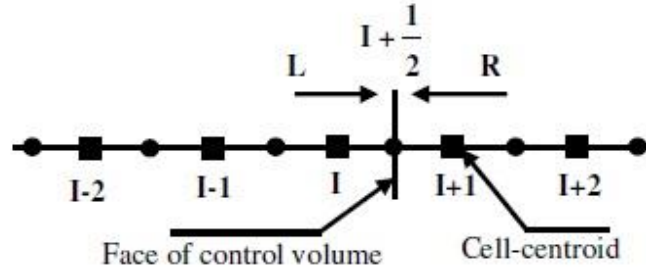


Figure 1. A cell-interface and its left and right states

In the case of van Albada limiter, $\hat{k} = 0$ and the update equations for left and right states are given by [24]

$$U_L = \bar{U}_I + \frac{1}{2} \delta_L \quad (4)$$

$$U_R = \bar{U}_{I+1} - \frac{1}{2} \delta_R \quad (5)$$

where the function δ is typically the same for both the states given by

$$\delta = \frac{a(b^2 + \varepsilon) + b(a^2 + \varepsilon)}{a^2 + b^2 + \varepsilon} \quad (6)$$

so that

$$a_R = \bar{U}_{I+2} - \bar{U}_{I+1}, \quad b_R = \bar{U}_{I+1} - \bar{U}_I \quad (7)$$

$$a_L = \bar{U}_{I+1} - \bar{U}_I, \quad b_L = \bar{U}_I - \bar{U}_{I-1} \quad (8)$$

The additional parameter ε is required in order to prevent the activation of the limiter in smooth-flow regions owing to small-scale oscillations. It has to be set proportional to the local grid scale. Based on extensive numerical experiments the present work considers ε in terms of the cell volume V as

$$\varepsilon = 10 \times (V)^{1.25} \quad (9)$$

Inside the viscous-shear layers owing to the presence of physical diffusion the extrema in the flow field are attained smoothly. This offers an opportunity to switch off the slope limiter $\Phi_{I \pm \frac{1}{2}}^\pm$ inside the boundary layer, provided the presence of the boundary layer is sensed by a suitable sensor. If the slope limiter is switched off, the MUSCL reconstruction becomes

$$U_L = \bar{U}_I + \frac{1}{4} \left[(1 + \hat{k}) \Delta_{I+\frac{1}{2}} \bar{U} + (1 - \hat{k}) \Delta_{I-\frac{1}{2}} \bar{U} \right] \quad (10)$$

$$U_R = \bar{U}_{I+1} - \frac{1}{4} \left[(1 + \hat{k}) \Delta_{I+\frac{1}{2}} \bar{U} + (1 - \hat{k}) \Delta_{I+\frac{3}{2}} \bar{U} \right] \quad (11)$$

In order to regulate the slope limiter, a recently introduced boundary layer sensor is used in the present work. The boundary layer sensor r_{vg} is computed as the absolute ratio of velocity gradient across a cell-interface to the velocity gradient at the solid wall [18].

$$r_{vg} = \left| \frac{(\partial U_{\text{par}} / \partial \eta)_{\text{interface}}}{(\partial U_{\text{par}} / \partial \eta)_{\text{wall}}} \right| \quad (12)$$

where U_{par} is the velocity component parallel to the wall and η is the direction normal to the wall. Literature suggests scaling down the numerical diffusion in the wall-normal direction only for viscous-flow computations [26]. Following the same principle, the present work also suggests to control the action of limiters in the higher-order-accurate computations of fluxes only across the cell-faces that are “aligned” with the flow. The algorithm for switching on and off the limiters for MUSCL reconstruction in the higher-order-accurate computations of fluxes is as follows:

- (i) The parameter r_{vg} is used to track the location of the edge of the boundary layer. At the wall $r_{vg} = 1$. As one moves away from the wall towards the free stream, the value of the boundary-layer sensor decreases asymptotically till it attains a value zero far away from the wall. Thus, a critical height Y_{critical} is identified at the cell for which, $r_{vg} \leq 0.01$.
- (ii) In the cells where $Y < Y_{\text{critical}}$, the flow is considered to be inside the boundary layer, where physical viscosity plays a significant role. Therefore the limiter is switched off during the solution reconstructions within the boundary layer, and computations are carried out using Eq. (10) and Eq. (11).
- (iii) For cells located at $Y \geq Y_{\text{critical}}$ the solution reconstructions are done using Eq. (1) and Eq. (2). If the limiter used is van Albada, then these equations reduce to Eq. (4) and Eq. (5).

Numerical Simulation, Results and Discussion

In the present work, we choose to demonstrate the performance of higher-order-accurate AUSM scheme. The AUSM scheme is selected because of its high accuracy for viscous-flow computations. For illustration the solution reconstruction using the MUSCL approach with van Albada limiter is shown. However, our experience shows that the switching off other limiters inside the viscous-shear zones using the boundary-layer sensor also yield favourable results similar to that of the van Albada limiter. Two standard test cases are shown in the present paper, namely, viscous supersonic flow over an adiabatic flat plate at Mach 3 [18, 27, 28] and hypersonic flow over a ramped surface at Mach 6 [18, 28, 29]. The geometric and free-stream parameters for the two test cases are given in Table 1 and Table 2.

Table 1. The geometric and flow parameters for viscous supersonic flow over a flat plate

Parameter	Value
Length of the plate (L_c)	0.0000285 m
Free-stream pressure (p_∞)	101325 N/m ²
Free-stream temperature (T_∞)	288.15 K
Free-stream Mach number (M_∞)	3

Table 2. The geometric and flow parameters for hypersonic flow over a ramped surface

Parameter	Value
Length of the plate upto the compression corner (L_c)	0.05 m
Total length of the ramped surface	0.12 m
Ramp angle (θ)	15°
Reynolds number per unit length $Re_\infty (= \rho_\infty U_\infty / \mu_\infty)$	8×10 ⁵ m ⁻¹
Free-stream stagnation temperature	1747 K
Free-stream Mach number (M_∞)	6
Wall temperature (T_w)	298 K

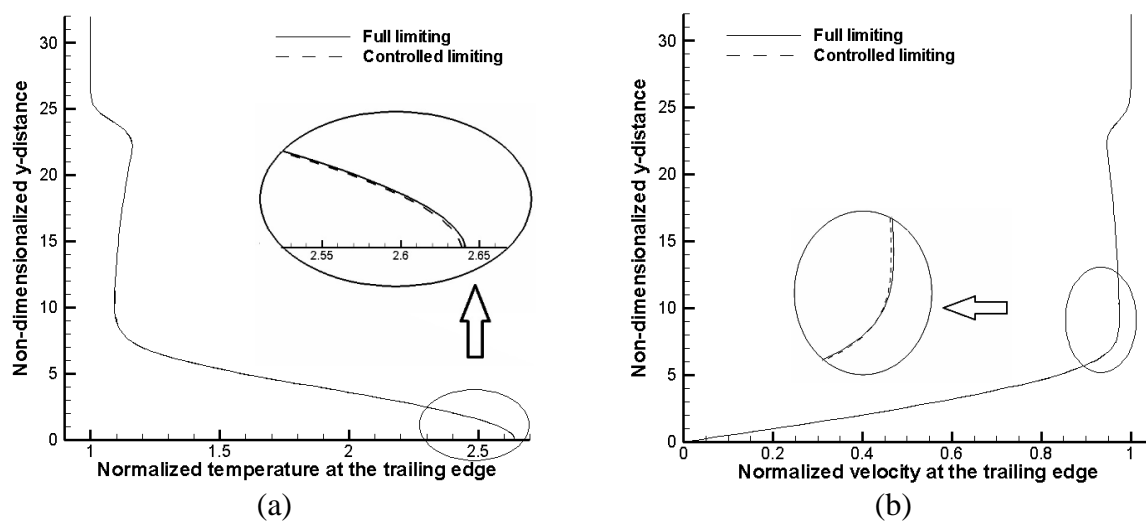


Figure 2. Viscous supersonic flow over an adiabatic-flat plate: (a) temperature profile at the trailing edge (b) velocity profile at the trailing edge

The normalized-temperature (T/T_∞) profiles for viscous supersonic flow over a flat plate under the adiabatic condition are compared in Fig. 2(a). The y-distance is non-dimensionalized as suggested by Van Driest [30]. It can be seen that full limiting predicts a marginally higher adiabatic wall temperature compared with the controlled limiting. This is due to the fact that switching off the limiter inside the boundary layer during controlled limiting results in a lower level of numerical diffusion. The normalized-velocity profiles are shown in Fig. 2(b). As expected, the controlled limiting results in marginally less smearing of the hydrodynamic boundary layer, which is evident from the encircled and zoomed-in portions.

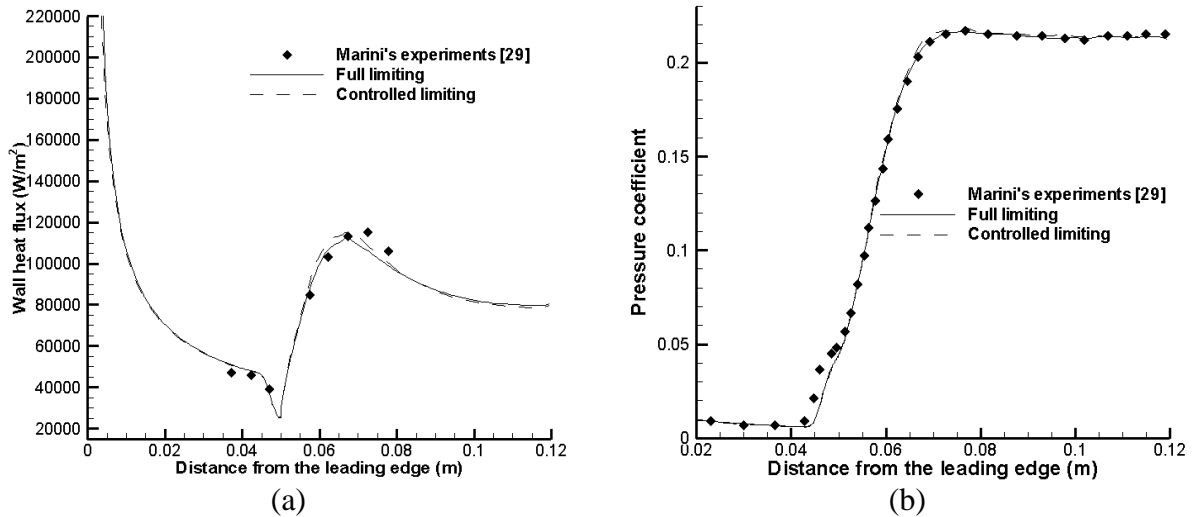


Figure 3. Hypersonic flow over a ramped surface: (a) wall-heat flux along the surface (b) pressure coefficient along the surface

The variations of wall-heat flux from Marini's experimental results for hypersonic flow over a ramped surface are compared with the present computations in Fig. 3(a). With full limiting, the peak-heat flux in the post-reattachment zone is lower than the corresponding value with controlled limiting. In other words, the prediction of wall-heat flux with controlled limiting is in better agreement with the experimental results. The reason for this can be ascribed to the fact that full limiting induces more numerical diffusion compared with controlled limiting. Accordingly, both the computed hydrodynamic and thermal boundary layers are smeared more by the former case than the latter one. The increased smearing of the computed thermal boundary layer results in a lower temperature gradient at the wall, which leads to an under-estimation of the wall-heat flux. The variations of pressure coefficients with both full and controlled limiting are in close agreement with Marini's experiments, as can be seen in Fig. 3(b).

Conclusions

In the present work, a novel approach is proposed to control the diffusive action of limiters inside the boundary layers for computation of viscous compressible flows. The method proposes an algorithm to switch off the limiters inside the viscous shear layers and switch on the same in the inviscid zone using a recently introduced boundary-layer sensor. This is important since the viscous-flow computations demand the minimum possible level of numerical diffusion so as to avoid the smearing of hydrodynamic and thermal boundary layers. The superior performance of the approach is demonstrated by choosing the higher-order AUSM scheme with MUSCL reconstruction using van Albada limiter. Two standard test cases, namely, viscous supersonic flow over an adiabatic flat plate and hypersonic flow

over a ramped surface are used to showcase the improved performance of the new approach. It is shown that controlling the diffusive effects of the limiters inside the boundary layers results in the decrease of smearing of boundary layers, thereby the improvement in accuracy of viscous-compressible-flow computations.

References

- [1] van Leer B. (2006) Upwind and high-resolution methods for compressible flow: From donor cell to residual distribution schemes, *Communications in Computational Physics*, **1**, 192-206.
- [2] Jaisankar S. and Raghurama Rao S.V. (2009) A Central Rankine-Hugoniot solver for hyperbolic conservation laws, *Journal of Computational Physics*, **228**, 770-798.
- [3] Roe P.L. (1981) Approximate Riemann solvers, parameter vectors, and difference schemes, *Journal of Computational Physics*, **43** 357-372.
- [4] van Leer B., Flux-vector splitting for the Euler equations, *Lecture Notes in Physics*, 1982, Springer-Verlag **170** 507-512.
- [5] Liou M.S. and Steffen C.J. (1993) A new flux splitting scheme, *Journal of Computational Physics*, **107** 23-39.
- [6] Nishikawa H. and Kitamura K. (2008) Very simple, carbuncle-free, boundary-layer-resolving, rotated-hybrid Riemann solvers, *Journal of Computational Physics*, **227** 2560-2581.
- [7] Mesaros L.M., Multi-dimensional fluctuation splitting schemes for the Euler equations on unstructured grids, PhD Thesis, University of Michigan, 1995.
- [8] Abgrall R. and Santis D.D. (2015) Linear and non-linear high order accurate residual distribution schemes for the discretization of the steady compressible Navier–Stokes equations, *Journal of Computational Physics*, **283** 329-359.
- [9] Balsara D. S. (2010), Multidimensional HLLE Riemann solver: Application to Euler and magnetohydrodynamic flows, *Journal of Computational Physics*, **229** 1970-1993.
- [10] Balsara D. S. (2012) A two-dimensional HLLC Riemann solver for conservation laws: Application to Euler and magnetohydrodynamic flows, *Journal of Computational Physics*, **231** 7476-7503.
- [11] Balsara D. S. (2014) Multidimensional Riemann problem with self-similar internal structure. Part I—Application to hyperbolic conservation laws on structured meshes, *Journal of Computational Physics*, **277** 163-200.
- [12] Dumbser M. and Balsara D. S. (2016) A new efficient formulation of the HLLEM Riemann solver for general conservative and non-conservative hyperbolic systems, *Journal of Computational Physics*, **304** 275-319.
- [13] Lax P.D. (1954) Weak solutions of nonlinear hyperbolic equations and their numerical computation, *Communications on Pure and Applied Mathematics*, **7** 159-193.
- [14] Lax P.D. and Wendroff B. (1960) Systems of conservation laws, *Communications on Pure and Applied Mathematics*, **13** 217-237.
- [15] LeVeque R.J. (2002) Finite Volume Methods for Hyperbolic Problems (Cambridge Texts in Applied Mathematics), Cambridge University Press.
- [16] Jameson A., Schmidt W. and Turkel E., Numerical solutions of the Euler equations by finite volume methods using Runge-Kutta time stepping schemes, *AIAA paper*, 1981, paper 81-1259.
- [17] van Leer B., Flux vector splitting for the 1990s, *Proceedings of Computational Fluid Dynamics Symposium on Aero-propulsion*, NASA, 1991, 203-214.
- [18] Kalita P. and Dass A.K. (2016) A diffusion-regulated scheme for the compressible Navier-Stokes equations using a boundary layer sensor, *Computers & Fluids* **129** 91-100.
- [19] Kalita P. and Dass A.K., Effect of numerical diffusion on the computation of hypersonic shock wave boundary layer interaction, *Proceedings of the International Symposium on Aspects of Mechanical Engineering & Technology for Industry*, North Eastern Regional Institute of Science and Technology India, 2014, 306-313.
- [20] Godunov S.K. (1959) A difference scheme for numerical computation of discontinuous solutions of the equations of hydrodynamics, *Matematicheskii Sbornik*, **47** 271-306.
- [21] van Leer B. (1979) Towards the ultimate conservative difference scheme V. A Second Order Sequel to Godunov Method, *Journal of Computational Physics*, **32** 101-136.
- [22] Laney C.B. (1998) *Computational Gasdynamics*, Cambridge University Press.
- [23] Jaisankar S. and Raghurama Rao S.V. (2007) Diffusion regulation for Euler solvers, *Journal of Computational Physics*, **221** 577-599.
- [24] van Albada G.D., van Leer B. and Roberts W. W. (1982) A comparative study of computational methods in cosmic gas dynamics, *Astronomy and Astrophysics*, **108** 76-84.
- [25] Blazek J. (2001) *Computational Fluid Dynamics; Principles and Applications*, 1st ed., Elsevier.
- [26] Swanson R.C., Radespiel R. and Turkel E. (1998) On some numerical dissipation schemes, *Journal of Computational Physics*, **147** 518-544.

- [27] Tu J., Yeoh G.H., and Liu C. (2007) *Computational Fluid Dynamics: A Practical Approach*, Butterworth-Heinemann.
- [28] Kalita P. and Dass A. K. (2017) A novel hybrid approach with multidimensional-like effects for compressible flow computations, *Journal of Computational Physics*, **340** 55-68.
- [29] Marini M., Effects of flow and geometry parameters on shock wave-boundary layer interaction in hypersonic flow, *AIAA paper*, 1998, paper 98-1570.
- [30] Van Driest E.R. Investigation of laminar boundary layer in compressible fluids using the Crocco method, *NACA Technical Note*, 1952, paper 2597.

The effects of solution techniques on the results of the simulation of human motion

†*Behzat B Kentel¹

¹Mechanical Engineering Program, Middle East Technical University Northern Cyprus Campus, Turkey.

*Presenting author: kbehzat@metu.edu.tr

†Corresponding author: kbehzat@metu.edu.tr

Abstract

Computer simulations of human movements are used for understanding the dynamics of the motion. These simulation models using individual muscles or torque generators can be separated into two parts as the dynamics of the body segments and the muscular mechanics part. The solution of the governing equations for these parts is necessary for the simulation. In this study, two different solution techniques will be compared and discussed. The first technique includes the solution of each part separately whereas all equations solved simultaneously in the second technique. In the first technique, the solutions can be obtained with relatively less computational effort but in the second technique more accurate results are expected. The comparison of the results will show whether the improvement in the accuracy worth the increase in the computational effort. As a result, a particular solution technique can be proposed to all current and forthcoming studies.

Keywords: Computation, biomechanics, motion analysis, simulation modelling

Introduction

Computer simulations of human movements are used for understanding the dynamics of the motion. In general, a detailed information is obtained from the simulations at a level where no direct measurement is possible. In the literature, there are many examples of simulation models with different complexity level using different methods. The most common methods for the modelling of the human motion include individual muscles or torque generators as actuators of the model. These elements represent the effect of force/torque generated at the muscle fibres. For example, Anderson and Pandy [1] developed an individual muscle model to simulate walking whereas Kentel et al. [2] developed a model with torque generators for simulating backhand ground strokes in tennis.

The models that use individual muscles or torque generators can be separated into two parts as the dynamics of the body segments and the muscular mechanics part causing the force or torque generation. These parts are directly related to each other and have to be considered together during the simulation. Each part has their own modelling structure but have common variables that affect both parts. The dynamics of the body segments simply deals with the motion of the body using the equation of motion. On the other hand, muscular mechanics part deals with the amount of force/torque generated due to muscle fibres.

The simulation of the human motion depends on the solution of the differential equations relating the variables in each part. The difficulty arises on how to solve these equations since they have common variables. No explicit solution of these equations were present in the literature but two possible ways of solutions may be used. As a first technique, the parts solved

separately at each integration step using the data transferred from the other part. In the second technique, all equations are solved simultaneously.

This study focusses on the effects of these two ways of solution techniques on the results. Although the second technique promises more accurate results, the computational effort would be higher. By analysing a sample human motion such as dumbbell curl, this study investigates whether the improvement in the accuracy worth additional computational effort. The results of both techniques are compared and the difficulties as well as advantages of both techniques are discussed

Model Development

This study focusses on the solution technique rather than the motion itself. Therefore, a simple and realistic motion could be selected for the computer simulation. For this purpose, a planar 5 kg dumbbell curl is used. The upper arm is assumed to be fixed and there is no relative motion at the wrist joint. In short, only one segment (forearm and hand) is moving about the joint (elbow) on the sagittal plane. The free body diagram of the forearm with the weight can be seen in Fig. 1. The equation of motion using the free body diagram is given in Eq. (1).

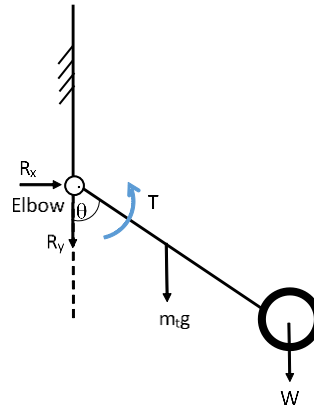


Figure 1. Free body diagram of the forearm with the weight

$$T - I_E \ddot{\theta} - m_t g l_1 \sin\theta - W l_2 \sin\theta \quad (1)$$

As the actuator of the model, a single torque generator located at the elbow joint is used. The anthropometric parameters of the forearm and the torque-strength parameters of the muscles are obtained from the literature [3]. The torque generator represents the moment effects of all muscle fibres going through the elbow joint. It has two separate units each corresponding the flexor and extensor muscle groups for agonistic and antagonistic action. The algebraic sum of the flexor and extensor units is the net torque on the elbow joint. Torque generators can be considered as a rotational form of Hill's muscle model (Fig. 2).

In this rotational form, muscle fibre length, tendon length and musculotendon length is represented by muscle angle (θ_m), tendon angle (θ_e) and joint angle (θ_j), respectively. The relation between these variables are expressed in Eq. (2) (Fig. 2b). The joint angle is the relative angle between the upper arm and the forearm at any instant.

$$\theta_m + \theta_e = \theta_j \quad (2)$$

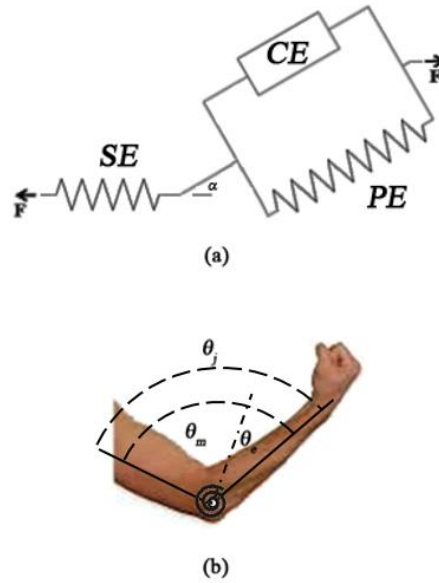


Figure 2. (a) Hill-type muscle model; (b) torque generator

The contractile element (*CE*) of the muscle model determines the relation between the torque, the muscle angle and the muscle angular velocity (contraction velocity). This torque value (T_{con}) is the maximum voluntary torque and has to be multiplied with the muscle activation, $a(t)$. In addition, differential activation, $d(\dot{\theta}_m)$, may be used as a correction considering the depression in muscle extension (Eq. 3). The torque can also be calculated through the series elastic element (*SE*) using the torsional spring constant, k_t , which represents the tendon stiffness.

$$T_{con}(\theta_m, \dot{\theta}_m) * a(t) * d(\dot{\theta}_m) = T = k_t \theta_e \quad (3)$$

The muscle activation is a function of time and ranging from 0 (no activation) to 1 (full activation). Muscle activation profiles describe how the activation changes during the period of simulation. Parametric curves are used for muscle activation profiles and the parameters are determined via an optimisation process to match the simulation result with the actual motion. A sample activation profile is presented in Fig. 3.

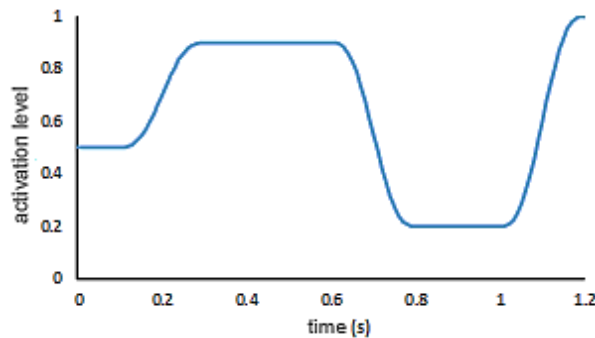


Figure 3. A sample activation profile generated with 9 parameters

The simulation model can be separated into two parts as the dynamics of the body segments and the muscular mechanics part. Eq. (1) describes the dynamics of the body segments whereas Eq. (2) and (3) describe the muscular mechanics part. The torque and joint angle terms appear in both parts and relate the two parts. The estimated torque values are input to the body segments

part and the output is the motion of the forearm i.e. change in the muscular geometry. On the other hand, muscular geometry is used to determine the amount of torque generated at the elbow.

Eq. (1) and (2) are two differential equations and can be solved by applying numerical methods. In this study, 4th order Runge-Kutta method is used to solve the differential equations in both muscular part and the dynamics of the body segments. Two different techniques are used for the solution of the equations.

In the first technique, two parts solved separately at each integration step. The output from the solution of one part in a step is used as an input to the other part. This can be summarized in Fig. 4. The same procedure continues at each integration step until the end of the simulation

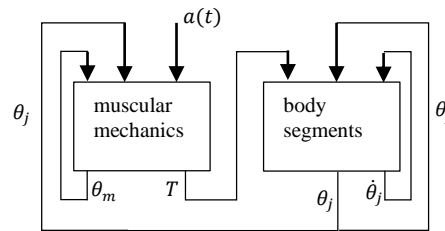


Figure 4. A schematic representation of the first solution technique

In the second technique, all equations are considered simultaneously at each integration step. Therefore, there is no distinct parts in the model. A combinatory Runge-Kutta method is applied to both differential equations.

Currently, the model is developed and two solutions method are being applied. Numerical results will be presented once the results from the both methods are obtained.

Conclusions

Two different solution techniques of the equations that govern the human motion will be compared and discussed. The first technique uses separate solutions of each part of the model at each integration step. This brings modularity to the model and therefore, the solutions can be obtained with relatively less computational effort. However, the values of the variables in one part appears to be unchanged with respect to other at each integration step and the change of the variables in one part can affect the other part only in the next step.

The second technique uses all equations simultaneously at each integration step. Since a combinatory Runge-Kutta is applied, the computational effort is higher than the first technique. The results are expected to be more accurate. The changing of all variables taken into account at the same time that result in less error during the numerical solution. After having the comparison of both methods a particular solution technique can be proposed to all current and forthcoming studies.

References

- [1] Anderson, F. C., and Pandy, M. G. (2001) Dynamic optimization of human walking, *Journal of Biomechanical Engineering*, **123**(5), 381-390.
- [2] Kentel, B. B., King, M. A., and Mitchell, S. R., (2011) Evaluation of a Subject-Specific, Torque-Driven Computer Simulation Model of One-handed Tennis Backhand Ground Strokes, *Journal of Applied Biomechanics*, **27**, 345-354.
- [3] King, M. A., and Yeadon, M. R., (2002) Determining Subject-Specific Torque Parameters for Use in a Torque-Driven Simulation Model of Dynamic Jumping, *Journal of Applied Biomechanics*, **18**(3), 207-217.

Seismic risk of masonry buildings: methods compared for ICCM

*†Mariateresa Guadagnuolo¹, Marianna Aurilio¹, and Giuseppe Faella¹

¹Department of Architecture and Industrial Design, Università degli Studi della Campania Luigi Vanvitelli,
S.Lorenzo ad Septimum Abbey, Aversa (CE) Italy.

*Presenting author, †Corresponding author: m.guadagnuolo@unicampania.it

Abstract

The Italian Guidelines for Seismic Risk Classification of Buildings, issued by D.M. 28.02.2017, point to a vulnerability assessment to be understood primarily in terms of risk mitigation and in terms of optimization of the interventions. Considering the current serious condition of the Italian built heritage, which has been devastated by earthquakes because of its high vulnerability, the main goal is to provide a seismic classification methodology of existing buildings, before and after any interventions aimed to improving the vulnerability class. For masonry buildings, the Guidelines also provide a simplified approach based on a classification of buildings depending on their wall masonry type, their structural peculiarities and site hazard, identified through the zoning defined by the OPCM 3274/2003. The method allows to evaluate both the seismic vulnerability and the effectiveness of the interventions to be implemented to mitigate the risk. The purpose of the paper is a constructive critical analysis of the application of the above simplified method applied to buildings located in the province of Caserta, hypothesizing some interventions of improvement. The possible interventions of mitigation included in the procedures, are directed to strengthen in local way the building, to increase its safety index and to decrease the class of risk, they are the product of a study aimed to individualize its real necessity both in relationship to the type of intervention that to the part of the structure that is in the optics of the optimization of the same interested by it. The analyses were conducted with reference to two types of buildings particularly recurrent and representative of the built heritage of the province of Caserta and located into areas with different seismic hazards.

Keywords: Classification, Seismic Risk, Masonry, Interventions, Province of Caserta.

Introduction

The building patrimony of our Country is represented by constructions in masonry of which a wide part is situated in the historical centres. Such constructions deserve particular attention as bearers of inestimable values due to their existence in the time that makes her a rich historical, artistic and cultural palimpsest, and in how much the totality of them has not almost been realized with criterions seismic. Besides the existing buildings are characterized by problems concerning phenomena of degrade and longevity [1].

The actual strategy of seismic prevention is based on an unitary approach that foresees the seismic classification of the territory, the seismic planning of the new constructions and its projected toward the adjustment or the improvement [2], [3], [4]. The Directive the Ministers' President of the Council [5] (Directive 2011) furnishes indications for the evaluation and the reduction of the seismic risk for the protected cultural patrimony; compiled with the intent to specify a run of knowledge, evaluation of the safety level towards the seismic actions and project of the possible interventions conceptually analogous to that anticipated for the

constructions not protected, but opportunely suitable to the demands and peculiarity of the cultural patrimony [6].

Due to the recent code developments and of the growing attention given to the seismic safety of structures, especially after the last Italian earthquakes, the analysis and verification of existing building heritage have become a fundamental tool to assess the seismic vulnerability, to safeguard human lives and to plan structural interventions.

The Italian building heritage is characterized by high complexity and heterogeneity, both from architectural and structural points of view. For all these reasons, it is important to define a methodology to obtain comparable results to plan the future activities of risk analysis, assessment and management.

A significant number of older stone and masonry buildings are not in accordance with any of actual provisions code. The seismic risk assessment and seismic vulnerability assessment of existing building stock is essential for establishing priorities in a long-term prevention policy. Vulnerability index method uses collected information of parameters of the building (plan, height, structural and non-structural elements, type and quality of materials). This method is used as one of several general methodologies for vulnerability assessment and seismic risk assessment.

Due to the scale and number of buildings involved, the methodologies currently available to assess the seismic vulnerability of urban areas usually require the treatment of a massive volume of data associated with the inspection and survey work, and for this reason the use of more simplified approaches is becoming more popular.

In [7] is to identify a methodology of verification easily manageable and adaptable to many different buildings, but at the same time able to determine the actual state of structure in terms of critical steps and structural deficiencies.

Actual condition obviously pushes toward an evaluation expeditious type, based on the logic to adjust how much patrimony possible but the choice of the typology of interventions adopted following the evaluations it doesn't result to always have been both decisive in comparison to a fragile and unstable starting condition that economic in comparison to a condition of urgency [8], many interventions adopted following the seismic phenomena are partially revealed ineffective to withstand the intense seismic actions [9].

In the present paper is effected a seismic analysis applying the method LV1 for the calculation simplified of the safety index before and after the interventions. Particularly it has been made an evaluation of the typology and the quantity of local interventions to adopt. The analyses were conducted with reference to two types of buildings particularly recurrent and representative of the built heritage of the province of Caserta and located into areas with different seismic hazards.

The aim of this paper is to provide the first steps in assessing seismic risk in Campania, which has been achieved through an investigation of the building typology by site investigation and existing plans and documentation.

Cases Study

The present study analyses two residential buildings, representative of the large majority of the existing constructions in the historical center of the province of Caserta, in which it is possible to individualize two periods of construction, the XV century and the period between 1800 to the beginnings of 1900. The greatest part of the buildings is made by simple or massive stones and develops around a court or a central courtyard, generally raising for at least two or three floors. The constructions are whether isolated or inserted in united and they usually have gable roof not pushing and timber and metallic planking's.

Palazzo Petrucci-Novelli

The first building examined in Fig. 1 and Fig. 2, is a construction situated in Carinola as showed in Fig. 3, in the province of Caserta and it constitutes a typical example of the constructive typology of the area made by simple stone in regular blocks and it's located in a town which has an average-low seismic dangerousness. The thickness of the construction, realized around the XIV century, is 60 cm and it develops on two levels around a central courtyard where is an external staircase which colleague the ground floor to the loggia as showed in Fig. 4 and in Fig. 5. The decks have crossvaults, barrel vaults, plan floor and a roof made by wooden trusses. It has a compact form and an irregular morphology in both plant and elevated.



Figure 1. Building's exterior



Figure 2. Building's exterior



Figure 3. Layout of the historical center

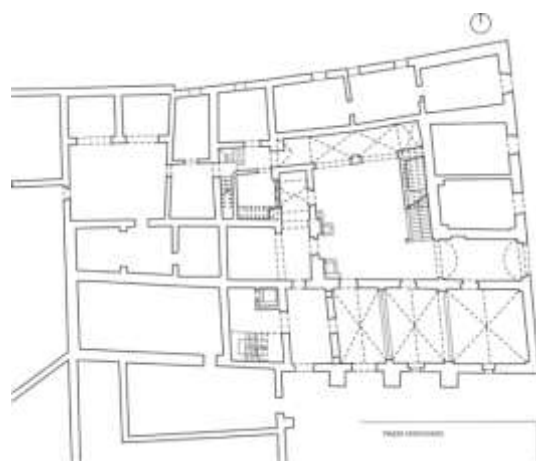


Figure 4. Floor Plan



Figure 5a. Elevations

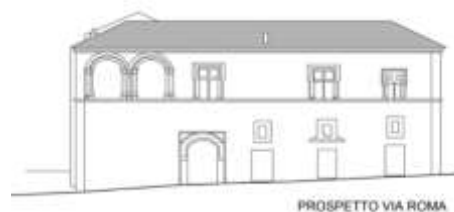


Figure 5b. Elevations

Palazzo Ducale

This building showed in Fig. 6 and Fig. 7, situated in the commune of Piedimonte Matese Fig. 8, in the province of Caserta, it's a prototype of the constructive typology of the Middle Volturno's areas, next to the Appennino therefore a district to high seismic dangerousness. The palace, built in the XVI century, is made by a massive stone mixed to fieldstone and blocks of bricks with a thickness around 80 cms. The building has four levels with attics in wood articulated around a central court and it has a compact form and an irregular morphology in both plant and elevated as showed in Fig. 9 and Fig. 10.



Figure 6. Building's exterior



Figure 7. Building's courtyard



Figure 8. Layout of the historical center



Figure 9. Floor Plan



Figure 10a. Elevations



Figure 10b. Elevations

The Guidelines for the Reduction of Seismic Risk on Cultural Heritage

The Guidelines for the Reduction of Seismic Risk on Cultural Heritage (code DPCM 2011), linked to the NTCs 2008, has been compiled in order to specify a path of knowledge, evaluation and reduction of the seismic risk for masonry buildings. To evaluate the seismic safety, three different levels of increasing completeness have been identified of which the level LV1 concerns the evaluations of the seismic safety at a territorial scale. It is necessary to quantitatively evaluate the ground acceleration leading to the achievement of the structure ultimate limit state (SLV), with a pre-set probability of overcoming, and that attended in the site on a reference's period defined on the building's characteristics and its use. The seismic safety index is estimated by the relation between the return period of the seismic action provoking the generic limit state and the corresponding return period of reference related to the earthquake expected on the site. This is useful to underline the critical situations and to establish a priority for the future interventions. In the same way it's possible to define an acceleration factor defined by the relation between the acceleration which provokes the ultimate limit state and the acceleration expected on the site.

Palazzo Petrucci Novelli_The LV1 method

According to the procedure defined in the Directive 2011 for the method LV1-Palace and (X), the construction has nominal life equal to 50 year-old (buildings with ordinary performance levels), use class II, coefficient of equal use 1,00 (construction that foresees normal overcrowdings without dangerous elements for the environment), reference period for the seismic action equal to 50, category of soil class A specific for "very rigid appearing on the surface or terrestrial rocky heaps", coefficient of subsoil equal to 1,00 and topographical category T1 for "level surfaces, slant and isolated reliefs with middle inclination $i < 15^\circ$ ". Using a NTCs software, inserting the coordinates of the structure (Lat. 41.188625° , Long. 13.976842°) and the values mentioned, the ground peak acceleration results a_g / g : 0,098g, the F_0 factor equal to 2,675. The construction, made by simple stone has a proper weight equal to 16,0 KN/m³s, the permanent load G_k changes by the levels. For example, the first floor has both metal decking and desks with vaults, the desks of the second floor are in metallic profiles and timber, and the third floor introduces a timber's roof. Analyzing them, three values of G_k are gotten, respectively 12 KN/m²s, 6 KN/m²s and 1,7 KN/m²s. The confidence factor assumed is $FC=1,35$ (corresponding to complete survey of the building geometry and to limited knowledge of the mechanical properties of materials). Due to the characteristics and peculiarity of the structure, the shearing force is equal to 0,028 MPas for each level and the failure index is equal to 0,8 for the piers and the strength of the spandrel beams for all the levels in both the directions. The minor acceleration factor (0,45) isn't sufficient to make forehead to a seismic event.

Palazzo Ducale_ The LV1 method

Using the same procedures above mentioned, the second construction has nominal life equal to 50 year-olds, use class II, coefficient of equal use 1,00, reference period for the seismic action equal to 50, category of soil class A, coefficient of subsoil equal to 1,00 and topographical category T1. Inserting the geographical coordinates in the software (Lat. $41.41.365277^\circ$, Long. 14.383055°) and the other values above-mentioned, the design ground acceleration is a_g / g : 0,249g, the F_0 factor equal to 2,304. In contrast to the previously construction analyzed, this building has a mixed masonry composition in simple stone and fieldstone, so its proper weight is 19,0 KN/m³s, and has a permanent load G_k equal to 5,4 kN/m²s unchanged for all the floor excluding the roof for which it's 0,5 kN/m²s. The level of knowledge assumed for the construction is $FC=1,35$ and thanks to its characteristics and

peculiarity the shearing force is equal to 0,028 MPas for each level and the failure index is equal to 0,8 for the piers and the strength of the spandrel beams for all the levels in both the directions. The smaller acceleration factor resulted by the analysis is equal to 0,25 indicate that the structures is unable to withstand the required seismic forces, provided by the seismic code.

Interventions

The procedure for the evaluation of the seismic safety also includes the potential interventions of mitigations, aims to strengthen in local way the building and to increase its seismic safety index; the two analyzed constructions, have a seismic safety index inferior to 1, so they are both unable to make forehead to a seismic event. In preliminary analysis the structure is verified before the intervention with identification of the lacks and the level of seismic action for which the SLU is reached. The choice of the improvement intervention must be a motivated strategy aimed to interest a select portion for which improve the structural performance. Subsequently the technical choices and the materials are verified in order to apply them with the preliminary sizing of the reinforcements and the additional structural elements. Post intervention, a structural analysis will be effected. The evaluation of the safety and the design of intervention must be wide to all the parts of the structure potentially interested by changes of behavior, as well as to the structure in its whole. Starting from the condition of the single construction, with the purpose to increase the safety index of the structure, a first operation of analysis consists of combining interventions through which increases the shearing force from the first floor to the top in order to produce an increase safety index. Consequently, there are planned interventions which engrave on the collapse method through a further increase of the coefficients of collapse and strength gradually increasing until reaching a safety index greater than beginning.

Palazzo Petrucci Novelli_Interventions

The Figure 11 shows the increase's curve of the safety index, obtained through a combination of gradual increases of the shearing force (I_{fy}) and the coefficients of collapse and strength of piers and the spandrel beams ($C_c - C_r$) from the initial safety index 0.45, thin to the ultimate 0.73. The points in black, represents the three values of the safety index reached increasing at the most the shearing force (I_{fy}) and the coefficients of collapse and strength of piers and the spandrel beams ($C_c - C_r$) at the first to the top floor and then adding the maximum increases to the inferior floors. The gradual increase's curve of the seismic safety index and the points of maximum increase, differ between them for the different design choice, in the specific one the points of maximum increase hypothesize diffused interventions on the whole construction strengthening at the most the coefficients of collapse and resistance of the single piers and the spandrel beams ($C_c - C_r$) and the shearing force (I_{fy}) to every floor without consider that, the result of the LV1 analysis and the consequent low safety index, could be due a specific lack located on the construction rather than to a general lack, contrarily the curve minimizes the interventions and the costs going to gradually increase the single coefficients and the single shearing force (I_{fy}) (whereas is necessary) in the optics to reach a result of improvement but optimizing choices and costs. Following the hypothesis of the targeted and optimized interventions, the safety maximum index is obtained increasing the shearing force (I_{fy}) by 50% to the third floor; 50% to the second floor and 20% to the first floor. The coefficients of collapse of piers and the strength of spandrel beams ($C_c - C_r$) (assumed equal in directions x and y) have been increased thin to 1 for the third and the second floor while they have been being unchanged at the first floor. In the second case globally maximizing the values, the curve quickly grows and steeply up to get a safety index equal to 0.73. In this case both the

shearing force (I_{fy}) (increased by 50% to each floor) and the coefficients of collapse of the piers and strength of spandrel beams ($C_c - C_r$) (passed by 0,8 to 1) have reached the maximum increase. The increase of the seismic safety index, is verified increasing by 20% the shearing force (I_{fy}) to the first floor, considered the weakest floor, rather than increasing of 50% shearing force (I_{fy}) to the other floors.

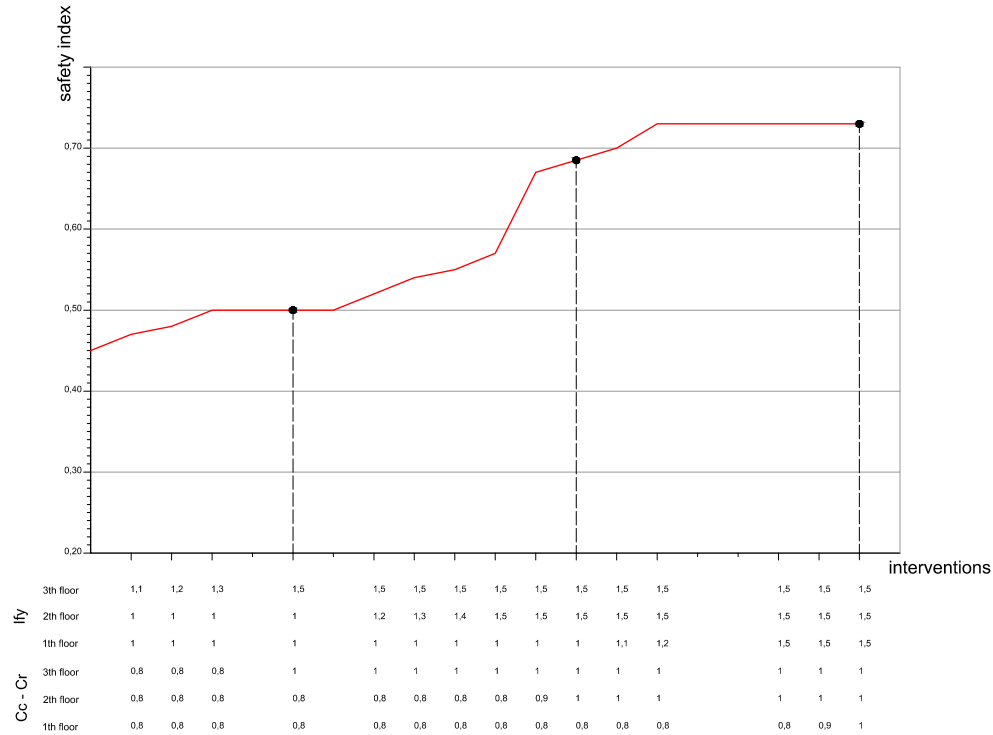


Figure 11. Safety Index curve of Palazzo Petrucci Novelli

Such analysis underlines as despite both design choices are succeed in optimizing the strength of the building increasing its safety index from 0.45 to 0.73, in the first case the interventions would be inferior, less expensive and really located on the parts interesting from structural lacks while in the second case it would intervene over the minimum, globally on the whole building and in very more elevated economic terms.

Palazzo Ducale_Interventions

In this case as underlines in Fig. 12, the LV1 method, returns an initial safety index equal to 0.25 that being smaller than 1 expresses a meaningful insufficiency to make forehead to a seismic event, condition mostly criticism in comparison to the first case study. Also for this second building has been effected a double analysis on two different design choices, the first one with a gradual increase of the factors while the second through the pursuit of the maximum result gotten contemporarily maximizing the increases. Otherwise from the previously case analyzed, for this typology of construction, the real raising of the index safety and relative structural improvement, is gotten for both the design choices, increasing at the most the shearing force (I_{fy}) to every floor in both the directions that the coefficients of collapse of the piers and the strength of the spandrel beams ($C_c - C_r$). In the specific case, the curve, that represents the gradual increase that conducts to a passage of the safety index from 0.25 (initial scenery) to 0.52 are obtained increasing by 50% the shearing force (I_{fy}) to the fourth floor; by the 50% to the third floor; by the 50% to the second floor and finally by the 50% to the first floor. The coefficients of collapse of piers and the strength of spandrel beams

($C_c - C_r$) (assumed equal in directions x and y) have been increased thin to 1 for all the levels.

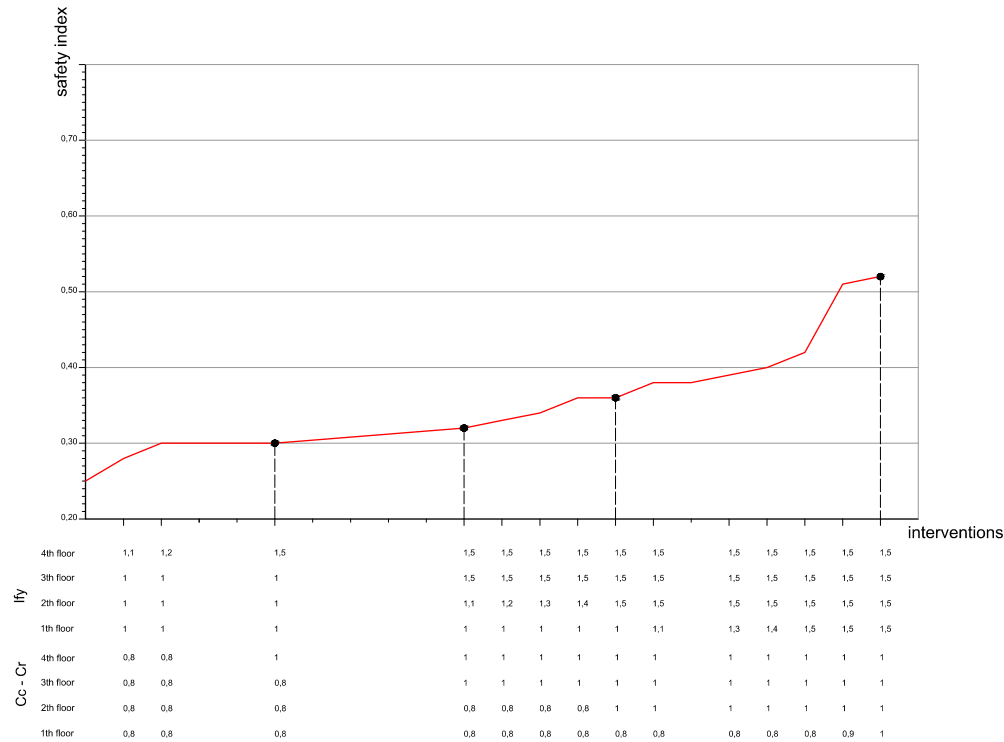


Figure 12. Safety Index of Palazzo Ducale

The points in black instead, representative of the second typology of design choices, show as increasing all the values from the fourth floor up to the first one, it's obtained a seismic safety index equal to that gotten from the first scenery of intervention, passing from 0,25 to 0,52. For the analysis of this construction, otherwise from the first one, the attainment of the increase of the safety index is contextually achieved for both the sceneries maximizing the increases of the resistances, delineating therefore as the real improvement of the strength ability of the building object of analysis is bound to a design choice of intervention globally put in work on the whole construction.

Conclusions

This paper deepens some aspects linked both to the estimation of the vulnerability of the masonry buildings and both to the choice of the interventions to adopt, with specific reference to those reverting in the area of Caserta.

Particular attention has been set on the simplified methods of evaluation which requesting a less deepened knowledge on the structure, they conduct to results that can be more reliable, but being of faster application, they are particularly suitable for the analyses of consistent champions, as in the case of the historical centers. The simplified method LV1 has been applied, on two representative buildings of an ample quantity of residential buildings of Caserta: a building of the historical center of Carinola, area with a low-average seismic dangerousness, representative of the regular buildings realized in simple stone and a second construction built in the historical center of Piedimonte Matese, area with an elevated seismic dangerousness, strongly irregular in its configuration and realized with mixed material of various typology. Analogous results have been caught by numerous constructions similar to

the examine ones, confirming the extendibility of the conclusions that are derived to an ample portion of the built patrimony of the northern area of Caserta.

Seismic vulnerability has been valued using a quantitative type of procedure: the method LV1 for the simplified calculation of the safety index introduced by the Directive of the 2011. For the first building (*Palazzo Petrucci-Novelli*), representative of the area of Carinola, the result caught through the LV1 method, based on the shearing force of the building walls, conduct to a middle vulnerability. The results achieved from the second Building (*Palazzo Ducale*) underlining a high vulnerability.

The LV1 method, operate in order to appraise the SLV seismic action of a building, choosing subsequently to effect interventions of reparation and developing a LV2 analysis which is necessary to confirm the real necessity of the interventions, considering the maximum acceleration to the floor of reference in the site; in the elements in which the SLV acceleration is already superior to this last, wouldn't be necessary to proceed would to the seismic improvement of that part.

It is clear that a simplified type of approach, needs reduced time but able, at the same time, to optimize the interventions. LV1 allows us to get a series of coherent information for the attainment of the objective and addresses toward the calculations that will allow to increase the safety index, studying the global/geometric state of the construction to be able to found a project of intervention. Individualized the formality of collapse, through LV1 that allows us to individualize the weak portions, on which it is necessary to intervene of it, it's simple to use the LV2 method, to extend to these parts of the construction to be able to develop the project of intervention and to choose its typology. The analyzed cases are representative of many others sceneries verified in the time in which, as in the case of *Palazzo Ducale*, in the optics to improve much possible patrimony, has been select to realize a series of interventions on the whole construction at the expense of costs and times but as they had showed, that determined benefits would be produced in every cases opting for a number of interventions done on the whole construction, otherwise in the case of *Palazzo Petrucci-Novelli*, the strength increase of the safety index and the strength ability of the structure would be achieved choosing to adopt interventions focused to located and weak parts of the structure that would have produced the same effects and the same benefits of interventions realized on the whole structure. Proceeding, for simplified models, could be neglected different peculiar characters but surely a whole series of fundamental parameters wouldn't be neglected neither in phase of analysis nor choosing interventions.

References

- [1] Masi, A., Santarsiero, G., Ventura G. (2017) Strategie per la riduzione del rischio sismico applicate agli edifici scolastici: un caso studio., Anidis 2017 Pistoia. 2232 – 2240.
- [2] Giovinazzi, S., Lagomarsino, S. (2004). A macroseismic models for the vulnerability assessment of the buildings. 13th World Conference on Earthquake. August 1.6.2004. Paper no 896.
- [3] Zuccaro, G., Cacace, F. (2009). Modello per la simulazione di scenari sismici per la Regione Campania. . L'Ingegneria Sismica in Italia . Bologna, Italy.
- [4] Dolce, M., Di Pasquale, G., Speranza E. (2012). A multipurpose method for seismic vulnerability assessment of urban areas. 15 WCEE LISBOA 2012.
- [5] MIT (2011) Direttiva del Presidente del Consiglio dei Ministri 9 Febbraio 2011; riferimento NTC D.M. 14/01/08. Valutazione e riduzione del rischio sismico del patrimonio culturale.
- [6] Guadagnuolo, M., Aurilio, M., Faella, G. (2017) Rischio sismico di edifici in muratura: metodi a confronto., Anidis 2017 Pistoia. 2171 – 2180.
- [7] G. Cimino, I. Ricci, G. Gasparini, T. Trombetti (2017) Seismic vulnerability of building heritage of the University of Bologna: methodology and analysis 16th World Conference on Earthquake, 16WCEE 2017 Santiago Chile, January 9th to 13th 2017.
- [8] Caterino, N., Cosenza E. (2017) Evaluation of seismic retrofit techniques via a multicriteria approach accounting for Italian tax incentives., Anidis 2017 Pistoia. 2181 – 2190.
- [9] Petrucci, E. (2017) Considerazioni sulle procedure attuate dopo il sism del 1997 nella Regione Marche: nuovi contributi al consolidamento delle opere murarie., Anidis 2017 Pistoia. 2386 – 2396.

An innovative Pedestrian-Bicycle Bridge Shape for Environmental Sustainability and Structural Efficient Improvement

***Lucci Melania¹, †Viskovic Alberto¹**

¹Department of Engineering and Geology, University "G. D'Annunzio" of Chieti-Pescara, Italy.

*Presenting author: melanialucci@hotmail.it

†Corresponding author: alberto.viskovic@unich.it

Abstract

Planning pedestrian and bicycle bridges to overpass rivers, highways or roads with intense traffic, represents a smart way to enhance a territory and to reduce the massive and improper use of public and private transports with high environmental impact. In this research work we would like to show how, thinking different, could be possible to improve the concept of mobility in a sustainable way. Proper design choices on shape, with high structural efficiency, and material, with low footprint on nature, are fundamental for this process. Thanks to its great shape stiffness, the hyperbolic paraboloid surface has proven to be good choice. It represents a ruled surface capable of realizing complex curvatures simply derived from straight lines crossed between themselves. About the material, we have chosen the Cross Laminated Timber (CLT) as an interesting alternative to the traditionally used materials for this kind of structure. The very good strength/weight ratio, the orthotropic behaviour and the flexibility in designing and forming structural elements of different sizes and shapes are some of the main features of the CLT. The geometric modelling of the bridge is created in Rhinoceros using a parametric approach. The algorithm, developed through a visual programming language Grasshopper, is written in such way that the size of the geometric element can be rapidly and parametrically modify as necessary. The structural analysis is based on a Finite Element Method (FEM) through MIDAS/Gen. A software chosen for its flexibility, the wide material and finite element library and the possibility of modelling the orthotropic behaviour. In particular, we have chosen "plate" as finite element with proper material characteristics according to an implemented well oriented mesh. In conclusion, the aim of this research work is to design a cycle and pedestrian bridge in which choices on material and shape move towards a sustainable mobility. The possibility of in-site rapid assembly process should not be neglected.

Keywords: Sustainability, Efficiency, Wood, Bridge, Paraboloid, Parametric, FEM

Introduction

In recent years, the urban development of Western countries has produced a clear separation between the art of building the city and the attention to places. The starting point of the following research work is the awareness that engineering, architecture and urban planning must return to assume a central role within the city, looking at the sustainable development of the urban landscape to "satisfy two human instances: [...] the physical and the spiritual one [...]" [1].

The design of cycle walkways is one of the concrete ways to follow this philosophy. It offers the possibility of experimenting with innovative forms and materials to create something that is not just a simple means of connection between two places

Modelling and Structural Analysis

The installation of a walkway that allow passing over rivers or busy roads is an intelligent way of enhancing a territory and reducing the massive use of public and private polluting vehicles.

At the base of the following research work there is the will to enhance, in an alternative way, the concept of slow and sustainable mobility. The decision to design a "structural architecture" was fundamental to this process. It is intended as an aptitude for combining forms and structures by attributing aesthetic value to the "backbone" of an architectural project. Therefore, the search for design solutions different from the classic types used for the design of driveway bridges. The adopted choices regarding the form, the material, the structural, constructive and economic feasibility, have the aim to obtain a result with a low environmental impact and high structural efficiency.

Thanks to its great stiffness in shape, with a membrane behaviour and an optimal thickness distribution, the meeting point between structural efficiency and sustainability was obtained through the choice of the hyperbolic paraboloid. A quadratic striped surface for the realization of only apparently complex curvatures, obtainable thanks to simple translation of straight lines into the space, through a simple and flexible construction process.

The structure has been designed with Cross-Laminated Timber (CLT) wooden plates making the material work according to its natural attitude of orthotropic and combining its technical performances with the energetic and ecological ones. The excellent strength-weight ratio, the properties of orthotropic, the excellent resistance and reaction to fire and the great flexibility in the design of structural elements of different sizes and shapes are some of the important peculiarities of CLT.

A key factor for a successful approach in the design of this boardwalk was the integration between architectural-parametric and structural modelling.

In particular, in the geometric modelling, a parametric approach was chosen using the Grasshopper software. A modelling therefore that does not follow the logic of CAD or 3D, but algorithmic. It is generated from a list of instructions that process inputs through a logical sequence of descriptive geometry operations. Only with the algorithm terminated, the surface has been edited with the help of the Rhinoceros Software.

From a square-based hyperbolic paraboloid [2] a hyperbolic paraboloid of rhomboidal base has been made. The surface has been cut out of the dimensions and features compatible with those of the boardwalk in question, implementing equations and parametric, as shown in Fig.1. Therefore, since the solution is not unambiguous in this way, design constraints have been inserted regarding the context, the structural and architectural properties and the normative apparatus in force. The obtained solution represents the fusion between geometrical features (derived from the most disadvantageous load conditions) and the optimization of results concerning the structural efficiency and the aesthetic dimension.

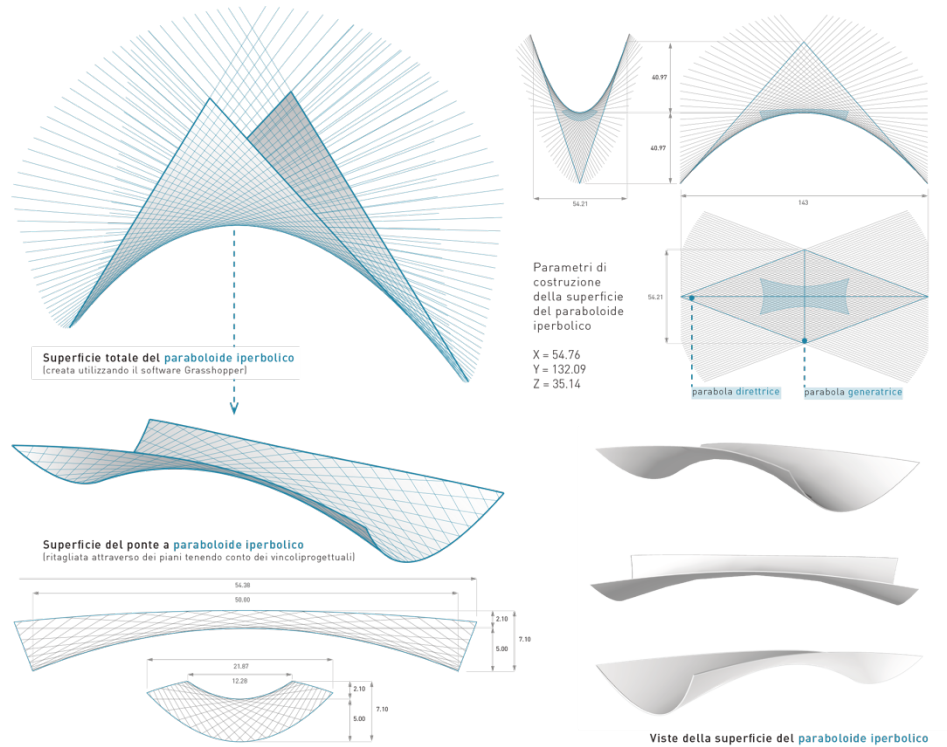


Figure 1. Geometrical and parametric modelling of the boardwalk surface [3]

The structural analysis, based on the Finite Element Method (FEM) has been realized through the MIDAS-Gen software. This choice derives from its flexibility, the possibility of having a very large library available to the user, both as regards the materials (customizable if necessary) and the finite elements, from the implementation of the orthotropic properties of the chosen material. from the ability to import and export files in many formats (it is perfectly oriented in the world of Building Information Modelling - BIM [4]).

The modelling was carried out following a computational method. In particular, established form, size, type of material and model loads, we have moved to its resolution. By determining deformation, strain and displacement, the structural verification of the element has been achieved.

The implementation of a simplified structural scheme with respect to the architectural model was fundamental. The inclusion of simplifying hypotheses has allowed to reduce the computational cost from the software point of view. However, the discretization of the model in 2D finite elements, called "plate", has been strongly influenced by the orthotropy of the material. The orthotropic behaviour was implemented by inserting the mechanical and physical properties of the material according to the orientations of the local (non-global) axes of the individual plate elements. Considering this choices, the creation of a manual mesh was necessary. Elements were created as close as possible near to be rectangles or squares, oriented in such a way that the values assigned to the mechanical properties in the material definition were consistent with the designed structure. It was specifically obtained by generating the transversal and longitudinal parabolas whose joint nodes are placed on the lines of connection between the various surfaces (surface, bridge deck and diaphragm in Fig.2).

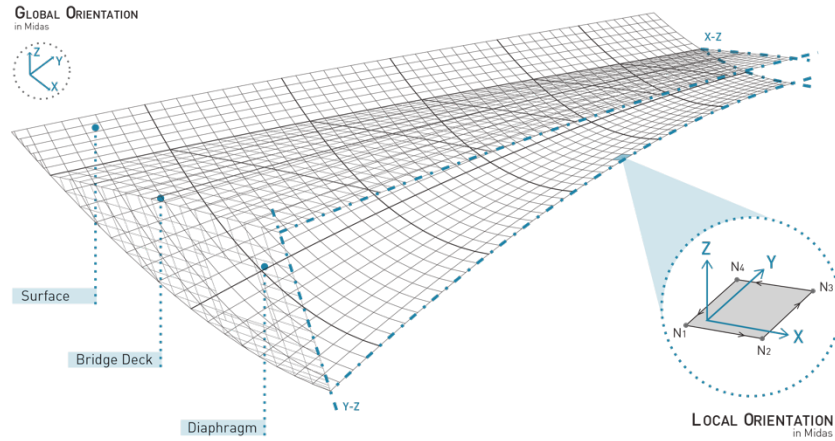


Figure 2. Structural model and mesh [3]

The first model created is certainly the most comprehensive of approximations with relatively poor reliability results. In it, the main approximation is made in the calculation of "equivalent" Young's modulus to be assigned to the mechanical properties of the material. One of the design constraints involved in the realization of different study models is represented by the orientation of the CLT panels in crossed layers. It does not follow that of the Cartesian reference system for which the values of the mechanical properties for the different directions are assigned by the literature [5], but it is dictated by the inclination of the generating lines of the ruled surface. Furthermore, since the 2D plate elements are finite, the different layers of each single panel are composed and the respective mechanical properties are taken into account in the introduction of the "equivalent" Young's modulus. The fundamental steps are:

- Computation of the projection of the "equivalent" Young's modulus, longitudinal ($E'0$) and transverse ($E'90$), based on the inclination dictated by the geometry of the surface compared to the XY global reference system;
- Multiplication of the previous components for the number of layers respectively provided for both directions;
- Considered an "equivalent" thickness (equal to $\frac{3}{4}$ of the real one) to have values of stiffness, axial and bending, not too far from the actual values (consequence of the approximation reported in the first point).

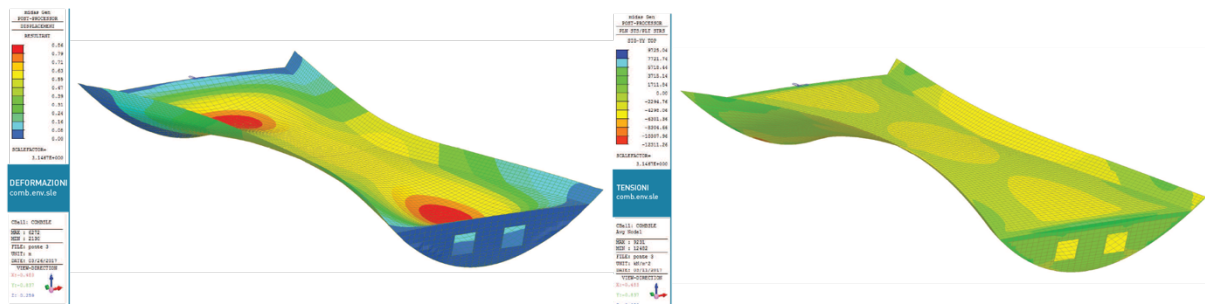


Figure 3. Displacements and strain results of the first model [3]

Because of the excessive approximations made, the results obtained by this model (shown in Fig.3) are unreliable. After several iterations in modelling, in which the transversal components ($E''0$ and $E''90$) of the Young's modulus were not neglected, a further division of

the plates in subgroups, was carried out. For each of them, the longitudinal component of the Young's modulus E ($E'0$ and $E'90$) has been recalculated and projected from time to time depending on the inclination of the elements in the global mesh. Furthermore, an equivalent thickness of $\frac{1}{2}$ of the real one was considered. This last model implemented show results that are the most reliable and all verified according to the limits set by the Italian code. In particular, keeping unchanged the mechanical properties assigned to the material once the approximations relative to the first model have been eliminated, in the last one a stiffening operation of the deck structure was carried out by inserting specially sized beams. In fact, despite being designed with an inclination of lamellar wood panels placed at 45° , it turns out to be the only element to have no enough stiffness in shape. As shown in the graphs in Fig.4), the deformation of the deck is much higher than those of the striped surface of the hyperbolic paraboloid.

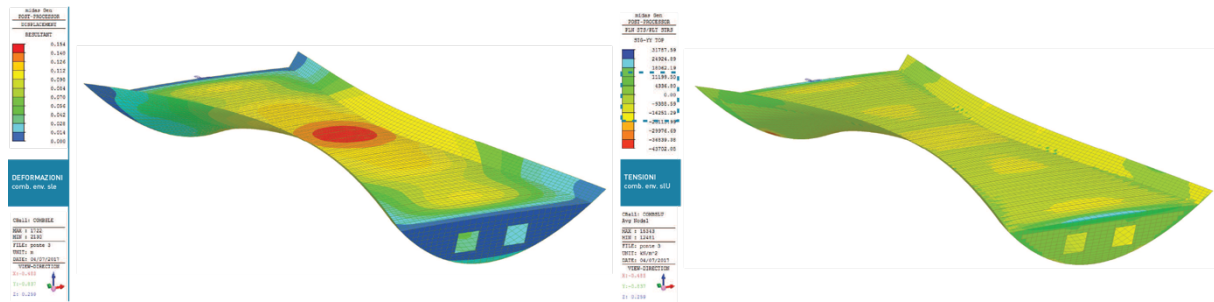


Figure 4. Displacement and Strain of the updated model [3]

In this model, the structural analysis is satisfied. However, from the graphs shown in Fig.4, it can be seen that in the diaphragm-bridge deck connection zone this check is not satisfied because of the geometry of the mesh and higher deformations in the deck.

The desire to enhance, in an alternative and innovative way, the concept of sustainable mobility design has been fulfilled by creating a meeting place; a new public space in which to stop, in a sensitive relationship with the surrounding environment. To complete the work and the tasks for this research work, an extension of the surface of the hyperbolic paraboloid used as a parapet was created in order to obtain a seat (in glulam wood) with soft shapes and in harmony with the whole boardwalk (Fig. 5). Furthermore, the upper surface of this additional element was exploited by inserting flexible and light solar panels used in the naval field [6]. The electric power produced allows the arrangement of a higher number of lighting devices than those provided for the classic walkways. So, a new urban space in which to experience a simple way of living, stopping, observing, listening, riding, walking.

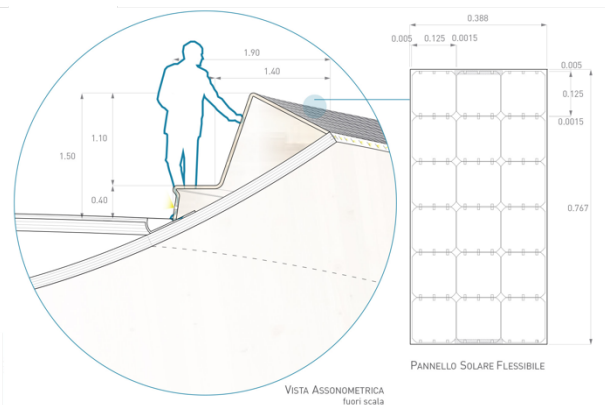


Figure 5. Detail of the rearing-seat [3]

Conclusions

The goal of this work is to give a new identity to the design of cycle-pedestrian walkways; a new way of conceiving form, linked not only to aesthetics but also to structural efficiency; an alternative design of sustainable mobility through the choice of unusual materials and shapes. Finally, it is important to underline how this project turns out to be a specific case but adaptable to any other context with different conditions. The simple construction (it may be assembled near the site and placed by two cranes in the right position), combined with flexible replicability are important features. Not last, the possibility of further optimizing the structure by pushing towards new structural models and more reduced thicknesses.

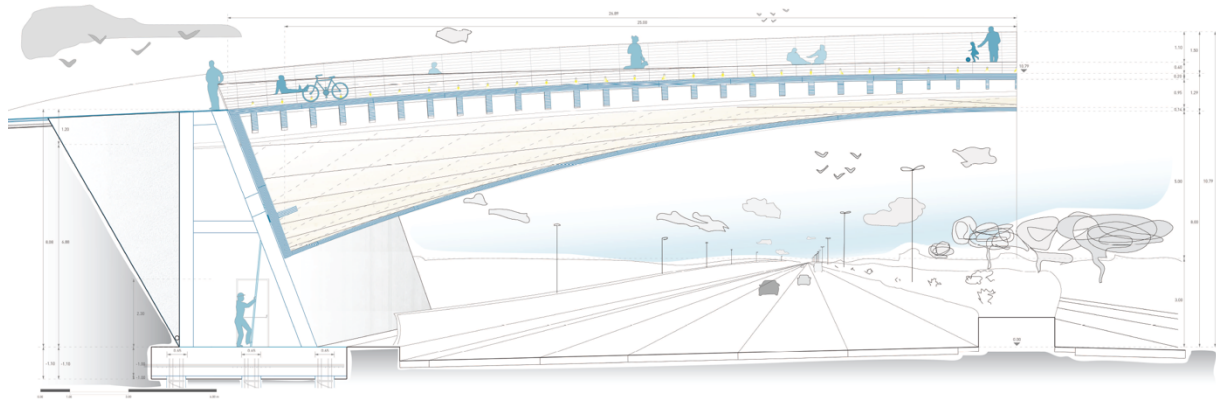


Figure 6. Longitudinal Section [3]

References

- [1] Cit. Eero Saarinen.
- [2] Antonio Turiello. (A.A.2015-2016) *Modellazione geometrica generativa: un approccio algoritmico*, Tesi. Facoltà di Ingegneria Civile dell'Università degli Studi di Salerno.
- [3] Melania Lucci. (A.A.2010-2011) *Ponte Ciclo-Pedonale in Legno dalla Forma Innovativa*, Tesi. Facoltà di Architettura dell'Università degli Studi "G. d'Annunzio"
- [4] www.cspfea.net/
- [5] www.promolegno.com/
www.dataholtz.it/it/
- [6] www.solbian.eu/it/content/9-nautica

Ball convergence for a multi-step Harmonic mean Newton-like method in Banach space

Ramandeep Behl¹, †Ali Saleh Alshomrani¹ and Ioannis K. Argyros²

¹Department of Mathematics, King Abdulaziz University, Jeddah 21577, Saudi Arabia.

² Cameron University, Department of Mathematics Sciences Lawton, OK 73505, USA

†Corresponding author: aszalshomrani@kau.edu.sa

Abstract

In this paper, we present a local convergence analysis of some iterative methods to approximate a locally unique solution of nonlinear equation in a Banach space setting. In the earlier study, Babajee et al. [1] demonstrate convergence of their methods under hypotheses on the fourth-order derivative or even higher. However, only first-order derivative of the function appears in their proposed scheme. In this study, we have shown that the local convergence of these methods depends under hypotheses only on the first-order derivative and the Lipschitz condition. In this way, we not only expand the applicability of these methods but also proposed the theoretical radius of convergence of these methods. Finally, a variety of concrete numerical examples demonstrate that our results even apply to solve those nonlinear equations where earlier studies cannot apply.

Keywords: Newton-like method, local convergence, Banach space, Lipschitz constant, radius of convergence.

Introduction

One of the most basic and important problem of Numerical analysis concerns with approximating a locally unique solution x^* of the equation of the form

$$F(x) = 0, \quad (1)$$

where F is a Fréchet -differentiable operator defined on a convex subset \mathbb{D} of a Banach space \mathbb{X} with value in a Banach space \mathbb{Y} .

Analytical methods for such type of problems are very rare or almost non existent. Therefore, it is only possible to approximate solutions by relying iterative methods. The convergence analysis of iterative methods is usually divided into two categories: semi-local and local convergence analysis. The semi-local convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iteration procedures. A very important problem in the study of iterative procedures is the convergence domain. Therefore, it is very important to propose the radius of convergence of the iterative methods.

We study the local convergence of the two step method defined for each $n = 0, 1, 2, \dots$ by

$$\begin{aligned} y_n &= x_n - \frac{2}{3}F'(x_n)^{-1}F(x_n), \\ x_{n+1} &= G_{4th\ HM}(x_n) = x_n - H_1(x_n)A(x_n)F(x_n), \\ z(x_n) &= F'(x_n)^{-1}F'(y_n), \\ H_1(x_n) &= I - \frac{1}{4}(z(x_n) - I) + \frac{1}{2}(z(x_n) - I)^2, \\ A(x_n) &= \frac{1}{2}(F'(x_n)^{-1} + F'(y_n)^{-1}) \end{aligned} \quad (2)$$

and

$$\begin{aligned}
 x_{n+1} &= G_{(2s+4)th\ HM}(x_n) = z_s(x_n) \\
 z_j(x_n) &= z_{j-1}(x_n) - H_2(x_n)A(x_n)F(z_{j-1}(x_n)), \quad j = 1, 2, \dots, s, \quad s \geq 1, \\
 H_2(x_n) &= 2I - z(x_n), \\
 z_0(x_n) &= G_{4th\ HM}(x_n),
 \end{aligned} \tag{3}$$

where, s is a natural number with $s = 0$, $x_0 \in D$ is an initial point and I is the identity operator. Notice that if $s = 0$, then method (3) reduces to method (2). These methods were studied in [1] in the special case when $\mathbb{X} = \mathbb{Y} = \mathbb{R}^i$ (i is a natural integer). Method (2) was shown to be of order four and method (3) was to be shown of order $2s + 4$. However, the local convergence was shown in [1], by using the Taylor series expansions and hypotheses reaching up to the fifth Fréchet derivative of involved operator F although only first order derivative appears in the proposed schemes. The hypotheses on the derivatives of F restrict the applicability of method (2) and method (3). As a motivational example, define function F on $\mathbb{X} = \mathbb{Y} = \mathbb{R}$, $D = [-\frac{5}{2}, \frac{1}{2}]$ by

$$F(x) = \begin{cases} x^3 \ln x^2 + x^5 - x^4, & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

Then, we have that

$$\begin{aligned}
 F'(x) &= 3x^2 \ln x^2 + 5x^4 - 4x^3 + 2x^2, \\
 F''(x) &= 6x \ln x^2 + 20x^3 - 12x^2 + 10x
 \end{aligned}$$

and

$$F'''(x) = 6 \ln x^2 + 60x^2 - 24x + 22.$$

Then, obviously, function $F'''(x)$ is unbounded on \mathbb{D} at the point $x = 0$. Hence, the results in [1], cannot apply to show the convergence of method (2) or its special cases requiring hypotheses on the fourth derivative of function F or higher. Notice that, in-particular there is a plethora of iterative methods for approximating solutions of nonlinear equations [2, 3, 4, 5, 6, 1, 7, 8, 9, 10, 11, 12, 13, 14, 15]. These results show that initial guess should be close to the required root for the convergence of the corresponding methods. But, how close initial guess should be required for the convergence of the corresponding method? These local results give no information on the radius of the ball convergence for the corresponding method. The same technique can be used to other methods.

In the present study we expand the applicability of method (2) using only hypotheses on the first order derivative of function F . We also proposed the computable radii of convergence and error bounds based on the Lipschitz constants. We further present the range of initial guess x^* that tell us how close the initial guess should be required for granted convergence of the method (2). This problem was not addressed in [1].

Local convergence in Banach space

We present the local convergence analysis that follows is based on some scalar functions and parameters. Let L_0 , $L > 0$ and $M \in [1, 3)$ be given parameters. Define functions g_1 , g_2 , h_2 , p ,

$q_j, h_{q_j}, j = 1, 2, \dots, s$ (s is a natural integer) on the interval $[0, \frac{1}{L_0})$ by

$$\begin{aligned} g_1(t) &= \frac{1}{2(1-L_0t)} \left(Lt + \frac{2M}{3} \right), \\ g_2(t) &= \frac{Lt}{2(1-L_0t)} + \frac{L_0M(1+g_1(t))t^2}{2(1-L_0t)^2} + \left[\frac{L_0(1+g_1(t))t^2}{4(1-L_0t)^2} + \frac{1}{2} \frac{L_0^2(1+g_1(t))^2t^4}{(1-L_0t)^4} \right] \frac{M^2}{(1-L_0t)^2}, \\ h_2(t) &= g_2(t) - 1, \\ p(t) &= \frac{M}{1-L_0t} \left(1 + \frac{L_0(1+g_1(t))}{1-L_0t} \right), \\ q_1(t) &= (1+p(t))g_2(t), \\ h_{q_1}(t) &= q_1(t) - 1, \\ q_j(t) &= (1+p(t))q_{j-1}(t) = (1+p(t))^j g_2(t), \quad j = 2, 3, \dots, s, \\ h_{q_j}(t) &= q_j(t) - 1 \end{aligned}$$

and parameters r_1 and r_A by

$$\begin{aligned} r_1 &= \frac{2 \left(1 - \frac{M}{3} \right)}{2L_0 + L}, \\ r_A &= \frac{2}{2L_0 + L}. \end{aligned}$$

We have that

$$0 < r_1 < r_A, \quad (4)$$

$g_1(r_1) = 1$ and for each $t \in [0, r_1] : 0 \leq g_1(t) < 1$. Moreover, by the definition of the functions g_2 and $h_2 : h_2(0) = -1$ and $h_2(t) \rightarrow +\infty$ as $t \rightarrow \frac{1}{L_0}^-$. It then follows from the intermediate value theorem that the function h_2 has zeros in the interval $(0, \frac{1}{L_0})$. Further, consider that r_2 is the smallest such zero. Similarly, we have that $h_{q_j}(0) = -1$ and $h_{q_j}(t) \rightarrow +\infty$ as $t \rightarrow \frac{1}{L_0}^-$. Denote by r_{q_j} the smallest zeros of the functions h_{q_j} , respectively on the interval $(0, \frac{1}{L_0})$. In particular, we have $h_{q_1}(r_2) = (1+p(r_2))g_2(r_2) - 1 = p(r_2) > 0$, since $1 - L_0r_2 > 0$ and $g_2(r_2) = 1$. Hence, $r_{q_1} < r_2$. Similarly, we get $h_{q_j}(r_{j-1}) = p(r_{q_{j-1}}) > 0$, since $1 - L_0r_{q_{j-1}} > 0$ and $q_{j-1}(r_{q_{j-1}}) = 1$. That is we obtain that

$$r_{q_s} < r_{q_{s-1}} < \dots < r_{q_1} < r_2. \quad (5)$$

Define

$$r = \min\{r_1, r_{q_s}\}. \quad (6)$$

Then, in view of (4) – (6), we have that

$$0 < r < r_A < \frac{1}{L_0} \quad (7)$$

and for each $t \in [0, r)$

$$0 \leq g_1(t) < 1, \quad (8)$$

$$0 \leq g_2(t) < 1, \quad (9)$$

and

$$0 \leq q_j(t) < 1. \quad (10)$$

Notice that if $s = 0$, then the radius of convergence should be defined by $r = \min\{r_1, r_2\}$. Let $U(w, \rho)$, $\bar{U}(w, \rho)$ stand, respectively for the open and closed balls in X with center $w \in X$ and of radius $\rho > 0$. Next, we present the local convergence analysis of method (3) using the preceding notations.

Theorem 1 *Let $F : D \subseteq \mathbb{X} \rightarrow \mathbb{Y}$ be a Fréchet differentiable operator. Suppose there exist $x^* \in D$ and $L_0 > 0$ such that for each $x \in D$*

$$F(x^*) = 0, \quad F'(x^*)^{-1} \in L(\mathbb{Y}, \mathbb{X}) \quad (11)$$

and

$$\|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \leq L_0\|x - x^*\|. \quad (12)$$

Moreover, suppose that there exist $L > 0$ and $M \in [1, 3)$ such that for each $x, y \in D \cap U(x^*, \frac{1}{L_0})$ the following estimates hold

$$\|F'(x^*)^{-1}(F'(x) - F'(y))\| \leq L\|x - y\|, \quad (13)$$

$$\|F'(x^*)^{-1}F'(x)\| \leq M, \quad (14)$$

and

$$\bar{U}(x^*, r) \subseteq D, \quad (15)$$

where the radius of convergence r is defined by (6). Then, the sequence $\{x_n\}$ generated by method (3) for $x_0 \in U(x^*, r) - \{x^*\}$ is well defined, remains in $U(x^*, r)$ for each $n = 0, 1, 2, \dots$ and converges to x^* . Moreover, the following estimates hold

$$\|y_n - x^*\| \leq g_1(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\| < r, \quad (16)$$

$$\|x_{n+1} - x^*\| = \|z_0(x_n) - x^*\| \leq g_2(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\|, \quad (17)$$

and

$$\begin{aligned} \|z_j(x_n) - x^*\| &\leq (1 + p(\|x_n - x^*\|))\|z_{j-1}(x_n) - x^*\| \\ &\leq (1 + p(r))^j g_2(r)\|x_n - x^*\| < \|x_n - x^*\|, \end{aligned} \quad (18)$$

where the “ g ” functions are defined previously. Furthermore, for $T \in [r, \frac{2}{L_0})$, the limit point x^* is the only solution of equation $F(x) = 0$ in $\bar{U}(x^*, T) \cap D$.

Proof: Mathematical induction shall be used to show estimates (16) – (18). By hypotheses $x_0 \in U(x^*, r) - \{x^*\}$, (7) and (12), we get that

$$\|F'(x^*)^{-1}(F'(x_0) - F'(x^*))\| \leq L_0\|x - x^*\| < L_0r < 1. \quad (19)$$

In view of (19) and the Banach Lemma on invertible operators [5, 13], we get that $F'(x_0)^{-1} \in L(Y, X)$, y_0 exists and

$$\|F'(x_0)^{-1}F'(x^*)\| \leq \frac{1}{1 - L_0\|x_0 - x^*\|}. \quad (20)$$

We also get that y_0 is well defined by the first sub step of method (3) for $n = 0$. We can write by (11)

$$F(x_0) = F(x_0) - F(x^*) = \int_0^1 F'(x^* + \theta(x_0 - x^*))(x_0 - x^*)d\theta \quad (21)$$

Notice that $\|x^* + \theta(x_0 - x^*) - x^*\| = \theta\|x_0 - x^*\| < r$, so $x^* + \theta(x_0 - x^*) \in U(x^*, r)$. Then, by (14) and (21), we have that

$$\|F'(x^*)^{-1}F'(x_0)\| \leq M\|x_0 - x^*\| \quad (22)$$

We can write by the first sub step of method (3) and (11)

$$\begin{aligned} y_0 - x^* &= \left(F'(x_0)^{-1}F'(x^*)\right) \int_0^1 F'(x^*)^{-1} (F'(x^* + \theta(x_0 - x^*)) - F'(x_0)) (x_0 - x^*)d\theta \\ &\quad + \frac{1}{3} \left(F'(x_0)^{-1}F'(x^*)\right) \left(F'(x^*)^{-1}F'(x_0)\right) \end{aligned} \quad (23)$$

Using (7), (8), (13), (20), (22) and (23), we obtain in turn that

$$\begin{aligned} \|y_0 - x^*\| &= \|F'(x_0)^{-1}F'(x^*)\| \left\| \int_0^1 F'(x^*)^{-1} (F'(x^* + \theta(x_0 - x^*)) - F'(x_0)) (x_0 - x^*)d\theta \right\| \\ &\quad + \frac{1}{3} \|F'(x_0)^{-1}F'(x^*)\| \|F'(x^*)^{-1}F'(x_0)\| \\ &\leq \frac{L\|x_0 - x^*\|^2}{2(1 - L\|x_0 - x^*\|)} + \frac{M\|x_0 - x^*\|}{3(1 - L\|x_0 - x^*\|)} \\ &= g_1(\|x_0 - x^*\|)\|x_0 - x^*\| < \|x_0 - x^*\| < r, \end{aligned} \quad (24)$$

which shows (16) for $n = 0$ and $y_0 \in U(x^*, T)$. It follows from (19), (20) and (24) for y_0 replacing x_0 , since $y_0 \in U(x^*, T)$ that $F'(y_0)^{-1} \in L(\mathbb{Y}, \mathbb{X})$

$$\begin{aligned} \|F'(y_0)^{-1}F'(x^*)\| &\leq \frac{1}{1 - L_0\|y_0 - x^*\|} \\ &\leq \frac{1}{1 - L_0g_1(\|x_0 - x^*\|)} \\ &\leq \frac{1}{1 - L_0\|x_0 - x^*\|} \end{aligned} \quad (25)$$

and x_1 is well defined by the second sub step of the method (3) for $n = 0$. Then, we can write by the second sub step of method (3) for $n = 0$ in turn that but using (9) instead of (8), we

obtain in turn that

$$\begin{aligned}
 x_1 - x^* &= (x_0 - x^* - F'(x_0)^{-1}F(x_0)) + \frac{1}{2}(F'(x_0)^{-1}F(x^*)) [F'(x^*)^{-1}(F'(x_0) - F'(x^*)) \\
 &\quad + F'(x^*)^{-1}(F'(x^*) - F'(y_0))] (F'(y_0)^{-1}F(x^*)) (F'(x^*)^{-1}F(x_0)) \\
 &\quad + \left[\frac{1}{4}(F'(x_0)^{-1}F(x^*)) (F'(x^*)^{-1}(F'(y_0) - F'(x^*)) + F'(x^*)^{-1}(F'(x^*) - F'(x_0))) \right. \\
 &\quad \times (F'(y_0)^{-1}F'(x^*)) F'(x^*)^{-1} + \frac{1}{2} \left((F'(x_0)^{-1}F(x^*)) (F'(x^*)^{-1}(F'(y_0) - F'(x^*)) \right. \\
 &\quad \left. \left. + F'(x^*)^{-1}(F'(x^*) - F'(x_0))) (F'(y_0)^{-1}F'(x^*)) F'(x^*) \right)^2 \right] \frac{1}{2}(F'(x_0)^{-1}F'(x^*)^{-1}) \\
 &\quad \times (F'(x^*)^{-1}F'(y_0) + F'(x^*)^{-1}F'(x_0)) (F'(y_0)^{-1}F(x^*)) (F'(x^*)^{-1}F(x_0)).
 \end{aligned} \tag{26}$$

Then, using the triangle inequality in (26), (7), (9), (20), (22) (for $x_0 = x_0$ and $x_0 = y_0$), (24) and (25), we get in turn that

$$\begin{aligned}
 \|x_1 - x^*\| &\leq \frac{L\|x_0 - x^*\|^2}{2(1 - L_0(\|x_0 - x^*\|))} + \frac{L_0M(\|x_0 - x^*\| + \|y_0 - x^*\|)\|x_0 - x^*\|^2}{2(1 - L_0\|x_0 - x^*\|)(1 - L_0\|y_0 - x^*\|)} \\
 &\quad + \left[\frac{L_0(\|x_0 - x^*\| + \|y_0 - x^*\|)\|x_0 - x^*\|}{4(1 - L_0\|x_0 - x^*\|)(1 - L_0\|y_0 - x^*\|)} \right. \\
 &\quad \left. + \frac{1}{2} \left(\frac{L_0(\|x_0 - x^*\| + \|y_0 - x^*\|)\|x_0 - x^*\|}{(1 - L_0\|x_0 - x^*\|)(1 - L_0\|y_0 - x^*\|)} \right)^2 \right] \frac{M^2\|x_0 - x^*\|}{(1 - L_0\|x_0 - x^*\|)(1 - L_0\|y_0 - x^*\|)} \\
 &\leq g_2(\|x_0 - x^*\|)\|x_0 - x^*\| < \|x_0 - x^*\| < r,
 \end{aligned} \tag{27}$$

which shows (17) for $n = 0$ and $x_1 \in U(x^*, T)$. Notice that by the definition of the method (3), $x_1 = z_0(x_0)$ and all iterates $z_j(x_0)$, $j = 1, 2, \dots, s$ are well defined. Then, we have by the third sub step of method (3) for $n = 0$

$$\begin{aligned}
 \|z_1(x_0) - x^*\| &\leq \|z_0(x_0) - x^*\| + \|H_2(x_0)A(x_0)F'(x^*)\|M\|z_0(x_0) - x^*\| \\
 &= (1 + M\|H_2(x_0)A(x_0)F'(x^*)\|)\|z_0(x_0) - x^*\|.
 \end{aligned} \tag{28}$$

We need an upper bound on the norm inside (28). It follows from the definition of A , H_2 and (27) that

$$\begin{aligned}
 \|H_2(x_0)A(x_0)F'(x^*)\| &\leq \frac{1}{2} \left(\frac{1}{1 - L_0\|x_0 - x^*\|} + \frac{1}{1 - L_0\|y_0 - x^*\|} \right) (\|I\| + \|I - u(x_0)\|) \\
 &\leq \frac{1}{1 - L_0\|x_0 - x^*\|} \left(1 + \|F'(x_n)^{-1}(F'(x_n) - F'(y_n))\| \right) \\
 &\leq \frac{1}{1 - L_0\|x_0 - x^*\|} \left(1 + \frac{L_0(\|x_0 - x^*\| + \|y_0 - x^*\|)}{1 - L_0\|x_0 - x^*\|} \right) \\
 &\leq \frac{1}{1 - L_0\|x_0 - x^*\|} \left(1 + \frac{L_0(1 + g_1(\|x_0 - x^*\|)\|x_0 - x^*\|)}{1 - L_0\|x_0 - x^*\|} \right) \\
 &\leq \frac{p(\|x_0 - x^*\|)}{M}.
 \end{aligned} \tag{29}$$

Using (28) and (29), we get that

$$\begin{aligned}\|z_1(x_0) - x^*\| &\leq (1 + p(\|x_0 - x^*\|))\|z_0(x_0) - x^*\| \\ &\leq (1 + p(\|x_0 - x^*\|))g_2(\|x_0 - x^*\|)\|x_0 - x^*\| \\ &= q_1(\|x_0 - x^*\|)\|x_0 - x^*\| < \|x_0 - x^*\| < r,\end{aligned}\tag{30}$$

so $z_1(x_0) \in U(x^*, r)$ and (18) for $n = 0$ and $j = 1$. In an analogous way by using $z_j(x_0)$, $z_{j-1}(x_0)$ instead of $z_1(x_0)$, $z_0(x_0)$ in (28) and (29), we get that (18) holds for $n = 0$ and $j = 1, 2, \dots, s$. Hence, x_2 is well defined in $x_2 \in U(x^*, r)$ and by (18) and the definition of the method (3) $\|x_2 - x_1\| \leq \|x_1 - x^*\|$. Continuing in this way, we arrive at the estimates (16)–(18) and $\|x_{k+1} - x^*\| \leq c\|x_k - x^*\| < r$, $c = g_2(\|x_0 - x^*\|) \in [0, 1)$, which shows $x_{k+1} \in U(x^*, r)$ and $\lim_{k \rightarrow \infty} x_k = x^*$. Finally, to show the uniqueness part, let $y^* \in \bar{U}(x^*, T)$ be such that $F(y^*) = 0$. Set $Q = \int_0^1 F'(x^* + \theta(y^* - x^*)) d\theta$. Then, using (12), we get that

$$\|F'(x^*)^{-1}(Q - F'(x^*))\| \leq L_0 \int_0^1 \theta \|x^* - y^*\| d\theta = \frac{L_0}{2} T < 1.\tag{31}$$

Hence, $Q^{-1} \in L(Y, X)$. Then, in view of the identity $F(y^*) - F(x^*) = Q(y^* - x^*)$, we conclude that $x^* = y^*$. \square

Remark

(a) In view of (12) and the estimate

$$\begin{aligned}\|F'(x^*)^{-1}F'(x)\| &= \|F'(x^*)^{-1}(F'(x) - F'(x^*) + F'(x^*))\| \\ &\leq 1 + \|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \\ &\leq 1 + L_0\|x_0 - x^*\|,\end{aligned}$$

condition (14) can be dropped and M can be replaced by

$$M = M(t) = 1 + L_0 t$$

or $M = 2$, since $t \in [0, \frac{1}{L_0})$.

(b) The radius r_1 was shown in [5, 6] to be the convergence radius for Newton's method under conditions (12) and (13). It follows from (5) and the definition of r_1 that the convergence radius r of the method (2) cannot be larger than the convergence radius r_1 of the second order Newton's method. As already noted in [5, 6], r_1 is at least as large as the convergence ball give by Rheinboldt [13]

$$r_R = \frac{2}{3L}.$$

In particular, for $L_0 < L$ we have that

$$r_R < r_1$$

and

$$\frac{r_R}{r_1} \rightarrow \frac{1}{3} \quad \text{as} \quad \frac{L_0}{L} \rightarrow 0.$$

That is our convergence ball r_1 is at most three times larger than Rheinboldt's. The same value for r_R given by Traub [14].

- (c) It is worth noticing that method (2) is not changing if we use the conditions of Theorem 1 instead of the stronger conditions given in [1]. Moreover, for the error bounds in practice we can use the computational order of convergence (COC) [9]

$$\xi = \frac{\ln \frac{\|x_{n+2}-x^*\|}{\|x_{n+1}-x^*\|}}{\ln \frac{\|x_{n+1}-x^*\|}{\|x_n-x^*\|}}, \quad \text{for each } n = 0, 1, 2, \dots \quad (32)$$

or the approximate computational order of convergence (ACOC) [9]

$$\xi^* = \frac{\ln \frac{\|x_{n+2}-x_{n+1}\|}{\|x_{n+1}-x_n\|}}{\ln \frac{\|x_{n+1}-x_n\|}{\|x_n-x_{n-1}\|}}, \quad \text{for each } n = 1, 2, \dots \quad (33)$$

This way we obtain in practice the order of convergence in a way that avoids the bounds involving estimates higher than the first Fréchet derivative. Notice that the evaluation of ξ^* does not require that the usage of the solution x^* .

- (d) If $s = 0$ and $r = \min\{r_1, r_2\}$, then the results of Theorem 1 hold for method (2) replacing method (3) (except (18)).

Numerical example and applications

In this section, we shall check the effectiveness and validity of our theoretical results which we have presented in section 2 on the scheme proposed by Babajee et al. [1]. For this purpose, we shall choose a variety of nonlinear equations and system of nonlinear equations which are mentioned in the following examples including motivational example. At this point, we will choose the following methods

$$\left\{ \begin{array}{l} y_n = x_n - \frac{2}{3}F'(x_n)^{-1}F(x_n), \\ x_{n+1} = G_{4th\ HM} = x_n - H_1(x_n)A(x_n)F(x_n), \\ z(x_n) = F'(x_n)^{-1}F'(y_n), \\ H_1(x_n) = I - \frac{1}{4}(z(x_n) - I) + \frac{1}{2}(z(x_n) - I)^2, \\ A(x_n) = \frac{1}{2}(F'(x_n)^{-1} + F'(y_n)^{-1}) \end{array} \right. \quad (34)$$

$$\left\{ \begin{array}{l} y_n = x_n - \frac{2}{3}F'(x_n)^{-1}F(x_n), \\ z_n = x_n - H_1(x_n)A(x_n)F(x_n), \\ x_{n+1} = z_n - H_2(x_n)A(x_n)F(z_n), \\ H_2(x_n) = 2I - F'(x_n)^{-1}F'(y_n) \end{array} \right. \quad (35)$$

and

$$\left\{ \begin{array}{l} y_n = x_n - \frac{2}{3}F'(x_n)^{-1}F(x_n), \\ z_n = x_n - H_1(x_n)A(x_n)F(x_n), \\ w_n = z_n - H_2(x_n)A(x_n)F(z_n), \\ x_{n+1} = w_n - H_2(x_n)A(x_n)F(w_n), \end{array} \right. \quad (36)$$

having convergence order of $p = 4$, $p = 6$ and $p = 8$ which can be deduced by using $s = 0$, $s = 1$ and $s = 2$, respectively. For computational point of view, we denoted them by M_1 , M_2 and M_3 , respectively.

First of all, we shall calculate the values of r_R , r_1 , r_A , r_2 , r_{qs} and r which are displayed in the Tables 1, 3, 5, 6. So, we can obtain the radius of convergence of the above mentioned methods. Then, we will also verify the theoretical order of convergence of these methods for scalar equations on the basis of the results obtain from computational order of convergence and $\left| \frac{e_n}{e_{n-1}^p} \right|$ (where p is either $p = 4, 6$ or $p = 8$). In the Tables 2, 4 and 7, we displayed the number of iteration indexes (n), approximated zeros (x_n), residual error of the corresponding function ($|F(x_n)|$), errors $|e_n|$ (where $e_n = x_n - x^*$), $\left| \frac{e_n}{e_{n-1}^p} \right|$ and the asymptotic error constant $\eta = \lim_{n \rightarrow \infty} \left| \frac{e_n}{e_{n-1}^p} \right|$. In addition, we calculate the computational order of convergence by using the above formulas (32) and (33). Moreover, we calculate the computational order of convergence, asymptotic error constant and other constants up to several number of significant digits (minimum 1000 significant digits) to minimize the round off error.

In the context of system of nonlinear equations, we also consider a nonlinear system in example 3 to check the proposed theoretical results for nonlinear system. In this regards, we displayed the number of iteration indexes (n), residual error of the corresponding function ($\|F(x_n)\|$), errors $\|e_n\|$ (where $e_n = x_n - x^*$), $\left\| \frac{e_n}{e_{n-1}^p} \right\|$ and the asymptotic error constant $\eta = \lim_{n \rightarrow \infty} \left\| \frac{e_n}{e_{n-1}^p} \right\|$ in the Table 7. Moreover, we use the above mentioned formulas namely, (32) and (33) to calculate the computational order of convergence to further verifying the theoretical order of convergence of nonlinear system.

As we mentioned in the earlier paragraph that we calculate the values of all the constants and functional residuals up to several number of significant digits but due to the limited paper space, we display the values of x_n up to 15 significant digits and the values of other constants namely, r_R , r_1 , r_A , r_2 , r_{qs} , r , $\xi(COC)$, $\left| \frac{e_n}{e_{n-1}^p} \right|$, η and $\left\| \frac{e_n}{e_{n-1}^p} \right\|$ are up to 5 significant digits. Further, the residual error in the function/system of nonlinear functions ($|F(x_n)|$ or $\|F(x_n)\|$), and the error $|e_n|$ or $\|e_n\|$ are display up to 2 significant digits with exponent power which are mentioned in the following Tables corresponding to the test function. However, minimum 1000 significant digits are available with us for every value.

Furthermore, we consider the approximated zero of test functions when the exact zero is not available, which is corrected up to 1000 significant digits to calculate $\|x_n - x^*\|$. For the computer programming, all computations have been performed using the programming package *Mathematica* 9 with multiple precision arithmetic. In addition, the meaning of $ae(\pm b)$ is $a \times 10^{\pm b}$ in the tables 1–7.

Example 1 Let $S = \mathbb{R}$, $D = [-1, 1]$, $x^* = 0$ and define function F on D by

$$F(x) = \sin x. \quad (37)$$

Then, we get $L_0 = L = 1$ and $M = 1$. We calculate the different values of the radius of convergence, COC (ξ) etc., which are displayed in the following Tables 1 and 2.

Table 1: Different values of parameters which satisfy Theorem 1

Cases	r_R	r_1	r_A	r_2	r_{qs}	r
M_1	0.66667	0.44444	0.66667	0.33913	0.19871	0.19871
M_2	0.66667	0.44444	0.66667	0.33913	0.097223	0.097223
M_3	0.66667	0.44444	0.66667	0.33913	0.0401905	0.0401905

Table 2: Convergence behavior of different cases on example 1

Cases	n	x_n	$ F(x_n) $	$ e_n $	ξ	$\frac{e_n}{e_{n-1}}$	η
M_1	0	0.15	$1.5e(-1)$	$1.5e(-1)$	5.0045	$8.3497e(-3)$ $2.2442e(-7)$	$2.2442e(-7)$
	1	$4.22751609372639e(-6)$	$4.2e(-6)$	$4.2e(-6)$			
	2	$7.16819456281920e(-29)$	$7.2e(-29)$	$7.2e(-29)$			
M_2	0	0.070	$7.0e(-2)$	$7.0e(-2)$	7.0009	$6.3153e(-4)$ $6.57738e(-13)$	$6.5738e(-13)$
	1	$-7.42990061829384e(-11)$	$7.4e(-11)$	$7.4e(-11)$			
	2	$1.10589328064363e(-73)$	$1.1e(-73)$	$1.1e(-73)$			
M_3	0	0.020	$2.0e(-2)$	$2.0e(-2)$	9.0001	$2.9561e(-5)$ $1.1160e(-21)$	$1.1160e(-21)$
	1	$-7.56773592959322e(-19)$	$7.6e(-19)$	$7.6e(-19)$			
	2	$1.20053454004807e(-166)$	$1.2e(-166)$	$1.2e(-166)$			

Example 2 Let $X = Y = \mathbb{R}^3$, $D = \bar{U}(0, 1)$, $v = (x, y, z)^T$ and defined F on D by

$$F(v) = \left(e^x - 1, \frac{e-1}{2}y^2 + y, z \right)^T. \quad (38)$$

Then the Fréchet-derivative is given by

$$F'(v) = \begin{bmatrix} e^x & 0 & 0 \\ 0 & (e-1)y + 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Notice that $x^* = (0, 0, 0)^T$, $F'(x^*) = F'(x^*)^{-1} = \text{diag}\{1, 1, 1\}$, $L_0 = e - 1$, $L = 1.789572397$ and $M = 1.7896$. Hence, we calculate the different values of the radius of convergence, COC (ξ) etc., which are mentioned in the following Tables 3 and 4.

Table 3: Different values of parameters which satisfy Theorem 1

Cases	r_R	r_1	r_A	r_2	r_{qs}	r
M_1	0.37253	0.15440	0.38269	0.17719	0.067047	0.067047
M_2	0.37253	0.15440	0.38269	0.17719	0.0152413	0.0152413
M_3	0.37253	0.15440	0.38269	0.17719	0.0023526	0.0023526

Example 3 Let $\mathbb{X} = \mathbb{Y} = C[0, 1]$, and consider the nonlinear integral equation of Hammerstein-type defined by

$$x(s) = \int_0^1 G(s, t) \frac{x(t)^2}{2} dt, \quad (39)$$

Table 4: Convergence behavior of different cases on example 2

Cases, x_0	n	$\ F(x_n)\ $	$\ e_n\ $	ξ	$\left\ \frac{e_n}{e_{n-1}^p}\right\ $	η
$M_1, (0.017, 0.017, 0.018)$	0	$3.0e(-2)$	$3.0e(-2)$	3.8124	0.17667 1.7584	1.7584
	1	$1.4e(-7)$	$1.4e(-7)$			
	2	$7.5e(-28)$	$7.5e(-28)$			
$M_2, (0.0022, 0.0022, 0.0023)$	0	$3.4e(-3)$	$3.4e(-3)$	5.8845	0.19198 5.7481	5.7481
	1	$6.4e(-16)$	$6.4e(-16)$			
	2	$4.1e(-91)$	$4.1e(-91)$			
$M_3 (0.00022, 0.00022, 0.00023)$	0	$3.9e(-4)$	$3.9e(-4)$	7.9202	0.19642 18.010	18.010
	1	$9.9e(-29)$	$9.9e(-29)$			
	2	$1.6e(-223)$	$1.6e(-223)$			

where the kernel G is the Green's function defined on the interval $[0, 1] \times [0, 1]$ by

$$G(s, t) = \begin{cases} (1-s)t, & t \leq s \\ s(1-t), & s \leq t. \end{cases} \quad (40)$$

The solution $x^*(s) = 0$ is the same as the solution of equation (1), where operator $F : C[0, 1] \rightarrow C[0, 1]$ is defined by

$$F(x)(s) = x(s) - \int_0^1 G(s, t) \frac{x(t)^2}{2} dt. \quad (41)$$

Notice that

$$\left\| \int_0^1 G(s, t) dt \right\| \leq \frac{1}{8}. \quad (42)$$

Then, we have that the Fréchet- derivative is defined by

$$F'(x)(y(s)) = y(s) - \int_0^1 G(s, t) x(t) dt. \quad (43)$$

So, we get that $F'(x^*(s)) = I$ and

$$\left\| F'(x^*)^{-1}(F'(x) - F'(y)) \right\| \leq \frac{1}{8} \|x - y\|. \quad (44)$$

Hence, we can choose $L_0 = L = \frac{1}{8}$ and $M = 2$. We calculate the different values of the radius of convergence based on the methods, which are mentioned in the following Table 5.

Table 5: Different values of parameters which satisfy Theorem 1

Cases	r_R	r_1	r_A	r_2	r_{q_s}	r
M_1	5.3333	1.7778	5.3333	1.0014	0.61000	0.61000
M_2	5.3333	1.7778	5.3333	1.0014	0.33700	0.33700
M_3	5.3333	1.7778	5.3333	1.0014	0.16848	0.16848

Example 4 Returning back to the motivation example at the introduction on this paper, we have $L = L_0 = 14.5$, $M = 2$ and our required zero is $x^* = 1$. We calculate the different values of the radius of convergence, COC (ξ) etc., which are given in the following Tables 6 and 7.

Table 6: Different values of parameters which satisfy Theorem 1

Cases	r_R	r_1	r_A	r_2	r_{q_s}	r
M_1	0.045977	0.015326	0.045977	0.030850	0.002334	0.002334
M_2	0.045977	0.015326	0.045977	0.030850	0.000052082	0.000052082
M_3	0.045977	0.015326	0.045977	0.030850	$1.01954e(-6)$	$1.01954e(-6)$

Table 7: Convergence behavior of different cases on example 4

Cases	n	x_n	$ F(x_n) $	$ e_n $	ξ	$\frac{e_n}{e_{n-1}^p}$	η
M_1	0	0.318	$7.3e(-5)$	$3.1e(-4)$	4.0007	1636.0 1615.17	1615.17
	1	0.318309886198877	$3.5e(-12)$	$1.5e(-11)$			
	2	0.318309886183791	$2.0e(-41)$	$8.4e(-41)$			
M_2	0	0.3183	$2.3e(-6)$	$9.9e(-6)$	6.0000	$4.5854e(+5)$ $4.5825e(+5)$	$4.5825e(+5)$
	1	0.318309886183791	$1.0e(-25)$	$4.3e(-25)$			
	2	0.318309886183791	$6.6e(-142)$	$2.8e(-141)$			
M_3	0	0.3183	$2.3e(-6)$	$9.9e(-6)$	8.0000	$1.3008e(+8)$ $1.2997e(+8)$	$1.2997e(+8)$
	1	0.318309886183791	$2.8e(-33)$	$1.2e(-32)$			
	2	0.318309886183791	$1.2e(-248)$	$5.1e(-248)$			

Results and discussion

It is worthy to note that the radii of convergence in the Table 6 are very small. Actually, the radius of convergence depends on the considered function, corresponding bounds and the body of structure of the iterative methods. We can see these things in the above Tables 1, 3 and 5 where we obtain better radii of convergence rather than Table 6. In addition, we also want to check the convergence behavior of the listed methods, when we consider initial approximation out of the convergence domain which can be seen in the Table 6. So, we can say that the proposed iterative methods will always converge to the required root whenever we consider the initial approximation inside of convergence domain however, it is also possible outside the domain. No doubts, the radius of convergence of the listed scheme is decreasing by increasing the number of sub-steps. But, we are getting better and faster convergence towards the required root which can be seen the Tables 2, 4 and 7.

References

- [1] Babajee, D.K.R., Madhu, K. and Jayaraman, J. (2015) On some improved Harmonic mean Newton-like methods for solving systems of nonlinear equations. *Algorithms* **8**(4), 895–909.
- [2] Amat, S., Busquier, S. and Plaza, S. (2005) Dynamics of the King and Jarratt iterations *Aequationes Math.* **69**(3), 212–223.
- [3] mat, S., Busquier, S. and Plaza, S. (2010) Chaotic dynamics of a third-order Newton-type method. *J. Math. Anal. Appl.* **366**(1), 24–32.
- [4] Amat, S., Hernández, M.A. and Romero, N. (2008) A modified Chebyshev's iterative method with at least sixth order of convergence. *Appl. Math. Comput.* **206**(1), 164–174.
- [5] Argyros, I.K. (2008) Convergence and Application of Newton-type Iterations. Springer.
- [6] Argyros, I.K. and Hilout, S. (2013) *Numerical methods in Nonlinear Analysis*. World Scientific Publ. Comp. New Jersey.
- [7] Behl R. and Motsa, R. (2015) Geometric construction of eighth-order optimal families of Ostrowski's method. *T. Sci. Wor. J.* **2015**, Article ID 614612, 11 pages.
- [8] Kanwar, V., Behl, R. and Sharma, K.K. (2011) Simply constructed family of a Ostrowski's method with optimal order of convergence. *Comput. Math. Appli.* **62** (11), 4021–4027.

- [9] Ezquerro, J.A. and Hernández, M.A. (2009) New iterations of R-order four with reduced computational cost. *BIT Numer. Math.* **49**, 325– 342 .
- [10] Magreñán, Á.A. (2014) Different anomalies in a Jarratt family of iterative root-finding methods. *Appl. Math. Comput.* **233**, 29–38.
- [11] Magreñán, Á.A. (2014) A new tool to study real dynamics: The convergence plane. *Appl. Math. Comput.* **248**, 215–224.
- [12] Petkovic, M.S., Neta, B., Petkovic, L. and Džunič, J. (2013) Multipoint methods for solving nonlinear equations. Elsevier.
- [13] Rheinboldt, W.C. (1978) An adaptive continuation process for solving systems of nonlinear equations. *Pol. Acad. Sci. Ban. Ctr. Publ.* **3**, 129–142.
- [14] Traub, J.F. (1964) Iterative methods for the solution of equations. Prentice- Hall Series in Automatic Computation, Englewood Cliffs, N.J. .
- [15] S. Weerakoon, T.G.I. Fernando, (2000) A variant of Newton’s method with accelerated third order convergence. *Appl. Math. Lett.* **13**, 87–93.

Numerical simulation of natural convection from a heated cylinder

†*Gregor Kosec¹, Jure Slak¹

¹Parallel and Distributed Systems Laboratory

“Jožef Stefan” Institute, Jamova 39

SI-1000 Ljubljana, Slovenia

†*Presenting and corresponding author: gregor.kosec@ijs.si

Abstract

In this paper, we present a meshless solution of natural convection from the heated cylinder. The numerical technique is constructed around Weighted Least Squares approximation that is used to evaluate derivatives needed to solve partial differential equations governing the problem at hand, i.e. Navier-Stokes, mass continuity, and heat transport. The results are presented in terms of temperature and velocity magnitude contour plots, as well as a more quantitative comparison with reference data in terms of Nusselt number and maximal velocity values. Three different cases are tackled, namely the standard de Vahl Davis case, natural convection from the heated cylinder, and cooling of overhead power line due to the natural convection under realistic conditions.

Keywords: Navier-Stokes, MLSM, de Vahl Davis, natural convection, cooling of overhead power lines.

Introduction

Natural convection from a horizontal cylinder plays a crucial role in many heat transfer related problems ranging from heat exchangers, solar heating systems, cooling of electronic packages, to cooling of overhead power lines. One of the first studies of convective heat transfer from circular cylinders goes back to 1892 when Ayrton and Kilgour investigated the thermal emission of thin, long horizontal wires [1]. Several similar experiments followed [2], and in 1975 Morgan collected experimental data in a comprehensive review paper [3]. Based on the collected data Morgan introduced the correlations between Grashof, Prandtl and Nusselt numbers [3]. In other words, he presented relationships between the power of cooling in dependence on the material properties and the temperature difference between the cylinder and the ambient, which still serve as a basis in operative models for predicting the temperature of overhead power lines [4, 5]. There are many other similar studies where authors investigate the heat transfer from the heated cylinder due to the natural or forced convection under different conditions [2, 6, 7].

Probably the most famous article on the numerical investigation of natural convection was published in 1983 by de Vahl Davis [8], who defined and solved a reference solution for natural convection of air in a closed rectangular cavity with differentially heated vertical and isolated horizontal walls. Many researchers followed his paper and solved proposed benchmark case with different numerical techniques [9, 10], gradually establishing high confidence in the numerical solution. Similar benchmark case for the natural convection around the homogeneously heated cylinder was introduced in 1992 by Demirdžić, et al. [11], again researchers responded with different numerical solutions [12].

Although numerical methods such as the Finite Volume Method, Finite Difference Method, or the Finite Element Method are typically used for solving natural convection problems, there has been also a considerable research done in an alternative meshless solution [13-15], which appeared in the seventies with Smoothed Particles Hydrodynamics (SPH) [16]. The SPH is an answer for attacking problems, where mesh-based methods fail entirely, e.g. breaking waves, however, at the cost of inconsistency due to the combination of Eulerian kernel and Lagrangian description of motion. Nevertheless, since the introduction of SPH a myriad of different strong and weak form meshless methods appeared [17]. The conceptual difference between meshless methods and mesh-based methods is in the treatment of relations between nodes. In the mesh-based techniques the nodes need to be structured into a mesh that covers the whole computational domain, while the meshless methods do not require any special relations between nodes and can be fully defined only through the relative inter nodal positions [18]. An immediate consequence of such simplification is greater generality regarding the approximation and the position of computational points and much higher flexibility in implementation.

In this paper, we present a local meshless solution that is based on Weighted Least Squares approximation of a natural convection problem in three scenarios: the de Vahl Davis case, the Demirdžić case, and finally the cooling of overhead power line.

The rest of the paper is organized as follows. In section Governing problem the problem is introduced in section Meshless solution procedure all the details for implementation of the solution procedure are presented, in section Results, the analyses of present study are given, and in the last section, paper offers conclusions and guidelines for future work.

Governing problem

The physical model for describing natural convection is well-established. The fluid mechanics is described with the Navier-Stokes equation, i.e. the Cauchy momentum equation with Newtonian stress tensor, and mass continuity, which is coupled with the heat transfer through the Boussinesq approximation. The model can be written nicely in the following system of partial differential equations

$$\nabla \cdot \mathbf{v} = 0, \quad (1)$$

$$\rho \frac{\partial \mathbf{v}}{\partial t} + \rho \nabla \cdot (\mathbf{v} \mathbf{v}) = -\nabla P + \nabla \cdot (\mu \nabla \mathbf{v}) + \mathbf{b}, \quad (2)$$

$$\rho \frac{\partial (c_p T)}{\partial t} + \rho \nabla \cdot (c_p T \mathbf{v}) = \nabla \cdot (\lambda \nabla T), \quad (3)$$

$$\mathbf{b} = \rho [1 - \beta_T (T - T_{\text{ref}})] \mathbf{g}, \quad (4)$$

with $\mathbf{v}(u, v)$, P , T , λ , c_p , \mathbf{g} , ρ , β_T , T_{ref} , μ and \mathbf{b} standing for velocity, pressure, temperature, thermal conductivity, specific heat, gravitational acceleration, density, the coefficient of thermal expansion, reference temperature for Boussinesq approximation, viscosity and body force, respectively. The natural convection can be characterised by two dimensionless numbers

$$\text{Ra} = \frac{|\mathbf{g}| \beta_T \Delta T \Omega^3 \rho^2 c_p}{\lambda \mu}, \quad (5)$$

$$\text{Pr} = \frac{\mu C_p}{\lambda}, \quad (6)$$

referred to as Rayleigh and Prandtl numbers, respectively, where Ω stands for the domain dimension.

We will consider three cases. First, the de Vahl Davis case to confirm the solution procedure, second, the heated cylinder case to demonstrate the flexibility of meshless method regarding the geometry, and finally a cooling of the overhead power line by natural convection under realistic conditions. The cases differ in geometry, boundary conditions and thermo-physical properties. The de Vahl Davis case is the most straightforward, since the closed square cavity with impermeable no-slip velocity boundary conditions, differentially heated vertical, and isolated horizontal walls are considered. The heated cylinder case is a bit more challenging due to the irregular geometry introduced by a cylinder in the domain. The last case, the cooling of the overhead power line is the most challenging due to realistic conditions and presence of velocity Neumann boundary conditions. All three cases are presented in Figure 1.

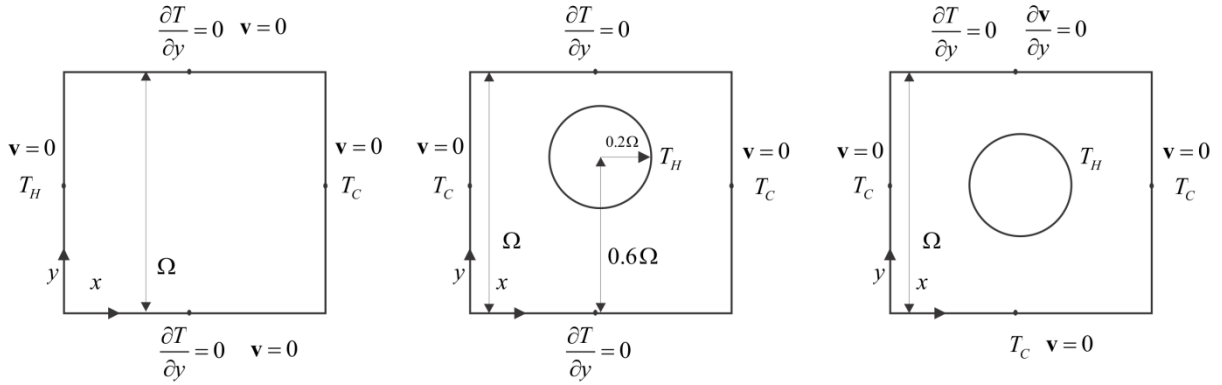


Figure 1: Geometry and boundary conditions for de Vahl Davis case (left), Demirdžić case (middle), and cooling of the overhead power line (right).

Meshless solution procedure

Spatial discretization

Spatial discretization is based on a local approximation of a considered field over the overlapping local support domains, i.e. in each node an approximation over a small local subset of neighbouring n nodes among all nodes N

$$\hat{u}(\mathbf{p}) = \sum_i^m \alpha_i b_i(\mathbf{p}) = \mathbf{b}(\mathbf{p})^T \boldsymbol{\alpha}, \quad (7)$$

with m , $\boldsymbol{\alpha}$, \mathbf{b} , $\mathbf{p}(p_x, p_y)$ standing for the number of basis functions, approximation coefficients, basis functions and the position vector, respectively, is used. Using the same number of basis functions as a number of support domain, i.e. $n = m$, the determination of coefficients $\boldsymbol{\alpha}$ simplifies to solving a system of n linear equations that result from expressing eq. (7) in each support node.

$$u(\mathbf{p}_j) = u = \mathbf{b}(\mathbf{p})^T \boldsymbol{\alpha}, \quad (8)$$

\mathbf{p}_j are positions of support nodes and \mathbf{u} are values of the considered field in the support positions. The above system can be written in a matrix form as

$$\mathbf{u} = \mathbf{B}\boldsymbol{\alpha}, \quad (9)$$

where \mathbf{B} stands for a coefficient matrix with elements $B_{ji} = b_i(\mathbf{p}_j)$. Using a higher number of support nodes than the number of basis functions, i.e. $n > m$, a Weighted Least Squares (WLS) approximation is used to solve the over-determined problem. In other words a norm

$$r^2 = \sum_j^n W(\mathbf{p}_j) (u(\mathbf{p}_j) - \hat{u}(\mathbf{p}_j))^2 = (\mathbf{B}\boldsymbol{\alpha} - \mathbf{u})^T \mathbf{W}(\mathbf{B}\boldsymbol{\alpha} - \mathbf{u}), \quad (10)$$

is minimized, where \mathbf{W} is a diagonal matrix with elements $W_{jj} = W(\mathbf{p}_j)$ with

$$W(\mathbf{p}) = \exp\left(-\left(\frac{\|\mathbf{p}_0 - \mathbf{p}\|}{\sigma p_{\min}}\right)^2\right), \quad (11)$$

where σ stands for weight parameter, \mathbf{p}_0 for the centre of support domain and p_{\min} for the distance to the first support domain node. The solution can be written in matrix form as

$$\boldsymbol{\alpha} = (\mathbf{W}^{0.5}\mathbf{B})^+ \mathbf{W}^{0.5}\mathbf{u}, \quad (12)$$

where $(\mathbf{W}^{0.5}\mathbf{B})^+$ stand for a Moore–Penrose pseudo inverse. By explicit expression of $\boldsymbol{\alpha}$ into (8) an equation

$$\hat{u}(\mathbf{p}) = \mathbf{b}(\mathbf{p})^T (\mathbf{W}^{0.5}(\mathbf{p})\mathbf{B})^+ \mathbf{W}^{0.5}(\mathbf{p})\mathbf{u} = \chi(\mathbf{p})\mathbf{u}, \quad (13)$$

is obtained, where χ stand for the shape function. Now, we can apply partial differential operator, which is our goal, on the trial function,

$$L\hat{u}(\mathbf{p}) = L\chi(\mathbf{p})\mathbf{u}, \quad (14)$$

where L stands for a general differential operator. In this paper, we deal with a Navier-Stokes equation and therefore only shape functions for Laplace operator and first derivatives are needed, which are pre-computed and stored

$$\chi^{\partial x}(\mathbf{p}) = \frac{\partial}{\partial x} \mathbf{b}(\mathbf{p})^T (\mathbf{W}^{0.5}(\mathbf{p})\mathbf{B})^+ \mathbf{W}^{0.5}, \quad (15)$$

$$\chi^{\partial y}(\mathbf{p}) = \frac{\partial}{\partial y} \mathbf{b}(\mathbf{p})^T (\mathbf{W}^{0.5}(\mathbf{p})\mathbf{B})^+ \mathbf{W}^{0.5}, \quad (16)$$

$$\chi^{\nabla^2}(\mathbf{p}) = \nabla^2 \mathbf{b}(\mathbf{p})^T (\mathbf{W}^{0.5}(\mathbf{p})\mathbf{B})^+ \mathbf{W}^{0.5}, \quad (17)$$

The presented formulation is convenient for implementation since most of the complex operations, i.e. finding support nodes and building shape functions, are performed only when nodal topology changes. In the main simulation, the pre-computed shape functions are then convoluted with the vector of field values in the support to evaluate the desired operator. We will refer to this approach as to the Meshless Local Strong Form Method (MLSM) in further discussions.

Positioning of computational nodes

The presented MLSM approach can be understood as a generalisation of FDM. Despite its simplicity it offers many possibilities for treating challenging cases, e.g. nodal adaptivity to address regions with sharp discontinuities or p-adaptivity to treat obscure anomalies in

physical field. Furthermore, the stability versus computation complexity and accuracy can be regulated merely by changing a number of support nodes, etc. In this paper, we will exploit the generality to solve the problem in an irregular domain. Although the above formulation does not need an exact mesh, it is expected that using regularly distributed nodes lead to more accurate and stable results [18-20]. Therefore, despite seeming robustness of meshless methods regarding the nodal distribution, a certain effort has to be invested into the positioning of the nodes [21], with the ultimate goal to maximize stability and accuracy and retain the generality of the meshless principle. A possible approach to achieve that is to distribute nodes with a straightforward algorithm based on Poisson Disc Sampling. Such algorithms have been already used in a meshless context [22] and will also be used here. The general idea is to put a seed node randomly within the domain. Then, add new nodes on a circle with centre in the seed node and radius supplied as a desired nodal density parameter δr where the value of δr represents the desired distance between nodes. In the next iteration one of the newly added nodes is selected as the new seed node and the procedure repeats. Example of nodes positioned within the domain that will be tackled in numerical examples with the Poisson Disk Sampling algorithm is depicted in Figure 2.

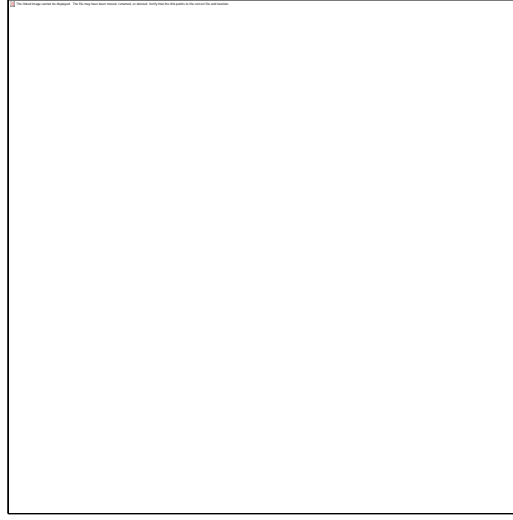


Figure 2: Nodes positioned with a Poisson Disk Sampling algorithm.

Solution procedure

Each time step begins with computing new intermediate velocity from the equation (2) without pressure term with explicit Euler's method. Since the intermediate velocity does not satisfy equation (1), a Poisson pressure correction equation

$$\nabla^2 p^{corr} = \frac{\rho}{\Delta t} \nabla \cdot \mathbf{v}^{iter}, \quad (18)$$

where Δt stand for time step and \mathbf{v}^{iter} for intermediate velocity, is solved with following boundary condition

$$\frac{\Delta t}{\rho} \frac{\partial p^{corr}}{\partial n} = \hat{\mathbf{n}} \cdot (\mathbf{v}^{iter} - \mathbf{v}^{BC}), \quad (19)$$

where $\hat{\mathbf{n}}$ stands for the outside unit normal vector. The pressure Poisson equation is, at given boundary conditions, defined only up to a constant and to avoid instabilities a unique solution is enforced with an additional condition, also referred to as a regularization

$$\int_{\Omega} p d\Omega = 0. \quad (20)$$

Once the pressure correction is known, a velocity is corrected accordingly

$$\mathbf{v}^{corr} = -\frac{\Delta t}{\rho} \nabla P^{corr}. \quad (21)$$

Finally, the equation (3) is solved, again with Euler method.

Results

De Vahl Davis benchmark case

De Vahl Davis investigated natural convection of air with $Pr = 0.71$ up to $Ra = 10^6$ in his original paper. However, more intense solutions followed with the most volatile case of $Ra = 10^8$. Here, we focus only on the $Ra = 10^8$ case. Results are presented in terms of temperature and velocity magnitude contour plots in Figure 3. A more quantitative result is shown in Figure 4, where a hot side Nusselt number and maximal vertical cross-section velocity with respect to the number of nodes used are compared against reference solutions [23] (a), [9] (b) and [10] (c). It can be clearly seen that the solutions of MLSM approach converge towards values that are within the dispersion of the reference solutions and we can conclude that the presented solution procedure can handle intense natural convection cases on regular nodal distributions.

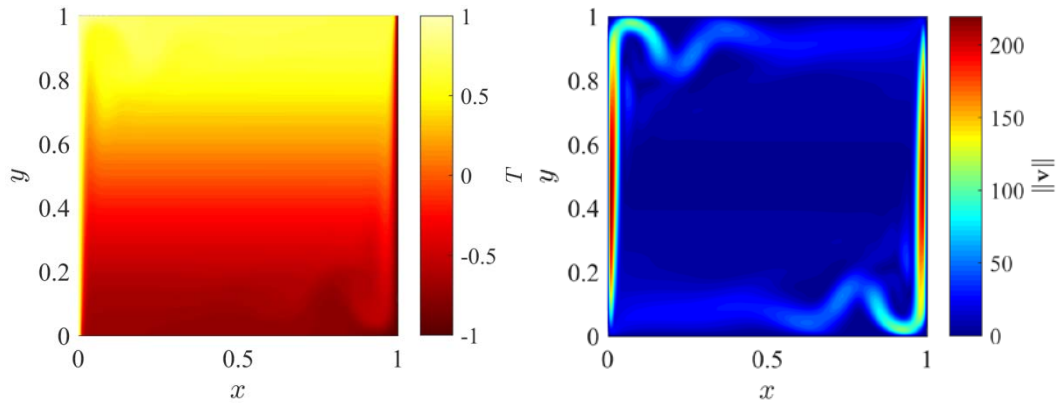


Figure 3: Temperature (left) and velocity magnitude (right) contour plots for de Vahl Davis case at $Ra = 10^8$.

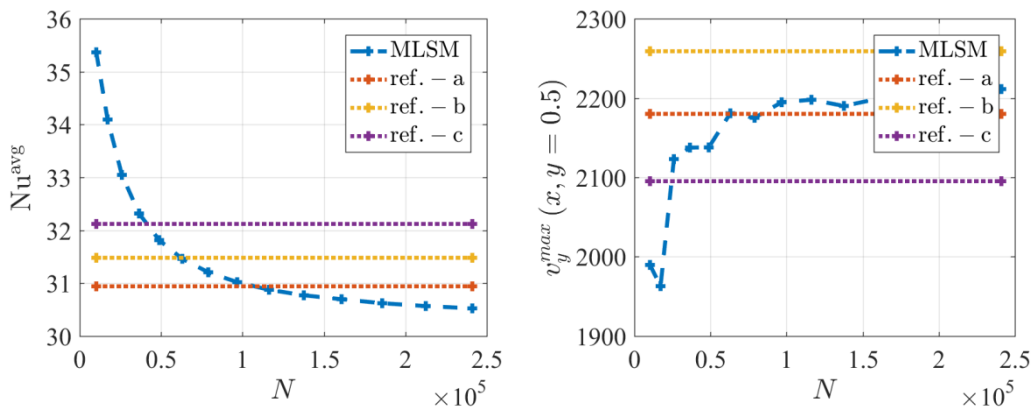


Figure 4: Average hot side Nusselt number (left) and maximal vertical velocity (right) with respect to the number of used nodes N for $Ra = 10^8$ case.

Heated cylinder case benchmark case

In a next case, the natural convection from a cylinder whose wall is maintained at a constant temperature T_H , enclosed by a square duct with vertical walls kept at constant temperature, and horizontal walls assumed adiabatic, is considered. The cylinder centre is displaced from the duct centre vertically for 10%, and its radius is 20% of domain height. The temperature and velocity magnitude contour plots computed by MLSM approach are presented in Figure 5. Furthermore, in Figure 6 MLSM solution is compared against reference data [11] in terms of isotherms and cold side Nusselt number. We can observe that the MLSM solution agrees well with the data provided by Demirdžić, et al.

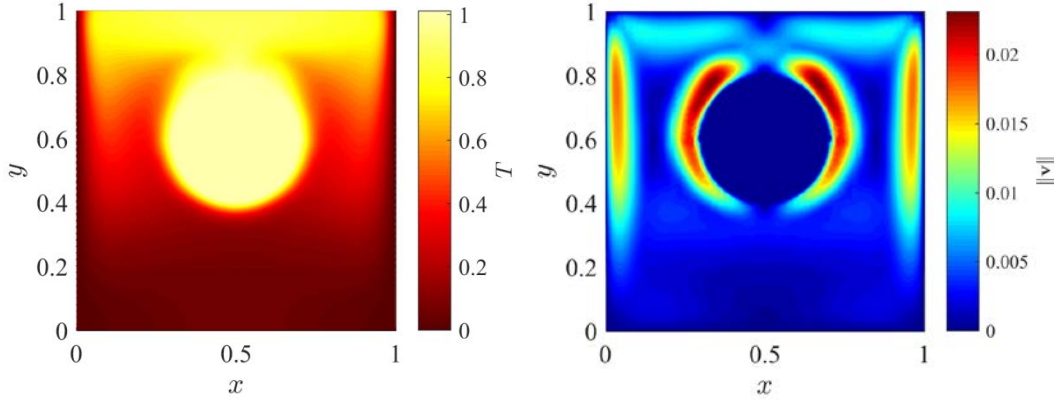


Figure 5: Temperature (left) and velocity magnitude (right) contour plots for Demirdžić case.

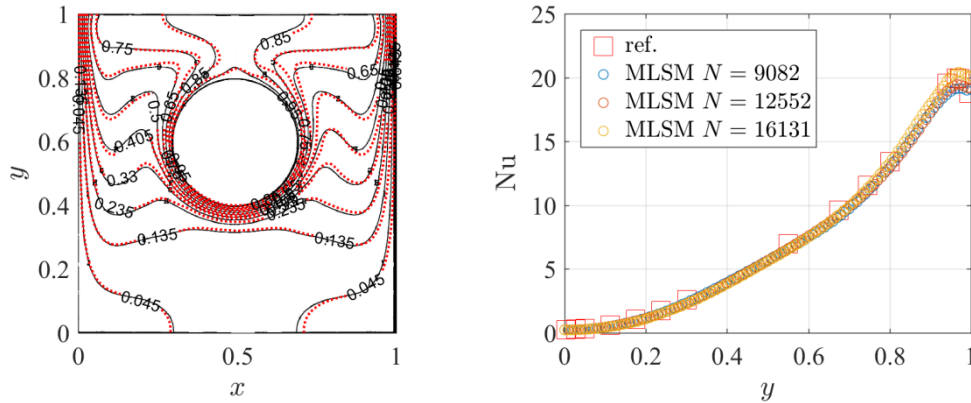


Figure 6: Comparison between MLSM solution and data from [11] in terms of isotherms (left) and cold side Nusselt number (right).

Cooling of overhead power line case

In the last numerical example, we examine a cooling of the overhead power line by natural convection. The problem is very similar to the previous one, with the main difference in a top boundary condition, where instead of confined cavity an open domain is assumed. Besides, the thermo-physical properties of real materials are considered and not dimensionless characterisation as in previous two examples. The power line 490-AL1/64-ST1A with radius 1.33 cm is positioned in the centre of in 5x5 cm square domain, and the air is modelled with following properties: $\rho = 1.29 \text{ kg/m}^3$, $c_p = 1005 \text{ J/kgK}$, $\beta = 0.00367$, thermal conductivity modelled as

$$\lambda_a = 2.368 \cdot 10^{-2} + 7.23 \cdot 10^{-5} T - 2.763 \cdot 10^{-8} T^2 \left[\frac{\text{W}}{\text{mK}} \right], \quad (22)$$

and viscosity modelled as

$$\mu = (17.239 + 4.635 \cdot 10^{-2} T - 2.03 \cdot 10^{-5} T^2) \cdot 10^{-6} [\text{Pa s}]. \quad (23)$$

First, the temperature and velocity magnitude contour plots for $T_C = 0^\circ \text{C}$ $T_H = 80^\circ \text{C}$ are presented in Figure 7. In a next analysis (Figure 8) we compare the power of convective cooling computed by presented MLSM solution procedure against two leading standards, namely CIGRE [4] and IEEE [5]. Although the agreement is not perfect, we are satisfied with the results. It is essential to understand that CIGRE and IEEE computations rely only on empirical relations, while the MLSM solution uses solely physical model and thermo-physical properties of air.

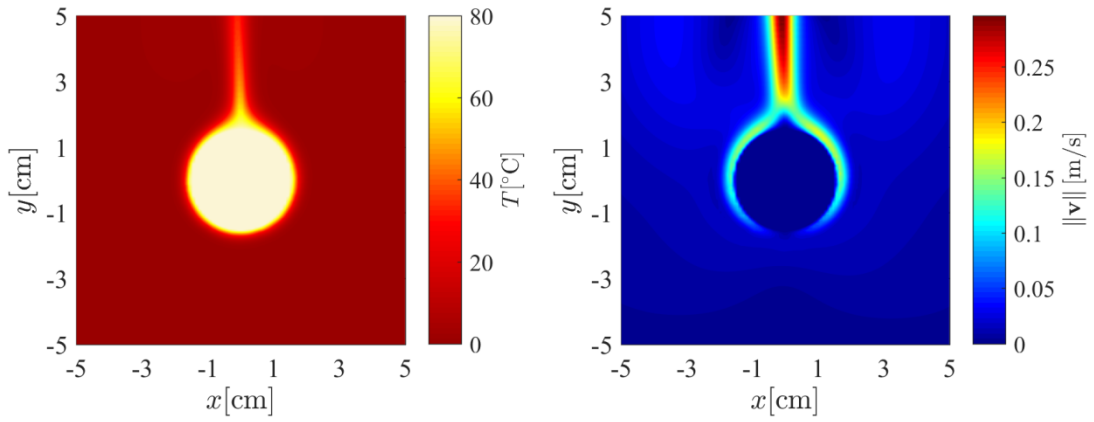


Figure 7: Temperature (left) and velocity magnitude (right) contour plots for cooling of overhead power line case at skin temperature 80°C .

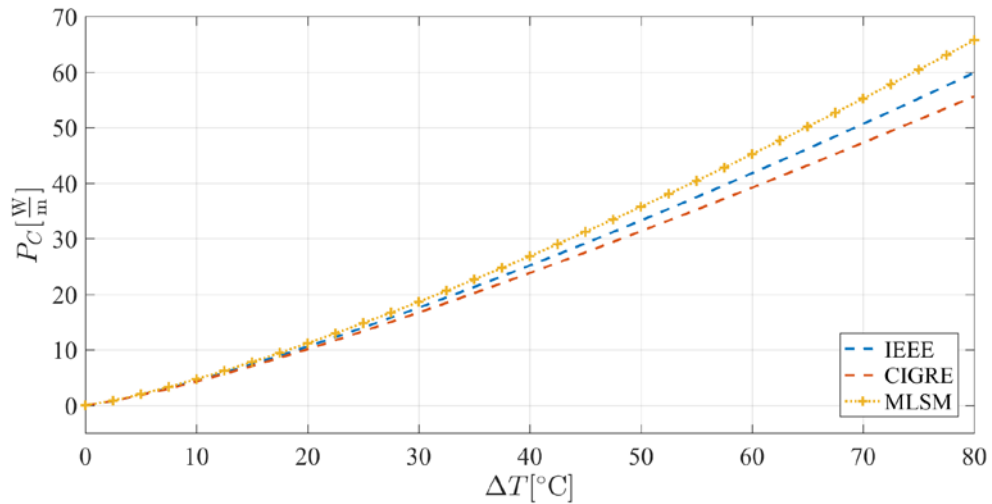


Figure 8: Power of convective cooling with respect to the difference between the skin temperature and ambient temperature ($\Delta T = T_H - T_C$).

Conclusions

In this paper, we demonstrated the usability of the meshless method in solving the natural convection from a heated cylinder. First, we established some confidence in the solution procedure by solving the standard de Vahl Davis case on the regular nodal distributions. Next, we attacked a bit more complicated case of natural convection in the irregular domain. As for the last numerical example, we demonstrated the simulation of cooling of the overhead power line by natural convection.

In all presented cases the results computed with MLSM are in good agreement with the reference data. In the future work we would like to present also a coordinate free implementation of MLSM and compare its performance with more established numerical libraries.

References

- [1] Ayrton, W. & Kilgour, H., The thermal emissivity of thin wires in air. . *Philos TR Soc A*, **183**, pp. 371–405, 1892.
- [2] Boetcher, S. K. S., *Natural Convection from Circular Cylinders*, Springer: Daytona Beach, 2014.
- [3] Morgan, V., The overall convective heat transfer from smooth circular cylinders. *Advances in Heat Transfer*, **1**, pp. 199-264, 1975.
- [4] Guide for Thermal Rating Calculations of Overhead Lines. pp. 2014.
- [5] IEEE Standard for Calculating the Current-Temperature Relationship of Bare Overhead Conductors. pp. 2014.
- [6] Atmane, Mohamed A., Chan, Victor S.S. & Murray, Darina B., Natural convection around a horizontal heated cylinder: The effects of vertical confinement. *International Journal of Heat and Mass Transfer*, **46**, pp. 3661-3672, 2003.
- [7] Cesinia, G. , Paroncinia, M. , Cortellab G. & Manzanb, M., Natural convection from a horizontal cylinder in a rectangular cavity. *International Journal of Heat and Mass Transfer*, **42**, pp. 1801-1811, 1999.
- [8] de Vahl Davis, G., Natural convection of air in a square cavity: a bench mark numerical solution. *International Journal for Numerical Methods in Fluids*, **3**, pp. 249-264, 1983.
- [9] Wan, D. C., Patnaik, B. S. V. & Wei, G. W., A new benchmark quality solution for the buoyancy-driven cavity by discrete singular convolution. *Numerical Heat Transfer*, **B40**, pp. 199-228, 2001.
- [10] Kosec, G. & Šarler, B., Solution of thermo-fluid problems by collocation with local pressure correction. *International Journal of Numerical Methods for Heat & Fluid Flow*, **18**, pp. 868-882, 2008.
- [11] Demirdžić, I., Lilek, Ž & Perić, M., Fluid flow and heat transfer test problems for non-orthogonal grids: Bench-mark solutions. *International Journal for Numerical Methods in Fluids*, **15**, pp. 329–354, 1992.
- [12] Roychowdhury, D. G., Das, Sarit K. & Sundararajan, T., Numerical simulation of natural convective heat transfer and fluid flow around a heated cylinder inside an enclosure. *Heat and Mass Transfer*, **38**, pp. 565-576, 2002.
- [13] Kosec, G., Depolli, M., Rashkovska, A. & Trobec, R., Super linear speedup in a local parallel meshless solution of thermo-fluid problems. *Computers & Structures*, **133**, pp. 30-38, 2014.
- [14] Arefmanesh, A., Najafi, M. & Musavi, S. H., Buoyancy-driven fluid flow and heat transfer in a square cavity with a wavy baffle—Meshless numerical analysis. *Engineering Analysis with Boundary Elements*, **37**, pp. 366–382 2013.
- [15] Wang, C. A., Sadat, H. & Prax, C., A new meshless approach for three dimensional fluid flow and related heat transfer problems. *Computers and Fluids*, **69**, pp. 136-146, 2012.
- [16] Gingold, R. A. & Monaghan, J. J., Smoothed particle hydrodynamics: theory and application to non-spherical stars. *Mon. Not. Roy. Astron. So*, **181**, pp. 375-389, 1977.
- [17] Nguyen, V. P. , Rabczuk, T., Bordas, S. & Duflot, M., Meshless methods: a review and computer implementation aspects. *Mathematics and computers in simulation*, **79**, pp. 763--813, 2008.
- [18] Trobec, R & Kosec, G., *Parallel scientific computing : theory, algorithms, and applications of mesh based and meshless methods*, Springer: 2015.
- [19] Trobec, R., Kosec, G., Šterk, M. & Šarler, B., Comparison of local weak and strong form meshless methods for 2-D diffusion equation. *Engineering Analysis with Boundary Elements*, **36**, pp. 310-321, 2012.

- [20] Amani, J., Afshar, M.H. & Naisipour, M., Mixed discrete least squares meshless method for planar elasticity problems using regular and irregular nodal distributions. *Engineering Analysis with Boundary Elements*, **36**, pp. 894-902, 2012.
- [21] Kosec, G., A local numerical solution of a fluid-flow problem on an irregular domain. *Advances in Engineering Software*, **120**, pp. 36-44, 2018.
- [22] Fornberg, B. & Flyer, N., Fast generation of 2-D node distributions for mesh-free PDE discretizations. *Computers Mathematics with Applications*, **69**, pp. 2015.
- [23] Sadat, H. & Couturier, S., Performance and accuracy of a meshless method for laminar natural convection. *Numerical Heat Transfer*, **B37**, pp. 455-467, 2000.

Highly Accurate Smoothed Finite Element Methods Based on Simplified Eight-noded Hexahedron Elements

Y.H. Li¹, M. Li^{1†}

¹ College of Mathematics, Taiyuan University of Technology, Taiyuan, Shanxi, 030024, China.

[†]Corresponding author: liming01@tyut.edu.cn

Abstract

Compared with tetrahedron elements, hexahedron elements are preferred for its high accuracy. However, coordinate mapping required in the hexahedron elements of FEM formulation costs huge running time, leading to poor performance. Besides, the high quality of Jacobian matrix and mesh is required, which affects the accuracy of the strain results greatly. In order to solve these problems, we propose a novel simplified integration technique based on smoothed finite element method (S-FEM) for the eight-noded hexahedron elements, where coordinate mapping is not demanded. The proposed new S-FEM models include: NS-FEM-H8 (using node-based smoothing domains) and FS-FEM-H8 (using face-based smoothing domains). In the work, we divide a non-triangular face segment of a smoothing domain into two triangular sub-segments, since the strain-displacement matrix can be calculated using a summation in S-FEM theory instead of the integration in FEM. Then we conduct the Gauss integration scheme in each triangular face sub-segment in order to avoid the coordinate mapping in quadrilateral face segments. The rest solving algorithm is the same as the standard S-FEM. Through intensive numerical examples, our simplified S-FEM-H8 is approved to have the following features: (1) the strain energy of simplified NS-FEM-H8 is an upper bound of the exact solution; (2) simplified NS-FEM-H8 can overcome the volume locking problems for incompressible materials.

Keywords: eight-noded hexahedron element, coordinate mapping, a novel simplified integration technique, S-FEM-H8.

1. Introduction

Since the late 1950s, researchers have used the finite element method (FEM) [1] as an essential and important tool for the modeling and simulation of practical problems. However, when using FEM we encounter some problems, such as overly stiff issues and the significant loss of accuracy caused by distortions of the mesh.

In recent years, a new numerical method-smoothed finite element method (S-FEM) [2][3] is presented by G.R. Liu to overcome the disadvantages of the standard FEM, which combines the FEM and the mesh-free technique. S-FEM always produces models that are softer than FEM and even softer than the exact model. Besides, S-FEM can obtain relative accurate strain solutions for the distorted mesh where FEM cannot work.

In practical problems, engineers prefer to use linear elements since the linear elements can be

easily obtained using automatically dividing program. However, the poor accuracy of the strain solution is the biggest problem for linear elements. Although FEM based on higher-order elements, such as quadrilateral element (Q4) and eight-noded hexahedron element (H8), can overcome the shortcoming of poor precision in lower-order elements, the coordinate mapping and the strict requirement of the quality of the mesh lead to the limitation of the high-order elements.

To solve the problems, we present a novel simplified integration technique for eight-noded hexahedron elements in S-FEM to avoid coordinate mapping. The proposed S-FEM-H8 includes simplified NS-FEM-H8 and simplified FS-FEM-H8. In order to compute the smoothed strain-displacement matrix effectively for smoothing domains, a simplified integration technique is also proposed, in which we divide a non-triangular face segment of a smoothing domain into two triangular sub-segments. Therefore, the surface integration in the S-FEM formulation can be performed without coordinate mapping.

2. Three-dimensional Smoothed Finite Elements

2.1 Discretized Linear Algebraic System of Equations

Standard discretized algebraic system of equations

$$\bar{\mathbf{K}}\bar{\mathbf{d}} = \tilde{\mathbf{f}} \quad (1)$$

where $\bar{\mathbf{d}} \in \mathbb{R}_0^{dN_n}$ is the vector of nodal displacements for the all nodes in the S-FEM model, and $\bar{\mathbf{K}}$ is the smoothed stiffness matrix given in the general form of

$$\bar{\mathbf{K}}_{IJ} = \int_{\Omega} \bar{\mathbf{B}}_I^T \mathbf{c} \bar{\mathbf{B}}_J d\Omega = \sum_{k=1}^{N_s} \int_{\Omega_k^s} \bar{\mathbf{B}}_I^T \mathbf{c} \bar{\mathbf{B}}_J d\Omega = \sum_{k=1}^{N_s} \bar{\mathbf{B}}_I^T \mathbf{c} \bar{\mathbf{B}}_J A_k^s \quad (2)$$

the smoothed strain-displacement matrix $\bar{\mathbf{B}}_I$ is computed by Eq. (3) 错误!未找到引用源。

$$\bar{\mathbf{B}}_I = \frac{1}{V_k^s} \int_{\Omega_k^s} \mathbf{L}_n(\mathbf{x}) \mathbf{N}(\mathbf{x}) d\Omega = \begin{bmatrix} \bar{b}_{Ix} & 0 & 0 \\ 0 & \bar{b}_{Iy} & 0 \\ 0 & 0 & \bar{b}_{Iz} \\ 0 & \bar{b}_{Iz} & \bar{b}_{Iy} \\ \bar{b}_{Iz} & 0 & \bar{b}_{Ix} \\ \bar{b}_{Iy} & \bar{b}_{Ix} & 0 \end{bmatrix} \quad (3)$$

with

$$\bar{b}_{Ih} = \frac{1}{V_k^s} \int_{\Omega_k^s} n_h(\mathbf{x}) N_I(\mathbf{x}) d\Omega = \frac{1}{V_k^s} \sum_{p=1}^{n_{\Omega_k^s}} n_{h,p} N_I(\mathbf{x}_p^G) A_p^{surf}, \quad (h = x, y, z) \quad (4)$$

The above surface integration along Ω_k^s can be carried out using the Gauss quadrature technique. Where $n_{\Omega_k^s}$ is the total number of boundary surfaces Ω_k^s and \mathbf{x}_p^G is the Gauss point of the boundary surfaces of Ω_k^s , whose area and outward unit normal are denoted as

A_p^{surf} and $n_{h,p}$, respectively.

When the displacement field along the boundary Ω_k^s is used, four Gauss points are in demand to a quadrilateral surface segment. The quadrilateral segment has in general an arbitrary shape in the physical coordinate system. When the general procedure is used to create the shape functions for Q4 elements, we will meet the problems with the compatibility issue. Thus, the formulation of H8 elements for 3D problems requires a coordinate mapping procedure which is a cost procedure.

However, for triangular surface segment, one Gauss point is sufficient for linear element. And using directly the physical Cartesian coordinate system Oxy, three shape functions have the following forms:

$$N_j = \frac{1}{6A_p^{surf}}(a_j + b_j x + c_j y) \quad (5)$$

in which subscript j varies from 1 to 3, A_p^{surf} is the area of triangular surface segment, and a_j, b_j, c_j are constants. It is clear that the strain-displacement matrix $\bar{\mathbf{B}}_i$ for the above segment is a constant matrix. And any coordinate mapping is not demanded.

Hence, we try to employ the advantage of the triangular surface segment to set up a simplified S-FEM-H8.

2.2 3D Simplified FS-FEM-H8 and NS-FEM-H8

In 3D problems, face-based smoothing domain is the simplest smoothing domain, which is created by joining the center points of the two adjacent elements to the four nodes of the face, shown in Figure 1(a). Node-based smoothing domain is a little complex, which is constructed by successively connecting the middle points of the edges connected to this node and the centroids of the faces containing this node, and the centroids of the faces and the center point of the element, shown in Figure 1(b). Edge-based smoothing domain is formed by sequentially connecting the center point of the element to the two nodes of this edge, the centroids of the faces containing the edge to the two nodes of this edge and the centroids of the faces containing the edge to the center point of the element, shown in Figure 1(c).

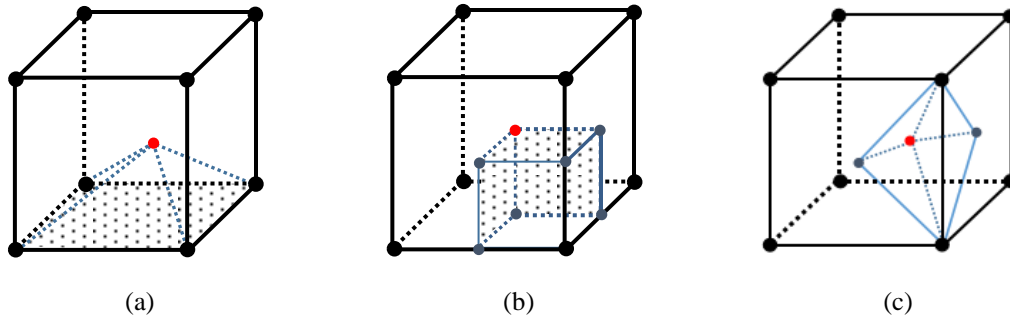


Figure 1 (a) a face-based smoothing domain (b) a node-based smoothing domain (c) a edge-based smoothing domain

From Figure 1, it is easily found that the smoothing domain has quadrilateral surface segments and triangle surface segments. For quadrilateral surface segments, the coordinate mapping is needed to obtain Jacobian matrix when the Gauss integration scheme implements in Eq.

(4). For triangular surface segments, such as the surface segments of edge-based smoothing domain and triangular surface segments of face-based smoothing domain, because the strain-displacement matrix is a constant matrix, no numerical integration is needed to compute the elemental stiffness matrix.

Therefore, in order to avoid the coordinate mapping, we divide a non-triangular surface segment of a smoothing domain into two triangular sub-segments to form simplified FS-FEM-H8 and simplified NS-FEM-H8, shown in Figure 2. For standard ES-FEM-H8, the smoothing surface segments are all triangles, so that we still use the standard ES-FEM-H8.

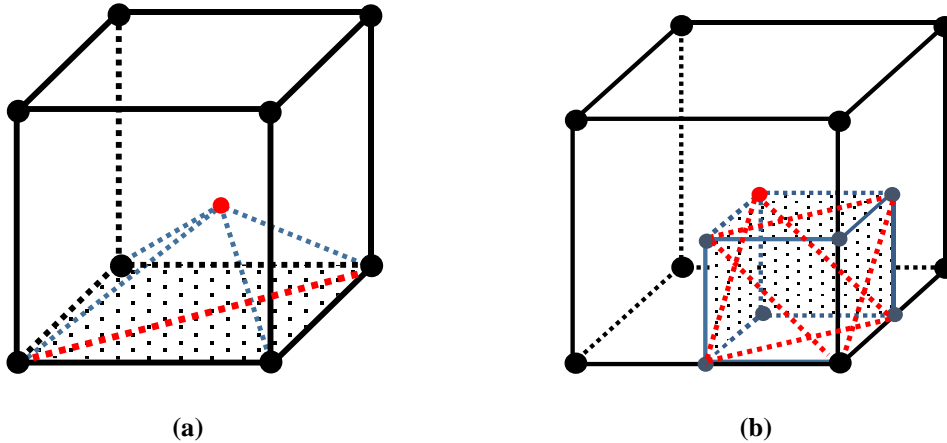


Figure 2 (a) Simplified face-based smoothing domains (FS-FEM-H8) (b) Simplified node-based smoothing domains (NS-FEM-H8)

3. Numerical Example

Consider a 3D Lamé problem consists of a hollow sphere with inner radius $a=1m$, outer radius $b=2m$ and subjected to an internal pressure $P=1 N/m^2$. As the problem is spherically symmetrical, only one-eighth of the sphere model is shown in Figure 3 and symmetry conditions are imposed on the three symmetric planes.

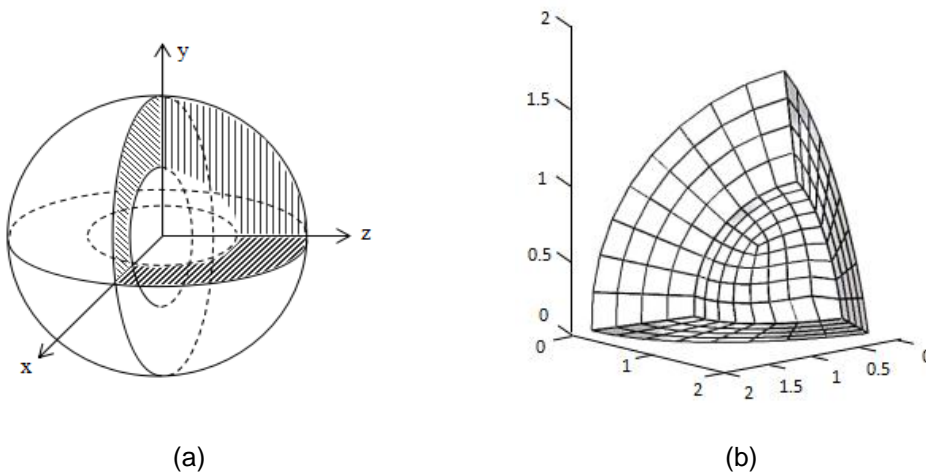


Figure 3 (a) hollow sphere problem domain and (b) one-eighth model discretized using eight-noded hexahedron elements.

Figure 4 shows that the convergence of the strain energy solution obtained using different

methods for the hollow sphere subjected to inner pressure. We can easily find that the strain energy of simplified FS-FEM-H8 and NS-FEM-H8 are both more accurate than that of FEM-H8. Besides, it confirms the upper bound property on the strain energy of simplified NS-FEM-H8 and the lower bound property of simplified FS-FEM-H8 and FEM-H8 for this 3D problem. Also, the distribution of the radial displacement, radial and tangential stresses using FEM-H8, simplified FS-FEM-H8, simplified NS-FEM-H8 and standard ES-FEM-H8 compared with the analytical solution is presented in Figure 5.

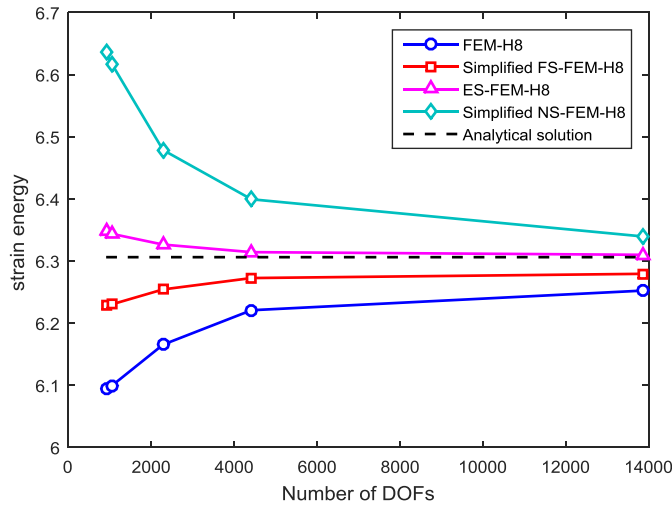


Figure 4 Convergence of the strain energy solution obtained using different methods for the hollow sphere subjected to inner pressure.

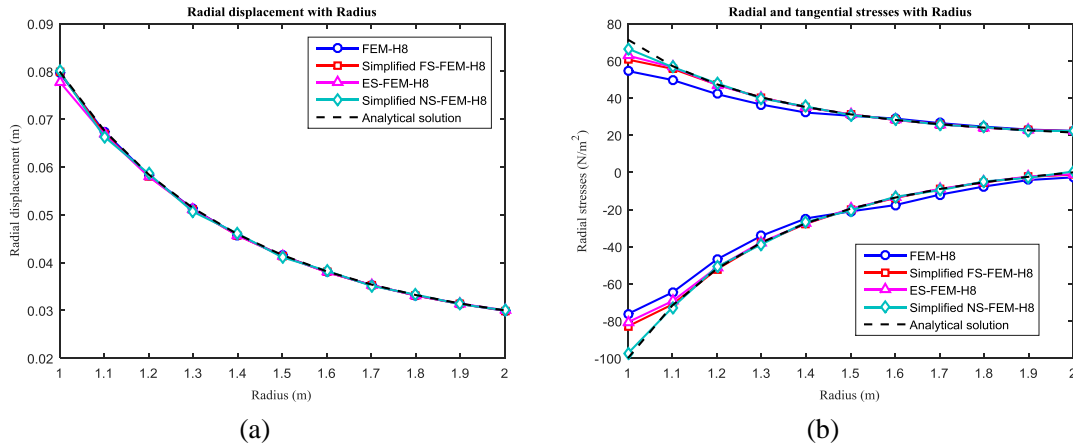


Figure 5 (a) Radial displacement and (b) radial and tangential stresses for the hollow sphere subjected to inner pressure.

Figure 6 plots the error in the displacement norm against Poisson's ratio changing from 0.4 to 0.4999999 obtained using eight-noded hexahedron elements. The results show that simplified NS-FEM-H8 is naturally immune from volumetric locking, while standard FEM-H8 is subjected to volumetric locking, resulting in a drastic accuracy loss in the numerical solutions.

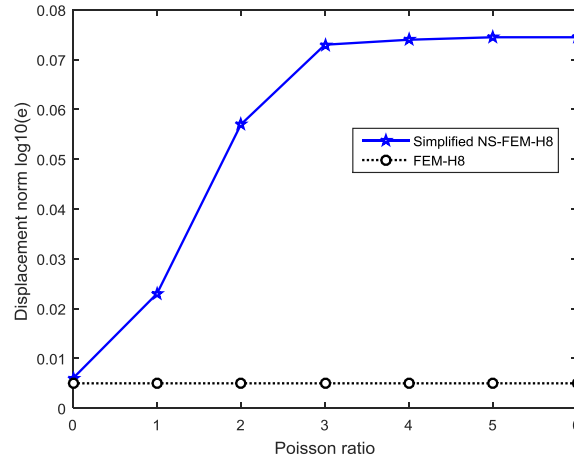


Figure 6 Displacement norm versus different Poisson's ratios for the hollow sphere subjected to inner pressure.

4. Conclusion

In the paper, we put forward a simplified NS-FEM-H8 and ES-FEM-H8 to avoid coordinate mapping for 3D problems, which can improve the efficiency of the algorithm. At the same time, the novel methods can maintain the high accuracy of the higher-order elements. Through numerical examples, some main conclusions are presented as follows:

- The strain energy of simplified NS-FEM-H8 is an upper bound of the exact strain energy, while the solutions of simplified FS-FEM-H8 and standard FEM-H8 are upper bound of the exact strain energy.
- For the almost incompressible problems, simplified NS-FEM-H8 has the characteristic of simplified integration that is free from volume locking.

References

- [1] Liu GR and Quek SS. 2002. The Finite Element Method: A Practical Course. Butterworth Heinemann, Oxford.
- [2] Nguyen-Thoi T, Liu GR, Dai KY, and Lam KY. 2007. Selective smoothed finite element method. *Tsinghua Science and Technology*; 12(5): 497–508.
- [3] Dai KY and Liu GR. 2007. Free and forced analysis using the smoothed finite element method (SFEM). *Journal of Sound and Vibration*; 301: 803–820.

A robust inversion-based Fourier Transformation algorithm used in the interpretation of non-equidistantly measured magnetic data

†Mihaly Dobroka^{1,2}, Daniel Nuamah¹, J. Somogyi-Molnar² and Tamas Ormos¹

¹Department of Geophysics, University of Miskolc, Hungary.

²MTA-ME, Geoengineering Research Group, H-3515, Miskolc-Egyetemváros, Hungary

†Corresponding author: dobroka@uni-miskolc.hu

Abstract

Fourier transformation is one of the most frequently used operation in data processing. In case of discrete data sets the Discrete Fourier Transformation (DFT) algorithm is often applied. As the measured data always contain noise, the noise sensitivity of the processing methods is an important feature. The noise registered in the time domain is directly transformed into the frequency domain. Therefore the traditional discrete variants of Fourier transformation are very noise sensitive procedures. In the field of inverse problem theory a variety of numerous procedures are available for noise rejection, so if the Fourier transformation is formulated as an inverse problem these tools can be used to reduce the noise sensitivity.

It is well-known from inverse problem theory that simple least square methods give optimal results only when data noises follow Gaussian distribution. The practice of geophysical inversion shows that the least square solutions are very sensitive to sparsely distributed large errors, i.e. outliers in the data set and the estimated model parameters may even be completely non-physical. There are various ways to address the question of statistical robustness: the Least Absolute Deviation (LAD) method (minimizing the L1 norm of the misfit between the observed and predicted data) and the Cauchy weighted Iteratively Reweighted Least Squares method are well-known. A more flexible method can be defined by modifying the weights with the help of Steiner's Most Frequent Value method. In the present paper the 2D Fourier transformation is handled as robust inverse problem using IRLS algorithm with Cauchy-Steiner weights. The discretization of the continuous Fourier spectra is given by a series expansion with the scaled Hermite functions as basis functions. The expansion coefficients are determined by solving an overdetermined inverse problem. In order to define a quick algorithm in calculating the Jacobi matrix of the problem, the special feature that the Hermite functions are eigenfunctions of the Fourier transformation was used. It is shown that the method can successfully be applied in the interpretation of geomagnetic data sets measured in 2D arrays. There is an important feature of the new inversion based Fourier Transformation: the measurement array should not be equidistant along the (x,y) directions. It will be shown in the presentation that the method gives accurate results even in the interpretation of geomagnetic data measured also in "random walk" measurement array.

Key words: Fourier Transformation, Inversion, Series Expansion, Cauchy Noise, Random walk measurement.

Introduction

The systematic improvement in geophysical data acquisition over the years demand for more innovative data processing methods. Traditional survey designs employ equidistant measurement on a regular grid. Unfortunately, measurements are sometimes taken out of grid due to several obstacles encountered in the field of survey. This has necessitated the development of methods for the effective processing of equal in random-walk geophysical measurements. Fourier transformation is one of the most frequently

used operation in data processing. It connects the time domain of signal registration and the frequency domain of signal processing. In case of discrete data sets, the DFT algorithm is often applied (Discrete Fourier Transformation). As the measured data always contain noise, the noise sensitivity of the processing methods is an important feature. The noise registered in the time domain is directly transformed into the frequency domain. Therefore, the traditional discrete variants of Fourier transformation are very noise sensitive procedures. Dobróka et al. (2015) presented an inversion based 1D Fourier transformation method (S-IRLS-FT) which proved to be an effective tool for noise reduction. The method was generalized to 2D, and an application is presented in solving reduction to the pole of the magnetic data set (Dobróka et al., 2017). In this paper, it is shown that the newly developed inversion-based Fourier transformation algorithm can also be used in processing non-equidistant (even in random walk) measurement geometry dataset.

Method Development (2D inversion-based Fourier transformation)

The 2D Fourier transform of a function $u(x,y)$ can be calculated by the integral

$$U(\omega_x, \omega_y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(x, y) e^{-j(\omega_x x + \omega_y y)} dx dy, \quad (1)$$

its inverse is given by the formula

$$u(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\omega_x, \omega_y) e^{j(\omega_x x + \omega_y y)} d\omega_x d\omega_y, \quad (2)$$

where x, y are the spatial coordinates, $U(\omega_x, \omega_y)$ means the 2D spatial-frequency spectrum and ω_x, ω_y indicate the spatial-angular frequencies. Following a powerful inversion strategy (the so-called series expansion-based optimization) developed at the Geophysical Department of the University of Miskolc (Gyulai and Szabó 2014, Dobróka et al. 2015), the discretization of the continuous spectrum is written in the form of a series expansion,

$$U(\omega_x, \omega_y) = \sum_{n=1}^N \sum_{m=1}^M B_{n,m} \Psi_{n,m}(\omega_x, \omega_y), \quad (3)$$

where $\Psi_{n,m}(\omega_x, \omega_y)$ are frequency dependent basis functions, $B_{n,m}$ are the expansion coefficients which represent the model parameters of the inverse problem. The basis function system should be square integrable in the interval $(-\infty, \infty)$. The Hermite functions meet this criterion with an additional advantage. The determination of Jacobian matrix in an inversion process requires the calculation of a complex integral for the interval $(-\infty, \infty)$, which is usually a time-consuming process. As it was derived by Dobróka et al. (2015) the elements of the Jacobian matrix can be considered as the inverse Fourier transformation of the basis function system. Therefore, they can be calculated more easily if the basis functions are chosen from the eigenfunctions of the inverse Fourier transformation. It can be shown, that the normed and scaled Hermite functions

$$H_n(\omega_x, \alpha) = \frac{e^{-\frac{\alpha \omega_x^2}{2}} h_n(\omega_x, \alpha)}{\sqrt{\frac{\pi}{\alpha} n! (2\alpha)^n}}, \quad \text{where } h_n(\omega_x, \alpha) = (-1)^n e^{\alpha \omega_x^2} \left(\frac{d}{d\omega_x} \right)^n e^{-\alpha \omega_x^2}, \quad (4)$$

$$H_m(\omega_y, \beta) = \frac{e^{-\frac{\beta \omega_y^2}{2}} h_m(\omega_y, \beta)}{\sqrt{\frac{\pi}{\beta} m! (2\beta)^m}}, \quad \text{where } h_m(\omega_y, \beta) = (-1)^m e^{\beta \omega_y^2} \left(\frac{d}{d\omega_y} \right)^m e^{-\beta \omega_y^2}, \quad (5)$$

are eigenfunctions of the inverse Fourier transformation and the Jacobian matrix of the inverse problem can be written as

$$G_{k,l}^{n,m} = \frac{(j)^{n+m}}{\sqrt[4]{\alpha\beta}} H_n^{(0)}\left(\frac{x_k}{\sqrt{\alpha}}\right) H_m^{(0)}\left(\frac{y_l}{\sqrt{\beta}}\right). \quad (6)$$

Here $H_n^{(0)}, H_m^{(0)}$ denote the non-scaled Hermite functions. Applying Eq. (6) provides a fast solution to the forward problem

$$u(x_k, y_l) = \sum_{n=1}^N \sum_{m=1}^M B_{n,m} G_{k,l}^{n,m}. \quad (7)$$

The introduction of single indices instead of double ones ($s=k+(l-1)K$, $i=n+(m-1)N$) makes the calculations much easier

$$u(x_k, y_l) = u_{k,l} = u_s, \quad B_{nm} = B_i, \quad u_s = \sum_{i=1}^I B_i G_{s,i}, \quad (i=1, \dots, I, \quad s=1, \dots, S), \quad (8)$$

where $I=N+(M-1)N=NM$ is the number of series expansion coefficients, $S=K+(L-1)N=KL$ is the maximal number of data. With these simplifications, the deviation of measured and calculated data (e_s) can be calculated as

$$e_s = u_s^{(measured)} - \sum_{i=1}^I B_i G_{s,i}. \quad (9)$$

The Iteratively Reweighted Least Squares (IRLS) method is a reliable data processing procedure if the data set contains outliers. It can be combined with Cauchy weights, where the scale parameters σ^2 have to be known a priori. Using the Most Frequent Value method (Steiner 1997), Szegedi and Dobróka (2014) proposed the use of Steiner weights because the scale parameters are derived from the statistical sample in an inner iteration cycle. To make the Fourier transformation more robust, a new Steiner-weighted IRLS method (S-IRLS) was introduced where the following weighted norm is minimized

$$E_w = \sum_{s=1}^S W_{ss} e_s^2. \quad (10)$$

where W_{ss} is the s -th Steiner weight. The normal equation for the j -th IRLS step can be written as

$$\underline{\underline{\mathbf{G}}}^T \underline{\underline{\mathbf{W}}}^{(j-1)} \underline{\underline{\mathbf{G}}} \underline{\underline{\mathbf{B}}}^j = \underline{\underline{\mathbf{G}}}^T \underline{\underline{\mathbf{W}}}^{(j-1)} \underline{\underline{\mathbf{u}}}^{(measured)}. \quad (11)$$

After reaching the stop criteria, the series expansion coefficients determined in the last iteration step are considered the solution of the problem.

Application

To prove the applicability of the 2D robust inversion based S-IRLS FT method, it was tested on synthetic magnetic data sets: one was noise-free and the other contained random noise following Cauchy distribution. 1089 measurement points were assumed along with a 5 m x 5 m grid which was further randomized to obtain non-equidistant measurements. Data were generated for a surface between -100 m, and 100 m both in the x and y directions above a 'C_L' shaped magnetic body (inclination $I=63^\circ$, declination $D=3^\circ$, magnetization 200 nT). The surface magnetic data were calculated by the Kunaratnam (1981) method and was subsequently reduced to the pole ($I=90^\circ$) by applying the formula in the frequency domain

$$R(u, v) = T(u, v) S(u, v), \quad (12)$$

where $T(u, v)$ is the 2D Fourier transform of the magnetic data set, $S(u, v)$ is the frequency domain operator of pole reduction and $R(u, v)$ is the reduced data set after the data reduction process. First, the reduction to the pole was performed by using conventional DFT algorithm. The map of noiseless magnetic data on equidistant grid and its reduced to pole are given in Fig. 1a and Fig.1b, respectively. The data was contaminated with Cauchy noise to produce Figure 2a which was reduced to pole using

the traditional DFT method. The poor noise reduction capability of the DFT method can be clearly seen in Figure 2b. This underlines the need for a new, robust and outlier-resistant Fourier Transformation method.

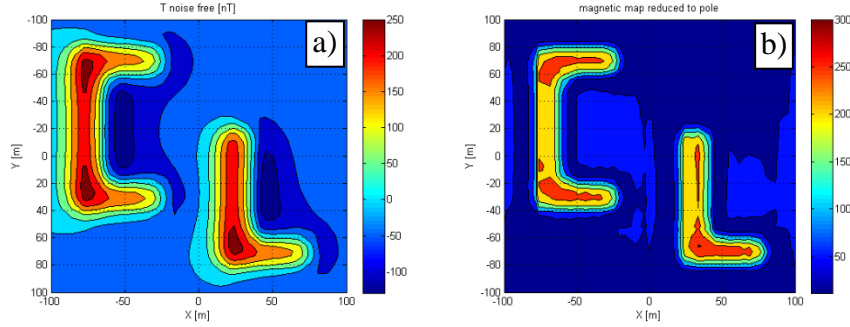


Figure 1 Noise-free data set and the conventional DFT. a) Magnetic map without pole reduction. b) Magnetic map reduced to the pole using conventional DFT.

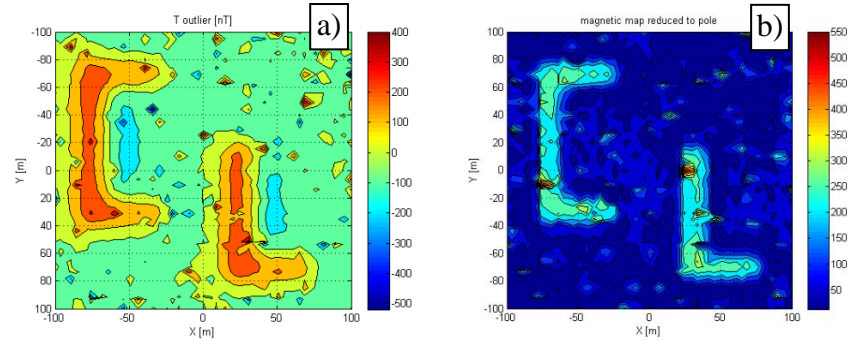


Figure 2 Data set with Cauchy noise and the conventional DFT. a) Magnetic map without pole reduction. b) Magnetic map reduced to the pole using conventional DFT.

Therefore, the magnetic data were processed using the new S-IRLS-FT algorithm instead of traditional DFT. The same map of noisy data (Fig. 3a) and its processing result are presented (Fig. 3b). Comparing Fig. 2b (the pole reduced map on noisy data set using conventional DFT) and Fig. 3b (the pole reduced map on noisy data set using S-IRLS-FT), sufficient improvement can be seen in the processed data using S-IRLS-FT method. The traditional DFT left more spikes of Cauchy noise after the pole reduction causing artifacts and a possible misinterpretation of magnetic anomalies. The high noise reduction capability of the new S-IRLS-FT algorithm is clearly observable.

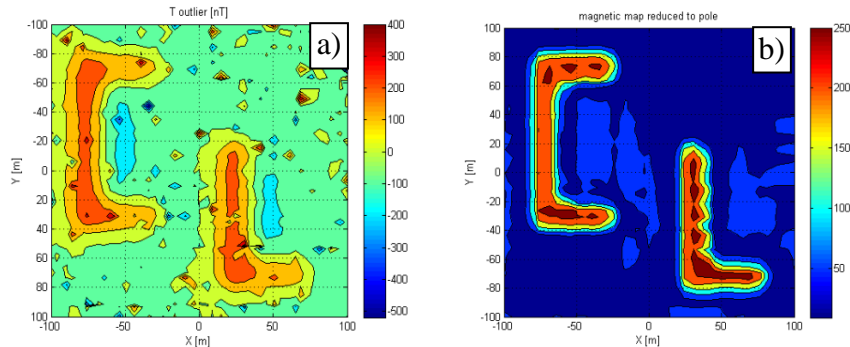


Figure 3 Data set with Cauchy noise and the new S-IRLS-FT. a) Magnetic map without pole reduction. b) Magnetic map reduced to the pole using the new S-IRLS-FT.

As an important feature of the new inversion-based Fourier Transformation, the measurement array should not be equidistant along the (x,y) directions. This is demonstrated in Fig. 4a and 4b. In Fig. 4a, (produced from Fig. 1a and same as Fig. 1b) the arrangement of the measurement points was equidistant. To generate a highly non-equidistant arrangement, we shifted randomly all the measurement point from

their “regular” place in the sampling interval before the S-IRLS FT method was applied. The result in the noise-free case is shown in Fig. 4b., which is similar to Fig. 4a. demonstrating that the method gives accurate results even in “random walk” measurements.

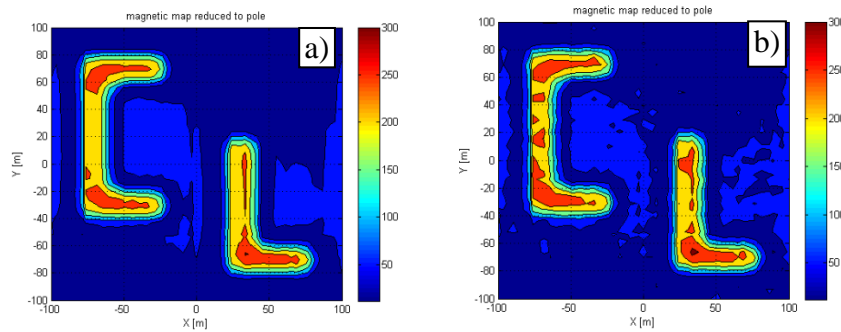


Figure 4 Reduction to the pole of noise-free data set. a) Equidistant measurement system. b) Random walk measurement system.

Conclusion

The Discrete Fourier Transformation (DFT) is a common data processing method but incorporate some level of noise in the transformation process. The introduced 2D S-IRLS-FT method treats the Fourier transformation as an inverse problem. The spectrum is discretized by series expansion and the inversion problem is solved for the series expansion coefficients by IRLS method using Steiner weights. Taking advantage of the good features of Hermite functions described in this paper, they were chosen as basis functions making the algorithm quicker. Comparatively, the newly introduced inversion-based Fourier transformation algorithm (2D S-IRLS-FT method) has a higher noise rejection capability than the traditional DFT Method as demonstrated in reduction to pole of magnetic data. In this paper, it was further shown that the inversion-based Fourier transformation algorithm can be effectively used in processing data set collected in non-equidistant (even in random walk) measurement geometry.

Acknowledgements

The study was supported by the National Research Development and Innovation Office (project No. K109441). The research was partly carried out within the GINOP-2.3.2-15-2016-00031 “Innovative solutions for sustainable groundwater resource management” project of the Faculty of Earth Science and Engineering of the University of Miskolc in the framework of the Széchenyi 2020 Plan, funded by the European Union, co-financed by the European Structural and Investment Funds.

References

- Dobróka, M., Szegedi, H., Somogyi Molnár, J. Szűcs, P. [2015] On the Reduced Noise Sensitivity of a New Fourier Transformation Algorithm. *Math. Geos.* **47**:(6) pp. 679-697.
- Dobróka M, Szegedi H, and Vass P. [2017] Inversion-based Fourier transform as a new tool for noise rejection, INTECH 2017, DOI:10.5772/66338
- Gyulai, Á., Szabó, N. P. [2014] Series expansion based geoelectric inversion methodology used for geo-environmental investigations. *Frontiers in Geosc.* **2**:(1) pp. 11-17.
- Kunaratnam, K. [1981] Simplified expressions for the magnetic anomalies due to vertical rectangular prisms. *Geophys. Prosp.* **29**. pp. 883-890.
- Steiner, F. [1997] *Optimum methods in statistics*. Akadémiai Kiadó, Budapest.
- Szegedi, H., Dobróka, M. [2014] On the use of Steiner's weights in inversion-based Fourier transformation - robustification of a previously published algorithm. *Acta Geodaetica et Geophysica*, **49**/1, 95-104, DOI 10.1007/s40328-014-0041-0.

Solving the singular Motz problem using Radial Basis functions

*†T. S. Li¹ and S. M. Wong¹

¹School of Science and Technology, the Open University of Hong Kong

*Presenting author: tsli@ouhk.edu.hk

†Corresponding author: tsli@ouhk.edu.hk

Abstract

Mathematical model has been used extensively in solving engineering problems with singularity. This paper introduces a mesh-free numerical scheme for solving problems with boundary singularity. The solution of the governing equation is approximated by a class of mesh-free radial basis functions. The proposed radial basis function is a continuously differentiable, positive definite and integrable function, it can easily be used to solve higher order of differential equations. In the vicinity of the singular point, we use a series to approximate the solution. Then domain decomposition is used to blend the two solutions together.

Keywords: Radial basis functions, singular problem, domain decomposition.

Introduction

The behaviour of many fatigue and fracture mechanics can be modelled as problems with boundary singularity. Studies of boundary singularity problems can be traced back to Motz [1] in 1946. The author adopted the classical finite difference scheme together with the relaxation method to overcome the discontinuity occurring from the crack tip over the boundary interface. The solution to the Motz's problem is singular at the origin and is often used by many researchers as reference example for testing numerical methods. The Motz's problem has the form $\nabla^2 u = -f$, where f is a specific function. In the vicinity of the singular point, the singular solutions u can be written as $u = \hat{u}_h + u_p$, where \hat{u}_h is the homogeneous solution and u_p is the particular solution. For the two-dimensional Laplace equations, the homogeneous solution \hat{u}_h in the vicinity of the singular points can be found in [2] and is given by the asymptotic series of the form

$$u_h = \sum_{i=1}^{\infty} A_i r^{\alpha_i} f_i(\theta), \quad r, \theta \in \Omega$$

over a connected region Ω , where A_i are the unknown expansion coefficients to be determined. These coefficients are termed as generalized flux intensity factors (GFIFs). The polar coordinates (r, θ) centred at the singular point and (α_i, f_i) associated with eigen-pairs, where α_i is the power arranged in ascending order as defined by $\alpha_i \leq \alpha_{i+1}$.

Yosibash *et al* [3, 4] in 1997 developed a mesh-dependent method derived from the least-square finite difference scheme for approximating solutions in the vicinity of the singular point and a conventional finite difference scheme was then applied to the remaining part of the given problem. Another comprehensive study on crack tip analysis of singularity problems was carried out by Li [5] using conformal mappings and several types of combined methods.

Recently Rao *et al* [6] developed an element-free Galerkin method (EFGM) for fracture analysis of cracks. However, the mathematical formulation is rather complicated and consists of three components: (i) moving least-squares approximation; (ii) choosing the weight functions and (iii) variational formulation and discretization. This paper introduces an efficient mesh-free numerical scheme which is derived from a class of radial basis functions (RBF). The RBF method possess a simple mathematical formulation and a truly mesh free property, which does

not require a global mesh for supporting computations. In addition, RBF are continuously differentiable and integrable, and is insensitive to dimension d . These features make it suitable for solving problems in higher dimensions with unsmooth boundary conditions.

Meshless Radial Basis Function Method for Solving PDEs

This paper discusses a mesh free approximation scheme based on the radial basis function for solving the problems with complex boundary conditions and singularities. The RBF methods have been found to have major advantages over the classical finite element or finite difference methods. One of these advantages is that it does not require the construction of an underlying mesh. This allows it to handle complicated boundaries with concave surface more efficiently. The basic concept of the RBF method is described below.

The RBFs were originally devised for scattered geographical data interpolation by Hardy [7], who introduced a class of functions called multiquadric functions in the early 1970's. The basic idea of the RBF interpolation is to approximate an unknown function, $\{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^d\}$ by an interpolant, say $\{\hat{f}(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^d\}$ at a set of N distinct data points $X = \{\mathbf{x}_j : j = 1, 2, \dots, N\}$. Let $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}$ be a set of positive definite basis functions defined by

$$\Phi = \{\phi(\|\mathbf{x} - \mathbf{x}_j\|)\}, \quad \mathbf{x}, \mathbf{x}_j \in \mathbb{R}^d$$

on a fixed space \mathbb{R}^d . Here $\phi(\|\mathbf{x} - \mathbf{x}_j\|)$ refers to a typical type of RBFs that is solely dependent on the Euclidean distance between \mathbf{x} and a fixed point $\mathbf{x}_j \in \mathbb{R}^d$. The RBF interpolant to the approximated solution of $f(\mathbf{x})$ can be expressed as a finite linear combination of $\phi(\|\mathbf{x} - \mathbf{x}_j\|)$:

$$\hat{f}(\mathbf{x}) = \sum_{j=1}^N \alpha_j \phi(\|\mathbf{x} - \mathbf{x}_j\|), \quad \mathbf{x}, \mathbf{x}_j \in \mathbb{R}^d, \quad (1)$$

where $\{\alpha_j : j = 1, 2, \dots, N\}$ are the unknown coefficients, which can be determined by setting the following condition:

$$\hat{f}(\mathbf{x}_i) = f(\mathbf{x}_i), \quad i = 1, 2, \dots, N. \quad (2)$$

This yields a system of linear equations, which can be expressed in the following matrix form

$$[\mathbf{A}_\phi] \vec{\alpha} = \vec{\mathbf{F}}, \quad (3)$$

where $[\mathbf{A}_\phi] = [\phi(\mathbf{x}_i - \mathbf{x}_j)]_{1 \leq i, j \leq N}$ is an $N \times N$ matrix, $\vec{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_N]^T$ and $\vec{\mathbf{F}} = [f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_N)]^T$ are $N \times 1$ column matrices. Provided that the chosen radial basis function $\phi \in \mathbb{R}^d$ is positive definite, the matrix $[\mathbf{A}_\phi]$ is non-singular so the linear system (3) has a unique solution. The unknown coefficients $\{\alpha_j\}$ can be obtained uniquely by solving the linear system (3).

Although the above-mentioned condition guarantees the uniqueness of some particular RBF interpolants, not all RBFs can satisfy the conditions of positive definiteness. A general theory on the existence, uniqueness and convergence of the RBFs interpolation was proven by Micchelli [8] in 1986. In accordance with the Micchelli's result, Powell [9], Madych *et al* [10] and Wu *et al* [11] extended the study and deduced some important non-singularity properties of the RBF interpolation. Their analysis concluded that the RBF interpolation method possess a super-convergent property and truly mesh-free algorithm. The RBF method has been demonstrated to be highly flexible for the approximation of high spatial dimensional problems. The accuracy of the RBF interpolant has an order of convergence $\mathcal{O}(h^{d+1})$, where h is the density of the

collocation points and d is the spatial dimension.

Many of RBF ideas can be easily generalized to the case where the basis function ϕ is only conditionally positive definite [12] in which one needs to add a finite number of polynomial of suitable degree to the interpolant $\hat{f}(\mathbf{x})$ in equation (1) and impose additional conditions to accomplish its uniqueness. Let $Q_m^d(\mathbf{x}) \in \Pi_m$ where Π_m is a set of d -variate polynomials of degree less than m . The RBF interpolant $\hat{y}(\mathbf{x})$ is now written as

$$\hat{f}(\mathbf{x}) = \sum_{j=1}^N \alpha_j \phi(\|\mathbf{x} - \mathbf{x}_j\|) + Q_m^d(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \quad 0 < m < N, \quad (4)$$

where

$$Q_m^d(\mathbf{x}) = \sum_{k=1}^L b_k p_k(\mathbf{x}), \quad L = \frac{(m+d-1)!}{(m-1)!d!}.$$

The terms $\{p_k(\mathbf{x}) \mid k = 1, 2, \dots, L\}$ are the basis of $Q_m^d(\mathbf{x})$. The approximation function of (4) has a unique solution if the system satisfies the conditions (2) and the following constraints

$$\sum_{j=1}^N \alpha_j p_k(\mathbf{x}_j) = 0, \quad k = 1, 2, \dots, L. \quad (5)$$

Note that, in this case, the matrix $[\mathbf{A}_\phi]$ is enlarged to order $(N+L) \times (N+L)$, and $\vec{\alpha}$ and $\vec{\mathbf{Y}}$ are $(N+L) \times 1$ column matrices. Although there are many possible radial basis functions, the followings are the most popular choices:

$$\phi(r_j) = \begin{cases} r_j^3, & \text{Cubic} & (a) \\ (r_j^2) \log r_j, & \text{Thin plate splines in } R^2 & (b) \\ e^{-\sigma r_j^2}, & \text{Gaussian, } \sigma > 0 & (c) \\ (r_j^2 + \delta^2)^{\frac{1}{2}}, & \text{Multiquadric, } \delta \in R & (d) \\ (r_j^2 + \delta^2)^{-\frac{1}{2}}, & \text{Reciprocal multiquadric, } \delta \in R & (e) \end{cases} \quad (6)$$

where $\{r_j = \|\mathbf{x} - \mathbf{x}_j\| \mid j = 1, 2, \dots, N\}$ is the Euclidean distance between \mathbf{x} and $\mathbf{x}_j \in \mathbb{R}^d$ and $\delta^2 \in \mathbb{R}$ is the shape parameter of the multiquadric functions in (d) & (e), which is used to control the fitting of a smooth surface to the data. These functions are globally supported and will generate a system of equations with a full matrix. However, as shown by Madych and Nelson [13], the multiquadric function (MQ-RBF) can be exponentially convergent so we can often use a relatively small number of basis elements to achieve a computational efficiency. As a consequence, the MQ-RBF method has been the most commonly used has been radial basis function and progressively refined recently by Kansa [14] and widely used by Hon *et al* [15] and Wong *et al* [16] to solve scientific and engineering problems. Their results from solving elliptic, parabolic and hyperbolic problems were shown to be better than other well established approximation methods.

The Algorithm

To study the performance of the proposed method, we apply it to solve a classical re-entrant corner problem. Re-entrant corner problem possesses a typical nature of singularity of solution, the singular point occurs at the origin forming an angle of $\gamma\pi$ which would result discontinuity. The model involves the Laplace equation satisfying some mixed Neumann and Dirichlet

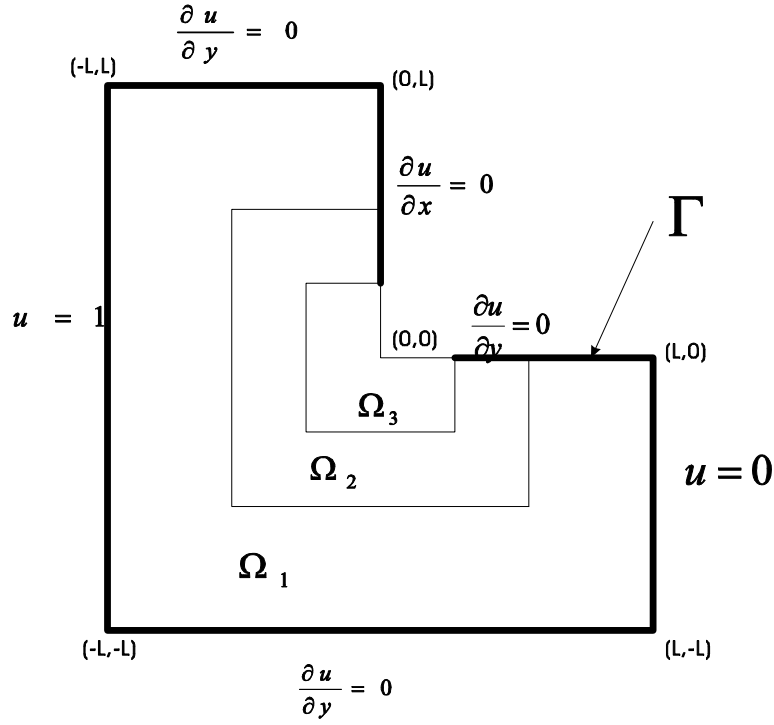


Figure 1: A region with an re-entrance angle.

boundary conditions. The re-entrant corner with the L-shaped domain as depicted in Figure (1), is a special case in which, $\gamma = \frac{3}{2}\pi$ at the origin.

The governing equation is:

$$\nabla^2 f = 0 \quad (7)$$

and the boundary conditions are depicted in the figure. We then divide the region into three as shown in the figure. Ω_1 is the region that is far away from the corner and Ω_3 is the region close to the corner. Ω_2 is the region between Ω_1 and Ω_3 . We are going to use a radial basis function to approximate the solution in Ω_1 and Ω_2 and the series solution to approximate the solution in Ω_2 and Ω_3 .

Let $X_1 = \{\mathbf{x}_i | i = 1, \dots, N_1\}$ be nodes in Ω_1 , $X_2 = \{\mathbf{x}_i | i = N_1 + 1, \dots, N_1 + N_2\}$ be nodes on the boundary Γ in the figure, $X_3 = \{\mathbf{x}_i | i = N_1 + N_2 + 1, \dots, N_1 + N_2 + N_3\}$ be nodes in Ω_2 . We now use these $N (= N_1 + N_2 + N_3)$ nodes in forming the approximation of the solution:

$$f_{rbf}(\mathbf{x}) = \sum_{j=1}^N a_j \phi(\|\mathbf{x} - \mathbf{x}_j\|) + \sum_{j=1}^M b_j p_j(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \quad 0 < M < N.$$

For solution close to the re-entrance corner, we would approximate the solution using the series solution:

$$f_{series}(r, \theta) = \sum_{i=1}^P c_i r^{\frac{2}{3}(i-1)} \cos \left[\frac{2}{3} (i-1) \theta \right], \quad -3\pi/2 \leq \theta \leq 0. \quad (8)$$

For each of the node in X_1 , we would set up an equation according to (7), so we have N_1 equations for them:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = [\mathbf{0}], \quad (9)$$

where \mathbf{A} is an $N_1 \times N$ matrix, \mathbf{B} is an $N_1 \times M$ matrix, \mathbf{a} is an $N \times 1$ matrix, \mathbf{b} is an $M \times 1$ matrix, and

$$\begin{aligned} \{\mathbf{A}\}_{ij} &= \nabla^2 \phi(|\mathbf{x} - \mathbf{x}_j|)|_{\mathbf{x}=\mathbf{x}_i}, \\ \{\mathbf{B}\}_{ij} &= \nabla^2 p_j(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_i}, \\ [\mathbf{a}] &= \{a_1, \dots, a_N\}^T, \\ [\mathbf{b}] &= \{b_1, \dots, b_M\}^T. \end{aligned}$$

Let the boundary condition be specified as:

$$B_{ou}(f) = v(\mathbf{x}), \text{ for } \mathbf{x} \in \Gamma, \quad (10)$$

where B_{ou} is the boundary condition operator and $v(\mathbf{x})$ is the prescribed boundary condition. So for each node in X_2 , we would set up an equation according to (10):

$$\begin{bmatrix} \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = [\mathbf{d}], \quad (11)$$

where \mathbf{C} is an $N_2 \times N$ matrix, \mathbf{D} is an $N_2 \times M$ matrix, and

$$\begin{aligned} \{\mathbf{C}\}_{ij} &= bou(\phi(|\mathbf{x} - \mathbf{x}_j|))|_{\mathbf{x}=\mathbf{x}_{i+N_1}}, \\ \{\mathbf{D}\}_{ij} &= bou(p_j(\mathbf{x}))|_{\mathbf{x}=\mathbf{x}_{i+N_1}}, \\ [\mathbf{d}] &= \{v(\mathbf{x}_{1+N_1}), \dots, v(\mathbf{x}_{N_2+N_1})\}^T. \end{aligned}$$

For nodes in X_3 we would we would set up an equation so that the radial basis solution equals to that of the series solution:

$$\begin{bmatrix} \mathbf{E} & \mathbf{F} & \mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{bmatrix} = [\mathbf{0}], \quad (12)$$

where \mathbf{E} is an $N_3 \times N$ matrix, \mathbf{F} is an $N_3 \times M$ matrix, \mathbf{E} is an $N_3 \times P$ matrix, and

$$\begin{aligned} \{\mathbf{E}\}_{ij} &= \nabla^2 \phi(|\mathbf{x} - \mathbf{x}_j|)|_{\mathbf{x}=\mathbf{x}_{i+N_1+N_2}}, \\ \{\mathbf{F}\}_{ij} &= \nabla^2 p_j(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_{i+N_1+N_2}}, \\ \{\mathbf{G}\}_{ij} &= -r^{\frac{2}{3}(j-1)} \cos \left[\frac{2}{3} (j-1) \theta \right] \Big|_{r=r_{i+N_1+N_2}, \theta=\theta_{i+N_1+N_2}}, \\ [\mathbf{c}] &= \{c_1, \dots, c_P\}. \end{aligned}$$

Then, we select additional nodes $X_4 = \{x_{N_1+N_2+N_3+1}, \dots, x_{N_1+N_2+N_3+N_4}\}$ in Ω_2 so that $N_4 >$

P , we would then use the least square method to set up the last P equations. First we find the sum of square of differences between the radial basis function and the series solution at nodes in X_4 :

$$S = \sum_{j=1}^{N_4} (f_{rbf}(\mathbf{x}) - f_{series}(r, \theta))^2.$$

Then, we would set up one equation for each of

$$\frac{\partial S}{\partial c_i} = 0, \text{ for } i = 1, \dots, P. \quad (13)$$

In matrix form, (13) is:

$$\begin{bmatrix} \mathbf{K}^T \mathbf{H} & \mathbf{K}^T \mathbf{J} & \mathbf{K}^T \mathbf{K} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{bmatrix} = [\mathbf{0}], \quad (14)$$

where \mathbf{H} is an $N_4 \times N$ matrix, \mathbf{J} is an $N_4 \times N$ matrix, \mathbf{K} is an $N_4 \times P$ matrix, and

$$\begin{aligned} \{\mathbf{H}\}_{ij} &= \nabla^2 \phi(|\mathbf{x} - \mathbf{x}_j|) \Big|_{\mathbf{x}=\mathbf{x}_{i+N_1+N_2+N_3}}, \\ \{\mathbf{J}\}_{ij} &= \nabla^2 p_j(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{x}_{i+N_1+N_2+N_3}}, \\ \{\mathbf{K}\}_{ij} &= -r^{\frac{2}{3}(j-1)} \cos \left[\frac{2}{3}(j-1)\theta \right] \Big|_{r=r_{i+N_1+N_2+N_3}, \theta=\theta_{i+N_1+N_2+N_3}}. \end{aligned}$$

We can then combine all the equations (9), (11), (12) and (14), we have the following system

$$\mathbf{LP} = \mathbf{Q}, \quad (15)$$

where

$$\mathbf{L} = \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} & \mathbf{0} \\ \mathbf{E} & \mathbf{F} & \mathbf{G} \\ \mathbf{K}^T \mathbf{H} & \mathbf{K}^T \mathbf{J} & \mathbf{K}^T \mathbf{K} \end{bmatrix}, \mathbf{P} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{bmatrix}, \text{ and } \mathbf{Q} = \begin{bmatrix} \mathbf{0} \\ \mathbf{d} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}.$$

The solution of \mathbf{P} from (15) can be obtained by

$$\mathbf{P} = \mathbf{L}^{-1} \mathbf{Q}.$$

Numerical Results

The numerical results from the proposed scheme are compared with that obtained by Yosibash [3], who produced the value of A_i by using finite element method. The computational results for $u's$ were generated 342 collocation points over the domain Ω_u . Table 1 analyses the first four intensity factors c_1, c_2, c_3 and c_4 of the series expression (8).

Table 1

Intensity factors	Results from Yosibash [3]	Results from MQ-RBF Method
c_1	0.6667	0.6676
c_2	-0.4520	-0.4514
c_3	-0.2149	-0.2139
c_4	0.0000	4.1516×10^{-5}

The maximum relative errors of the approximate results is $5.30\text{E-}02$ when comparing to the global solution obtained from series expansion. The small magnitude of relative errors reflects the proposed radial basis function method to produce a reasonable degree of accuracy. Figure 2 shows the predicted results over Ω_u of the underlying problem. The smooth distribution indicates a good performance of using the RBF method in the given model.

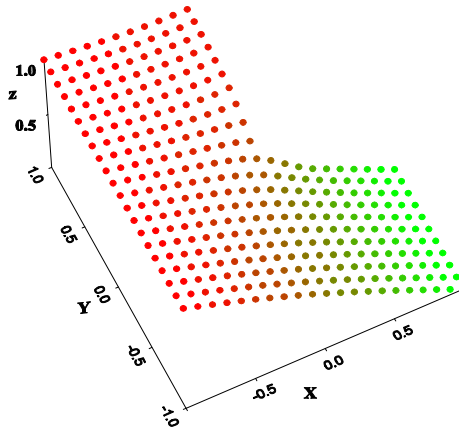


Figure 2: Predicted result of $u(x, y)$ over the L-sharp region.

From the numerical experience, we observed that the results of MQ-RBF method appear to have a same order of magnitude as those results achieved by Yosibash [3], where the authors reported that their first four coefficients are accurate up to the shown 4 decimal places. Our results indicates that the RBF method combined with overlapping domain decomposition is not only an efficiency scheme, it also produced a high level of accurate approximation. The present scheme has been shown to be very effective to overcome the shortcoming of RBF method as mentioned above.

Conclusions

In summary, the MQ-RBF method used in this paper is type of globally supported functions. The disadvantage of such global RBF is that the result in a full matrix which is computationally expensive and may cause instability if the matrix is ill-conditioned, which has seriously hindered its ability from solving large scale problem with a large number of nodal points. This shortcoming leads to the studies of domain decomposition scheme. The combination of RBF scheme and domain decomposition has been verified to be a very effective technique to overcome this shortcoming of RBF method. The overlapping domain decomposition scheme used in this paper is specially designed to overcome the discontinuity of the solution near the singular point. The special region which covers the singular point, is small and Ω_{12} is common region to both Ω_1 and Ω_2 .

On the other hand, the RBF method possesses a number of attractive properties. The greatest attractive properties are the mesh free configuration and the simple mathematical formulation, these properties make the RBF method more flexible in coupling with other remedial numerical schemes. In this paper, we would easily incorporate the domain decomposition and least square approximation scheme with RBF. We showed that the least square approximation with MQ-RBF lead to a small numerical discrepancy in the numerical experiments.

We have illustrated the efficiency of the proposed scheme. However, the scheme can be applied to any elliptic problems with boundary singularity, provided that the solution in the vicinity of the singular point in the form of asymptotic series expansion has been known explicitly. In addition, since the RBF method is insensitive to the dimension of the problem, the scheme can be used to solve higher dimensional problems.

Acknowledgements

The development of the research material was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (UGC/IDS16/14).

References

- [1] Motz, H. (1946) , The treatment of singularities of partial differential equations by relaxation methods, *Quart. Appl. Math.*, **4**, 371-377.
- [2] Grisvard P. (1985), *Elliptic Problems in Nonsmooth Domains*, Pitman Publishers, ISBN: 0-273-08647-2.
- [3] Yosibash, Z. (1997), An Accurate Semi-Analytic Finite Difference Scheme for Two-Dimensional Elliptic Problems with Singularities, *Numer. Methods Partial Differential Eq* 14:281-295.
- [4] Yosibash, Z., (1997), Numerical analysis on singular solutions of the Poisson equation, *Computational Mechanics* 20:320-330.
- [5] Li, Z. C., (1998) , *Combined Methods for Elliptic Equations with Singularities, Interfaces and Infinities*, Kluwer Academic Publishers.
- [6] Rao, B.N. and Rahman, S.,(2000) , An efficient meshless method for fracture analysis of cracks, *Computational Mechanics* 26 398-408.
- [7] Hardy, R. L., (1971) , Multiquadric equations of topography and other irregular surfaces, *J. Geophys. Res.* 176, pp. 1905-1915.
- [8] Micchelle, C. A., (1986), Interpolation of scattering data: distance matrices and conditionally positive definite functions, *Constr. Approx.* 2, pp. 11 - 22.
- [9] Powell, M. J. D.,(1990) , The theory of radial basis functions approximations, chapter 3, *Wavelets, subdivision algorithms and radial basis functions*, ed. Will Light, Vol. II, Oxford University Press, pp. 105 - 210.
- [10] Madych, W. R. and Nelson, S. A., (1988) , Multivariate interpolation and conditionally positive definite functions, *Approx. Theory Appl.*, Vol. 4, pp. 77 - 89.
- [11] Wu, Z. and Schaback, R. (1993) , Local error estimate for radial basis functions interpolation of scattered data, *IMA J. Num. Anal.*, Vol. 13, pp. 13 - 27.

- [12] Francis J. Narcowich and Joseph D. Ward (1994), Generalized Hermite interpolation via matrix-valued conditional positive definite functions, *Math. Comput.* Vol. 63, No. 208, pp. 661-687.
- [13] Madych, W. R. and Nelson, S. A., (1986) , Multivariate interpolation: a variational theory, *Approx. Theory Appl.*, pp. 11 -22, 1986.
- [14] Kansa, E. J., (1990), Multiquadric - a scattered data approximation scheme with applications to computational fluid dynamics - II. Solutions to hyperbolic, parabolic, and elliptic partial differential equations, *Comput. Math. Applic.*, Vol. 19, No. 8/9, pp. 127-161, 1990.
- [15] Hon, Y. C. and Wu, Z. (2000), Additive Schwarz domain decomposition with radial basis functions approximation, *Int. J. Appl. Math.* Vol. 4, No. 1. pp. 81 - 89.
- [16] Wong, S. M., Hon, Y. C., Li, T. S., Chung, S. L. and Kansa, E. J. (1999), Multi-zone decomposition for simulation of time-dependent problems using the multiquadric scheme, *Comput. Math. Appl.*, Vol. 37, No.8, pp. 23 - 43.

Application of Parametric Modeling in the Early Design Phase for an Interdisciplinary Design Approach for Adaptive Buildings

†Frederik Ernst¹, and Achim Menges¹

¹Institute for Computational Design and Construction, University of Stuttgart, Germany

† Presenting and Corresponding author: frederik.ernst@icd.uni-stuttgart.de

Abstract

The integration of adaptive elements to buildings might reduce significant amounts of material and energy. Load-bearing structures are extremely oversized for the predominant period of use. If it is possible to dampen the stress peaks, significant quantities of building materials could be saved.

As a result, the overarching objective is to enable material savings through the deployment of energy when required and to anticipate such an approach in early design through computational processes. For adaptivity, almost all design and form-finding methods embedded in the planning process require completely new exploration and development.

A building with adaptive qualities has no constant properties. Therefore, such structures can't be designed with the usual design methods. The related knowledge and components from mechanical and aerospace engineering need to be transferred, adapted, and redeveloped, since the building industry lacks the experience of adaptive qualities.

Furthermore, this means an increase in the number of disciplines participating in the planning process.

This Paper presents insight in the ongoing process of building a twelve story tower with adaptive qualities by an interdisciplinary research team from the fields of architecture, mechanical engineering, civil engineering, aerospace engineering, textile engineering, social science and computer science.

For this task, parametric modeling was used in the early design phase to investigate the advantages and limitations of this method in terms of adaptive behavior. The findings will be the foundation for a computational design approach in early design for adaptive buildings with the integration of knowledge from the different research fields in the future.

Keywords: Computational Design, Adaptive Buildings, Early Design, Material Savings, Interdisciplinary, Parametric Modeling, Building Information Modeling

Introduction

Globally, the building industry consumes 40% of our global resources, produces about 50% of the entire waste, and is responsible for 40% of energy consumption and emissions [1]. Furthermore, 2 billion people need new apartments, workplaces, and infrastructures within the next 16 years because of the rapid growth of the world population [2]. The creation of such a big construction volume with conventional approaches is impossible without tremendous consequences for our planet. Therefore, a new approach is vital.

For the future, the described problems require building more with less material, waste, energy consumption, and emission. The integration of adaptive elements inside a building can reduce significant amounts of material and energy. Nowadays all load-bearing structures are designed for peak loads, which usually occur in very rare cases. As a result, they are extremely oversized

for the predominant period of use. If it is possible to dampen the stress peaks by manipulating the stress conditions in the load-bearing construction, significant quantities of building materials could be saved. Adaptivity allows a façade or a supporting structure to assume a variety of different states. This is the major difference to the current building design, which assumes one state for any possible scenario. This raises the question of how we design and represent these multiple conditions. Since the building industry never worked with adaptive qualities before, the related knowledge and components from other disciplines like mechanical and aerospace engineering need to be transferred, adapted, and redeveloped.

The Collaborative Research Center (SFB) 1244 “Adaptive Building Skins and Structures for the Built Environment of Tomorrow”, funded by the German Research Foundation (DFG), focuses on the basic principles, as well as the potential and the impact of integrating adaptive elements in load-bearing structures, facades and interior work of buildings.

Here, the term “adaptive” is used in the context of precise modification of the geometry, adjustment of material characteristics and the properties of construction elements over time. The aim is to minimize the amount of required material and minimize embodied energy. The interdisciplinary research team consist of scientists on the topic of architecture, mechanical engineering, civil engineering, aerospace engineering, textile engineering, social science and computer science.

As a first milestone, a 36-metre-high tower with an adaptive load-bearing structure will be built for specific research into the topic of adaptivity [Figure 1]. At a later point in time, the demonstrator tower will also be equipped with different adaptive façade systems, which are still in development. The demonstrator serves as a proof-of-concept and gives the researches the ability to contribute and analyse the process of creating such a tall adaptable building.

The main part of the authors work was to assists and observes the design, planning and construction process of the demonstrator. For this task parametric models were created with

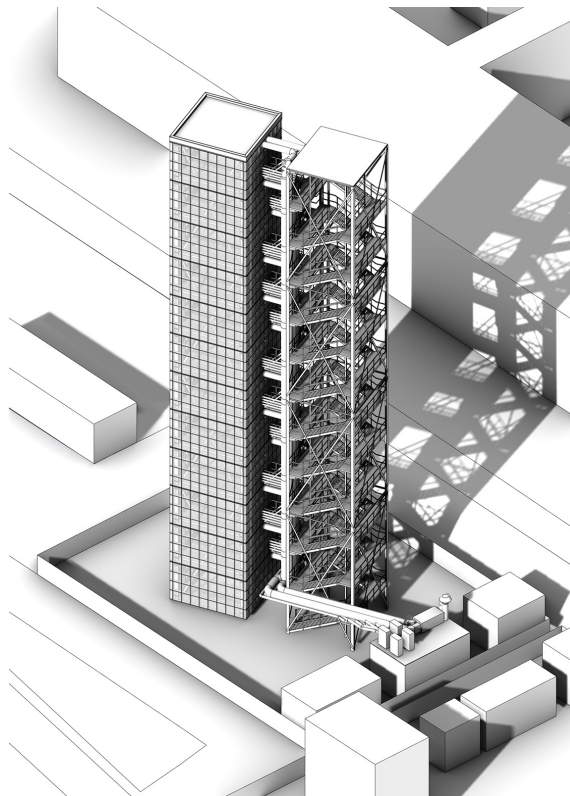


Figure 1. Twelve-story demonstrator tower

the Building Information Modeling Software “Autodesk Revit” and relevant information was shared with other research fields to guarantee an interdisciplinary design approach from the start. The gathered information of the building process will provide the framework for novel computational methods for a digital planning environment to help architects anticipate adaptive qualities in early design in a later stage. The findings from the planning of the demonstrator can be used to specifically identify the restrictive aspects of linear, conventional planning and to define and concretize requirements more precisely for the investigation of computational processes.

Also the design model of the demonstrator will be used in the future as a benchmark for an adaptive design model which is to be developed.

Current State

The established architectural planning process is hierarchical. A first draft developed by the architect can only be successively refined, constructively developed, and realized in the later phases in cooperation with experts like structural engineers.

The formulation of the task is initially vague and is only defined over time by the generation and evaluation of variation in the early design phase. For this reason, no optimization is possible right from the start.

Today, stationary and static planning is carried out. Each element of a building can be described precisely and has only one state for each occurring situation. For example, façade insulation has a set heat transfer coefficient, which remains identical regardless of external weather conditions and is unable to adapt.

The linearity and hierarchy of the existing process make it more difficult to integrate adaptivity into the necessary planning, structural, structural-physical, mechanic, and mechatronic considerations in the early design decisions. But these early design decisions account for 80% of a building's energy and material consumption [3]. On the other hand, this phase is only compensated with 24% of the commission fee according to the German Regulations for Architects and Engineers [4]. As a result, many architects lose the opportunity to examine far-reaching decisions and their implications in depth at this phase and postpone conflicts to later phases in the planning process at which a higher compensation is paid. But at this later stage it is not possible to explore the solutions space for better individuals anymore, because the design phase has already been completed and a final design has been selected.

Building Information Modeling is a method which attempts to break the linearity of the planning process and enables collaborative work on 3D models, avoidance of collisions, quantity and mass determination, coordination of processes on construction sites, etc. BIM is an approach for merging and enriching the planning and construction process with information. But the method lacks considerable influence on the early design phase. Therefore, BIM is usually applied after the conceptual design process, when there aren't many changes to be expected. Almost two-thirds of the UK BIM user think, that BIM won't help to reduce greenhouse gas emissions in the built environment by 50% [5]. Although there is research in the field of using BIM in the early design phase, nobody is approaching an interdisciplinary design approach for adaptive buildings. In contrast to existing information models, an extended information model should be able to map and quantify the different states of the structure and systems with dynamic and instationary properties, as well as their temporal change due to adaptivity.

The existing methods and approaches of architectural planning can therefore only be conditionally applied to adaptive structures, because they cannot fulfil the fundamentally interdisciplinary character right from the start. The essential requirement for the development of new methods and procedures for adaptive architecture is to enable integrative planning right from the beginning.

Adaptive Envelope Systems and Adaptive Structures form a linked system of sensors and actuators with a variety of interdependencies. In order to better investigate these interactions, the architect needs digital support to assist him in exploring the solution space. This process has to recognize or point out characteristics and thus create new solutions which the human being alone can hardly foresee.

Observations of the design and planning process

The task was to analyze the capabilities of current state-of-the-art Design and Planning software in the Building industry regarding the advantages and limitations for adaptivity on a real showcase. “Autodesk Revit” is the market leader for BIM Software with a usage rate in the UK of 41%, followed by “Graphisoft ArchiCAD” with 15% [5]. Therefore “Autodesk Revit” was used for parametric Modeling of the demonstrator building. The parametric models were created on the foundation of vertical and horizontal Gridlines with dependencies to each other. On top of these guides, family members were attached and could adapt to changes of the gridlines.

As in the Beginning, this approach of parametric modeling seemed the most logic. The Topology of the Building was set. A ten stories tower with a height of 36 Meters had no space allocation plan and therefore very little requirements concerning the architectural design.

The main task seemed to be massing, detail drawing and scheduling to support the construction planning. Later the BIM Model should be transferred to become a control model of the adaptive elements.

Over time, more Inputs and Outputs were added to the Model [Figure 2]. Some of these requirements are specific for adaptive design and helped the interdisciplinary approach of the planning process. The BIM Model was used as central hub for all information.

A small footprint of the building would help the oscillation behavior for the testing scenario of the damping actuators of the structure. For this reason, the access had to move to a separate tower standing nearby, with stairwell, an elevator and bridges connecting the demonstrator.

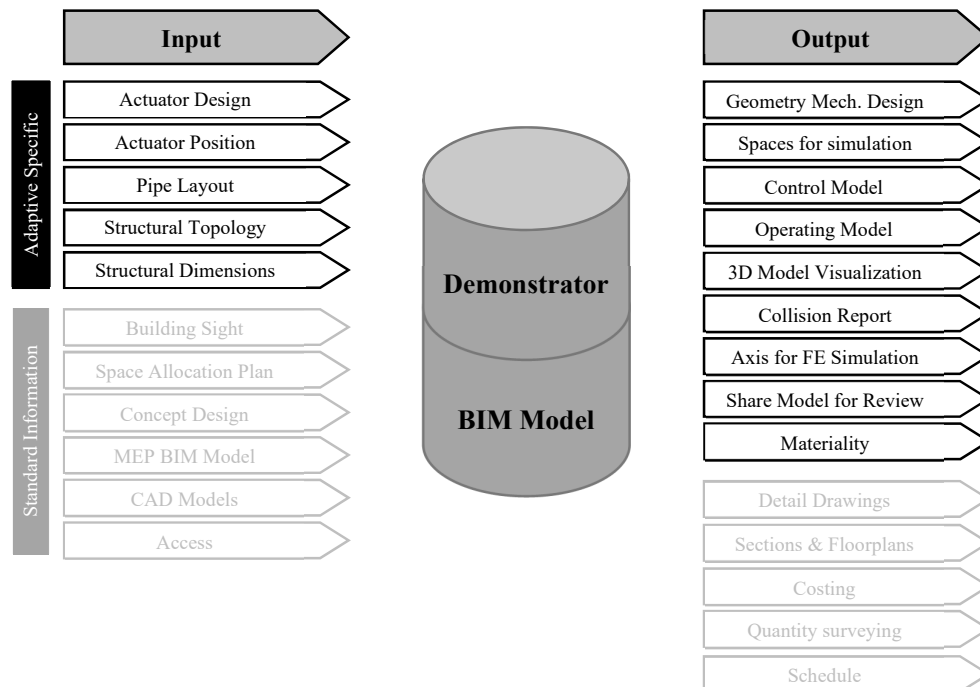


Figure 2. Inputs and Outputs of the BIM Model

This resulted in a conflict of the structural and the architectural design, because the bridges had to avoid the structural diagonal branching in order to access the demonstrator.

The principles for the adaptive structure were developed in a parallel process by civil and mechanical engineers. Over time, the tower had to become slender to increase the oscillating behavior. This was easily achieved by changing the parameter of the Center-to-Center distance. Over the course of one year the distance changed from 6000 millimeters, to 5300 Millimeters and 4300 Millimeters. Massing and drawings were updated on the fly without the need change multiple plans by hand. The process for the actuator integration [Figure 3.] was included into the linear and hierarchical planning process, because of the limited timespan to build the demonstrator. Also it was not the main goal to test novel planning methods in this phase and risk the fixed date for the completion of the building. Still this process leads to some interesting findings: Left side of the diagram shows similarities to the normal planning process, while the most adaptive activities happen on the right side. The concept design lacks adaptive qualities at the start. Therefore some feedback loops had to be made, which the architect didn't like, because it meant more work which wouldn't be reimbursed.

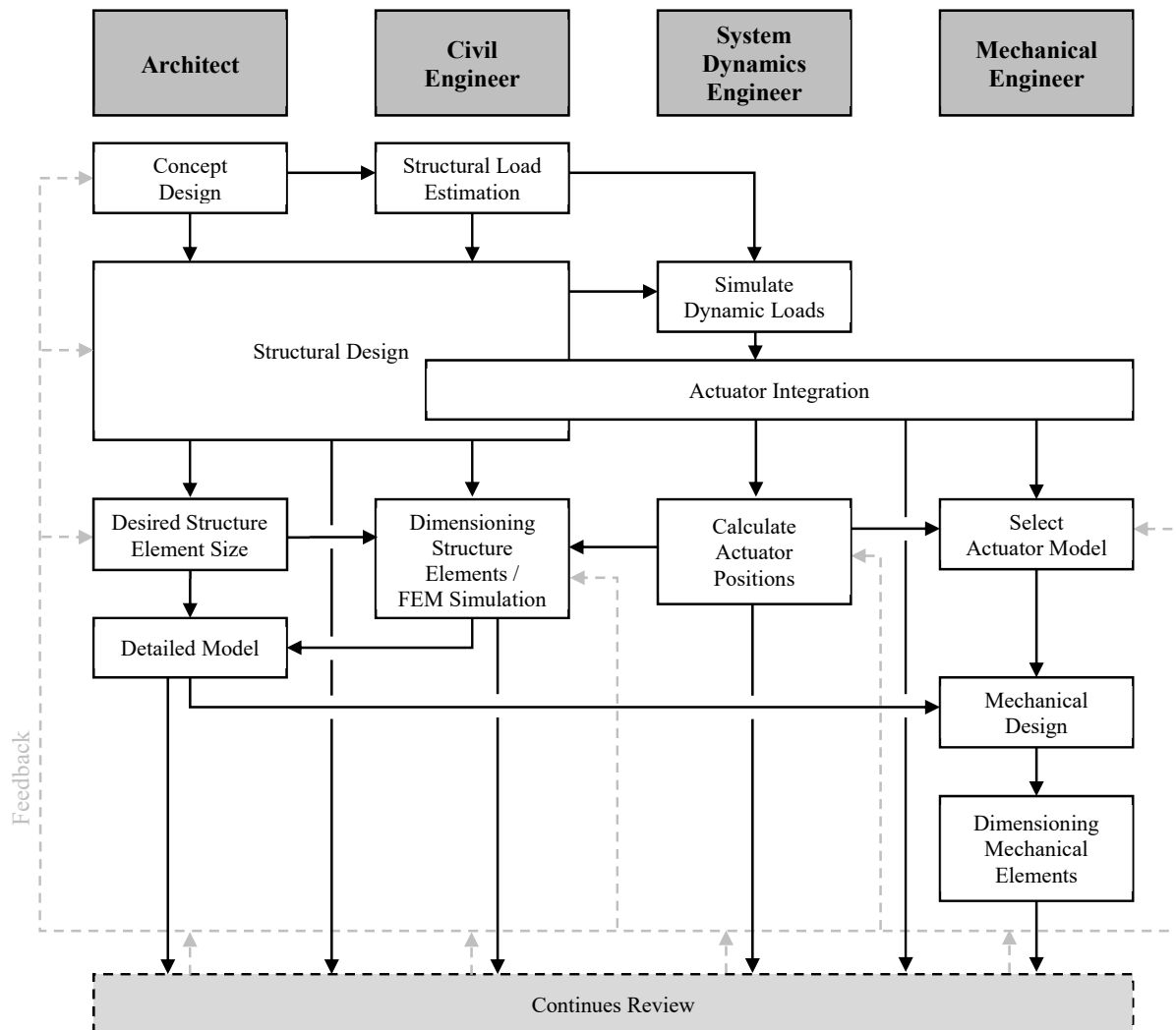


Figure 3. Design Process for the adaptive Structure

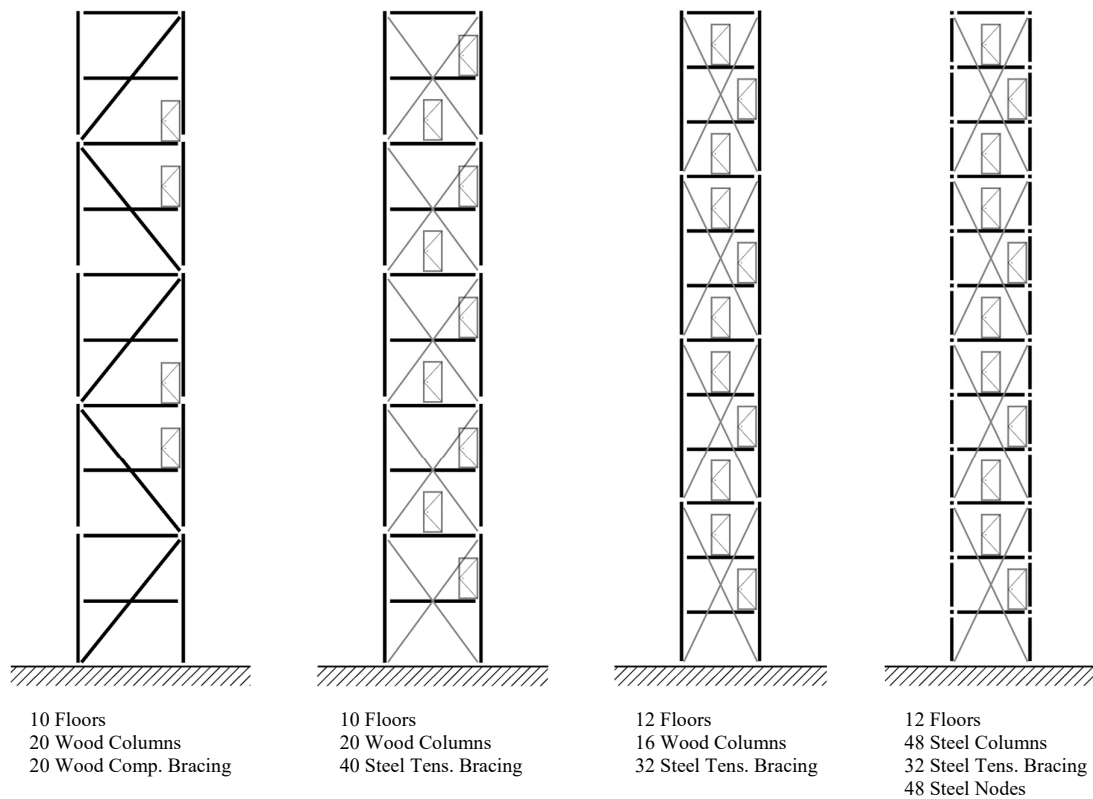


Figure 4. Changes of the Structural Topology in the Design Process

Fundamental topology modifications in the structure make parametric modeling difficult [Figure 4. & 5.]. Most of the time the parametric model can't adapt to topology changes and a new parametric model needs to be created, because the changes couldn't be foreseen, while creating the parametric model. Due to continuous improvements in the calculation of the actuator placement, changes in the adaptive structure topology had to be made more frequently. The first design envisaged a wooden structure. But a wooden column with an actuator could not be developed in the short time. So it was decided to use only a steel structure for the time being.

This also results in the idea of having maximum flexibility in changing the structure topology even after the building was finished. Therefore, the idea came up that every column or diagonal could integrate an actuator. Also every element of the structure can be replaced by a different design e.g. concrete or wood columns after the construction is finished. This is because there may be improved actuator placement models in the future, which might need further investigation.

So an attempt was made to change to Generative Parametric Modeling with the ability to solve the previous mentioned problems. This process is currently under investigation. However, implementing such an approach under the recognized conditions takes a considerable amount of time and still cannot guarantee to cover all upcoming scenarios of changes in the topology. But in the early design phase it is important to have no limitations in the creation and evaluation of many variations without spending too much time creating complex solutions. As mentioned before, the time in this phase is restricted and not sufficiently compensated.

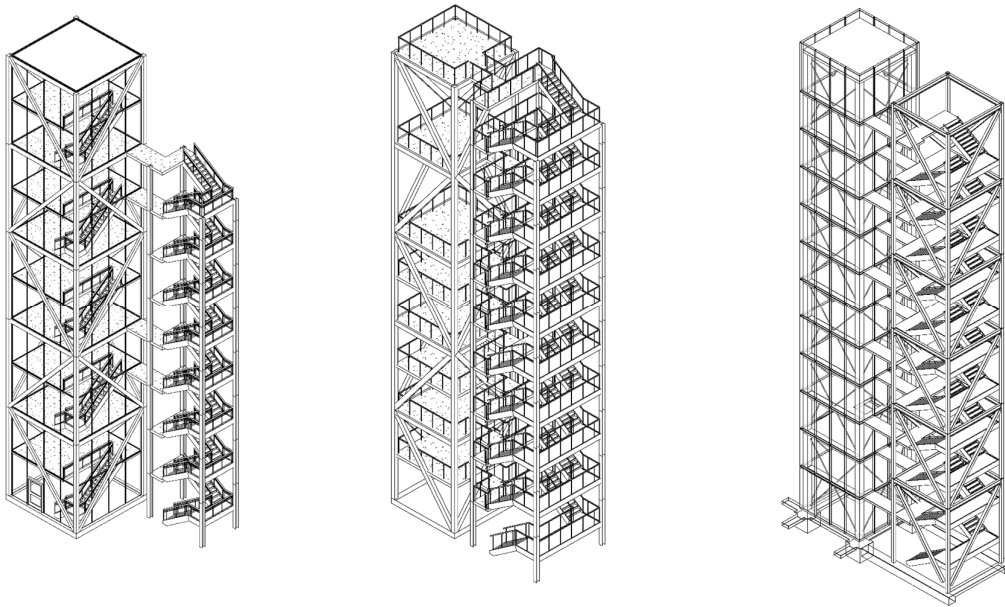


Figure 5. From Left to right: Evolution of the Building

Conclusions and Outlook

Parametric BIM modeling is capable of assisting in the early design process of adaptive buildings of an interdisciplinary team. Still it lacks the ability to describe dynamics and instationary properties of adaptive elements and is therefore limited in exploring the solution space in the early design phase. Hence, a holistic description of adaptive qualities in an extended building information model must be developed. Because of the necessity to change the structural topology during the exploration of the design space, further development of parametric models with regard to topological flexibility and variability is required. A graph-based approach for parametric modeling could be one feasible solution. Other longtime goals are the Development of computational processes and tools for the exploration and variation reduction in the early design phase, with the principles of adaptive buildings and the Integration of processes and methods into a digital design environment which enables feedback and optimization on the principle of simulation and analysis with the involved disciplines.

Acknowledgments

This work was supported by Collaborative Research Center (SFB 1244) " Adaptive Building Skins and Structures for the Built Environment of Tomorrow " (subproject A02) of the German Research Foundation (DFG).

References

- [1] UN Environment Programme (2009), Buildings and Climate Change: Summary for Decision-Makers
- [2] United Nations, Department of Economic and Social Affairs, Population Division (2014) World Urbanization Prospects: The 2014 Revision, <https://www.compassion.com/multimedia/world-urbanization-prospects.pdf>
- [3] Bogenstätter, U. (2000), Prediction and optimisation of life-cycle costs in early design, *Building Research & Information* **28.5/6**, 376–386.
- [4] Regulations for Architects and Engineers in Germany (HOAI)
- [5] NBS National BIM Report 2017

Comparison of Characteristic-wise WENO and Central Difference Schemes With Numerical Viscosity Models for the Unsteady Compressible Flow

Hyun-Jin Kwon¹ and †*Se-Myong Chang²

¹Department of Mechanical Engineering, Kunsan National University, Korea.

²School of Mechanical Convergence System Engineering, Kunsan National University, Korea

*Presenting author: smchang@kunsan.ac.kr

†Corresponding author: smchang@kunsan.ac.kr

Abstract

In the computation of unsteady compressible fluid dynamics, unnecessary numerical oscillations can appear in the domain near the sharp discontinuities such as shock or slip line, especially for the high-order spatial accuracy. The WENO(weighted essentially no oscillation) method can delete the oscillation of numerical dissipation error, and so can the central difference methods with artificial viscosity, which are very economical for the computational cost. The characteristics of conventional WENO and central difference schemes are compared with each other for a benchmark problem in this study where LF(Lax-Friedrichs) and Garnier filters are used with ACM(Artificial Compression Method) switch acting near the sharp-gradient discontinuities.

Keywords: Compressible Flow, WENO, Central Difference, Numerical Filter, ACM Switch, Numerical Viscosity

Introduction

Originally, the discontinuity of signal in the physical property can produce improper oscillation due to a loss of information in the sampling of continuous analog data, which can be analyzed as the truncation of high-order Fourier series terms even in the various experiments. In the similar principle, the truncation error in Taylor series expansion generates numerical oscillation for the finite difference approximation. For the spatial accuracy more than the second order, a dissipation error is inevitable in the central difference schemes. In the WENO schemes[1] developed by Jiang and Shu, the smooth indicator acts as a sensor for the gradient, giving a filtered solution like upwind or TVD(total variation diminishing) methods. However, this kind of schemes sacrifices an order of spatial accuracy that is an odd number: for example, three, five, and seven, etc. The central difference schemes with ACM sensors[2] developed by Yee et al. has shown a possibility to develop a filtered method with filters, and various filter models are proposed such as Garnier et al.[3] and Kim and Kwon [4]. Generally, so far the central-difference based methods show better result in the smooth region like vortex and slip layer, but the numerical oscillation can be very critical in the discontinuous waves. Therefore, systematic approach for the selection of schemes and filters is required in the development of numerical codes. In this study, we compared characteristics and performances related with three numerical schemes for the numerical simulation of unsteady compressible flow: WENO, central difference with Garnier filter, and central difference with LF filter.

Numerical Schemes

In this section, three numerical methods are explained, and the Euler equations is expressed for the conservative dependent variable vector $\mathbf{Q} = [\rho, \rho u_k, E]^T$:

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}_k}{\partial x_k} = 0 \quad (1)$$

where the flux is $\mathbf{F}_k = [\rho u_k, \rho u_k u_j + p \delta_{kj}, u_k(E + p)]^T$, and the equation of state is $p = (\gamma - 1)(E - \rho u_k u_j)/2$.

Eq. (1) is semi-discretized in space with the third-order Runge-Kutta algorithm for temporal integration[1]:

$$\frac{\Delta \mathbf{Q}_j}{\Delta t} = L(\mathbf{Q}_j^n) = -\frac{1}{2} \left(\mathbf{F}_{j+\frac{1}{2}}^n - \mathbf{F}_{j-\frac{1}{2}}^n \right) \quad (2)$$

where the time step Δt is restricted by CFL(Courant-Friedrichs-Lewy) condition. The numerical flux in Eq. (2) is computed with WENO and central difference methods.

WENO Scheme

The CW(characteristic-wise)-WENO method consists of the following algorithm[1]:

1. Projection to the characteristic field
2. Lax-Friedrichs flux splitting
3. WENO reconstruction
4. Transform back into physical projection

The left eigen-vector matrix from the Roe-averaged Jacobian matrix at the right face of the computational cell[5], $\mathbf{R}_{j+\frac{1}{2}}^{-1}$ is used for the projection to the characteristics field.

$$\mathbf{q}_l = \mathbf{R}_{j+\frac{1}{2}}^{-1} \mathbf{Q}_l \quad (3)$$

$$\mathbf{f}_l = \mathbf{R}_{j+\frac{1}{2}}^{-1} \mathbf{F}_l \quad (4)$$

Using Lax-Friedrichs flux splitting, the maximum eigen-value is calculated for the region of influence in the hyperbolic partial differential equation system, Eq. (1). The WENO interpolation applies the convex sum of weighted average in the numerical stencil. The weights are coefficients expressed as a function of smooth indicators and optimal coefficients for the finite difference. The weighted combination of ENO flux results in odd numbered order of spatial accuracy. In the last process, Eqs. (3-4) are transformed back to the primitive variables. The scheme is implemented as a fifth order of spatial accuracy in the present study.

Central Difference Schemes

The central difference flux is applied in Eq. (2), and it is filtered at the last procedure in each time marching:

$$\mathbf{Q}^{n+1} = \hat{\mathbf{Q}}^{n+1} + \Delta t L_f(\hat{\mathbf{Q}}^{n+1}) \quad (5)$$

where L_f is the spatial operator applying a low numerical viscosity filter, which can be controlled selectively for the large-gradient region with ACM switch function[2].

The numerical viscosity that is very similar with a flux in Eq. (5), ϕ can be modelled with various methods. Garnier filter[3] subtracts the central difference flux component, \hat{f}^0 from the WENO characteristic flux, which is similar with upwind method.

$$\phi_{i+\frac{1}{2}} = \hat{f}_{i+\frac{1}{2}}^+ + \hat{f}_{i-\frac{1}{2}}^+ - \hat{f}_{i+\frac{1}{2}}^0 \quad (6)$$

Kim and Kwon[4] designed a new filter based on the Lax-Friedrichs numerical viscosity using characteristic-wise WENO method as a filter.

$$\phi_{i+\frac{1}{2}} = -\frac{1}{2}\lambda_{max} \left(Q_{i+\frac{1}{2}}^R - Q_{i+\frac{1}{2}}^L \right) \quad (7)$$

where λ_{max} denotes the maximum eigen values.

Convergence Test

Eq. (1) under an initial condition, $\rho_0 = 1 + 0.2 \sin(\pi x)$ and $u_0 = p_0 = 1$ is solved for the convergence test with three methods described in the previous section. Errors and orders are given in Table 1 for the number of cells, N in the domain of $x \in [-1, 1]$. The order of accuracy satisfies fifth order at $N=160$ for all schemes.

Table 1. Result of convergence test

Numerical Method	N	L_1 error	L_∞ error	L_1 order	L_∞ order
WENO5	10	1.222E-02	1.757E-02	-	-
	20	6.495E-04	1.002E-03	4.23	4.13
	40	2.075E-05	3.735E-05	4.97	4.75
	80	6.479E-07	1.198E-06	5.00	4.96
	160	2.017E-08	3.640E-08	5.01	5.05
Central Difference with Garnier Filter	10	3.915E-03	5.377E-03	-	-
	20	1.060E-04	1.657E-04	5.21	5.02
	40	2.500E-06	5.387E-06	5.41	4.94
	80	3.948E-08	1.338E-07	5.98	5.33
	160	1.151E-09	3.612E-09	5.10	5.21
Central Difference with LF Filter	10	3.915E-03	5.377E-03	-	-
	20	3.072E-04	5.085E-04	3.67	3.40
	40	9.784E-06	1.955E-05	4.97	4.70
	80	3.058E-07	5.760E-07	5.00	5.09
	160	8.942E-09	1.504E-08	5.10	5.26

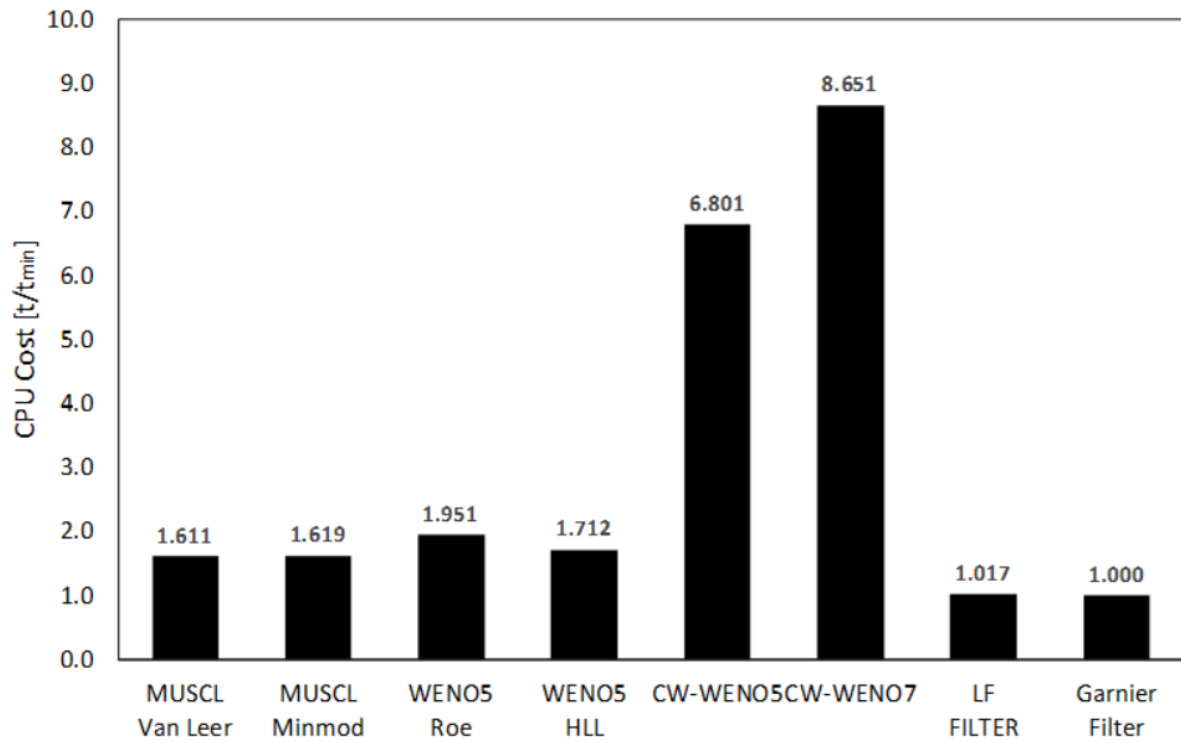


Figure 1. Comparison of CPU times

Computational Cost

The computational cost for the same problem with the same number of grids is compared in Fig. 1, which gives data from various schemes. The minimum time cost is achieved from the central difference with Garnier filter. It shows that the LF filter consumes only 1.7% computational time more than the minimum, but, however, the characteristic-wise WENO consumes about 6.8 times of time because it must perform the matrix inversion to transform back from Eq. (3-4). The time cost of WENO increases about 6.8 times because it should perform the matrix inversion at each intermediate Runge-Kutta time step, and the performance of LF filter is almost similar with Garnier model.

Summary

Three numerical methods are implemented for the computation of unsteady compressible flow: characteristic-wise WENO and central difference method with Garnier and LF filters. From the convergence test, all of them achieved the fifth-order spatial accuracy. The Garnier filter shows the best performance in the convergence test and the CPU cost. However, the test for more complicated problem related with shock and discontinuity waves can differ from the present simple benchmark problem.

Acknowledgement

This work is financially supported by NRF (National Research Foundation, 2016R1D1A3-B01015543) and KETEP (Korea Institute of Energy Technology Evaluation and Planning, No. 20174010201350).

References

- [1] Jiang G.S., and Shu. C.W. (1996) Efficient Implementation of Weighted ENO Schemes, *Journal of Computational Physics* **126**, 202-228.
- [2] Yee, H.C., Sandham, N.D., and Djomehri, M.D. (1999) Low-Dissipative High-Order Shock-Capturing Methods Using Characteristic-Based Filters, *Journal of Computational Physics* **150**, 199-238.
- [3] Garnier E., Sagaut, P., and Deville, M. (2001) A Class of Explicit ENO Filters with Application to Unsteady Flows," *Journal of Computational Physics* **170**, 184-204.
- [4] Kim, D. and Kwon, J.H. (2004) A Low Dissipative and Dispersive Scheme with a High Order WENO Dissipation for Unsteady Flow Analyses, AIAA 2004-2705.
- [5] Roe, P.L. (1981) Approximate Riemann Solvers, Parameter Vectors, and Difference Scheme, *Journal of Computational Physics* **43**, 357-372.

Timber shear walls: numerical assessment of the equivalent viscous damping

**†*Giorgia Di Gangi¹, Cristoforo Demartino², Giuseppe Quaranta¹, Marco Vailati¹,
Giorgio Monti^{1,2}, Marc'Antonio Liotta¹**

¹ Department of Structural and Geotechnical Engineering, Sapienza University of Rome, Italy

² College of Civil Engineering, Nanjing Tech University, PR China

†*Corresponding and presenting author: giorgia.digangi@uniroma1.it

Abstract

The seismic performance of timber shear walls is studied in this work, with focus on the energy dissipation ensured by sheathing-to-framing connections. Numerical non-linear analyses are carried out using a parametric numerical model developed in OpenSees and varying some basic design variables affecting the overall racking capacity of the wall, namely: aspect ratio, nails spacing and number of vertical studs. The equivalent viscous damping has been assessed by estimating the damping factor η through the Capacity Spectrum Method.

Keywords: Timber shear walls; racking capacity; energy dissipation; equivalent viscous damping; Capacity Spectrum Method.

Introduction

Timber light-framed constructions are widely used in North America, New Zealand and Northern Europe. These structural systems are very attractive for several reasons, including aesthetic pleasure, sustainability and rapid assembly of the elements. Moreover, they present a fairly good earthquake resistance, basically attributable to the high strength-to-density ratio of timber and to the remarkable ductility of joints with metal fasteners, which ensure reduced inertia forces and good energy dissipation, respectively.

Within this framework, a large amount of research on timber shear walls was carried out in the last decades. In fact, early researches on their mechanical performances date back to 1927 [1]. Existing studies on racking resistance, stiffness and ductility conducted by means of experimental, numerical and analytical methods have demonstrated the good mechanical performances of light-frame wall assemblies. Generally, timber has a poor dissipative capacity, unless it is properly reinforced [2], while the steel connections can ensure a good amount of plastic deformation and, as a consequence, a significant energy dissipation. In order to take into account this aspect, the designers can refer to the force-based design method proposed by EuroCode 8 [3], which allows to reduce the demand of the elastic acceleration spectrum by applying a reduction factor. Two approaches are commonly employed to this end, namely the N2 method [4] and the Capacity Spectrum Method [5][6]. The latter was considered in Ref. [7] to correlate structural damping and drift in timber-framed buildings. Overall, few efforts have been spent so far to analyze the mechanical behaviour and the energy dissipation of a single wall, and few parametric analyses are available that consider different wall configurations [8][9][10]. Therefore, an original parametric FE model has been implemented in the present work by means of the open-source software OpenSees [11] in order to assess the equivalent viscous damping of timber light-frame shear walls by estimating the damping factor η through the Capacity Spectrum Method.

Timber light-frame shear walls

Timber light-frame shear walls are employed in platform framing buildings. It is pointed out that only partially anchored walls will be investigated in this study. Details related to the classifications of timber shear walls can be found in [12] and [8] whereas the interested reader can refer to [13], [14] and [15] for detailed explanations about the differences between balloon and platform framing buildings. A timber light-frame shear wall is composed by vertical studs and horizontal joists (which belong to the frame) connected at their ends with internal constraints (typically modeled as hinges). This integrated system is braced by means of a sheathing panel, linked to the frame by using metal fasteners such as nails, screws and staples. The sheathing panel, in turn, can be built using different materials, like OSB, ply-wood, gypsum, glued laminated Guadua bamboo [16] and so on. The size of the sheathing panel, which could be used to brace one or both sides of the wall, sets the dimension of the frame. A typical size of a shear wall is $1.22\text{ m} \times 2.44\text{ m}$ or $2.44\text{ m} \times 2.44\text{ m}$, whereas the framing elements cross-sections are about $38\text{ mm} \times 89\text{ mm}$ and $38\text{ mm} \times 140\text{ mm}$ for internal and external wall studs, respectively [17]. A typical configuration with both sides braced with a sheathing panel is shown in Fig. 1.

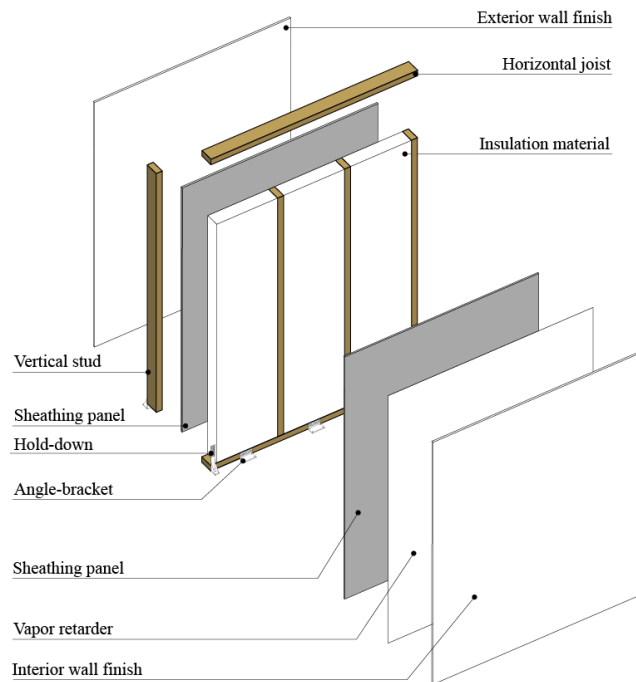


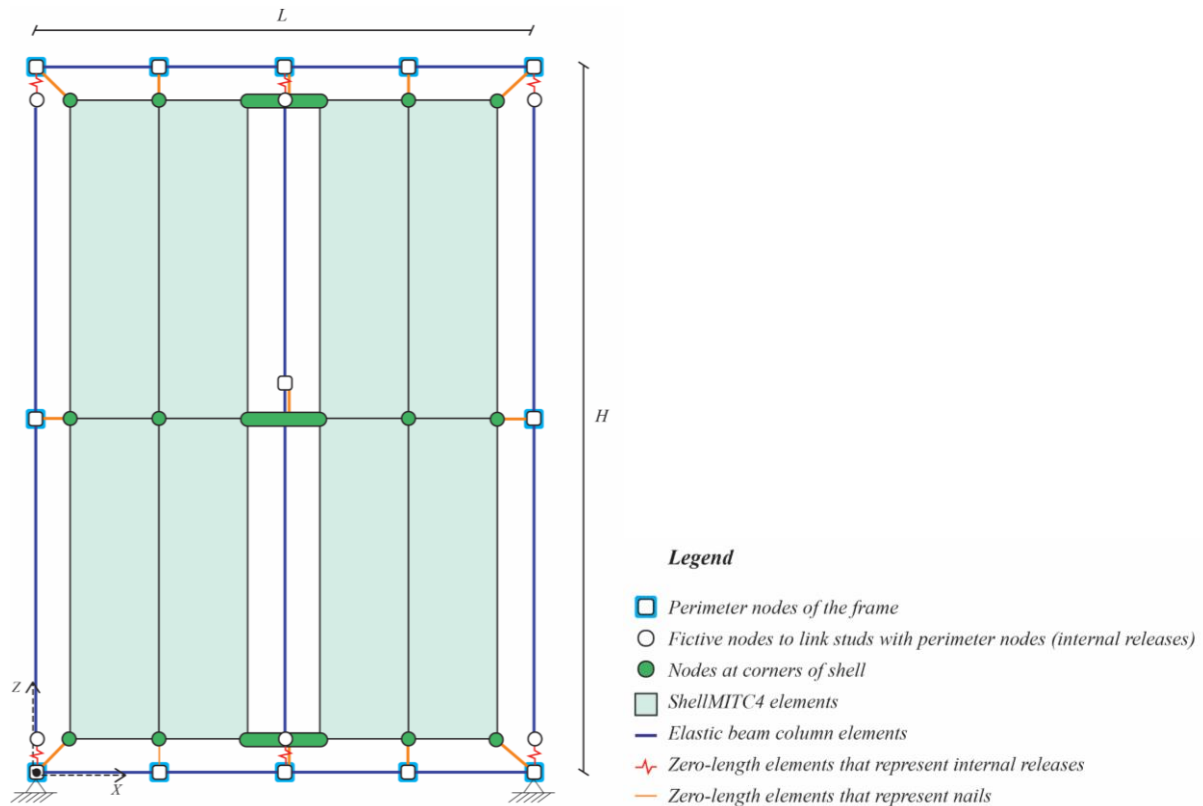
Figure 1. A typical configuration of a fully anchored timber light-frame shear wall braced on both sides, with further layers to improve thermal performances and fire-vapor resistances.

This integrated system is conceived to resist to different static, quasi-static and dynamic actions and its performances related to thermal insulation and fire-vapor resistances are often improved by adding further specific layers exploiting the thickness of external wall studs. Typically, the connections between framing elements and sheathing panels are made by means of 6D, 8D and 10D nails with thick shank placed on perimeter framing elements (the usual spacing is 50 mm, 75 mm or 100 mm) and intermediate studs. In the latter case, the spacing could be two or three times that on the perimeter studs because, as pointed out in Ref. [18], nails on the intermediate studs are only meant to prevent buckling of the sheathing panel and do not contribute to the

racking capacity of the wall. In the so-called fully anchored timber shear wall, the connections with either the foundation or the lower storey shear wall are made by means of steel brackets, which prevent both lifting and horizontal relative sliding. The effects induced on the overall mechanical response by sheathing-to-framing connections have been investigated in [19][20][21]. Further details about base and stud-joist connections can be found in [22] and [20], respectively.

Numerical modeling and validation

To the best authors' knowledge, the finite element (FE) model developed in this work is the first numerical model of timber light-frame shear wall implemented in the open-source software OpenSees [11]. The model has been implemented in the TCL environment in such a way to allow rapid definition of all geometric parameters affecting the racking capacity of the shear walls, namely: *i*) panel size, *ii*) horizontal and vertical nails spacing, *iii*) number of vertical studs. Once these parameters are defined, the number of nodes and elements are updated automatically. For this FE model, it is assumed that the base and height of the shear wall are aligned with the *x*-axis and the *z*-axis, respectively. The frame has been modeled using elastic beam column elements, whereas non-linear coupled zero-length link elements are adopted to represent sheathing-to-framing connections. The sheathing panels are modeled by means of ShellMITC4 elements, whose mesh size depends on the nails spacing. In order to consider the constraints at the ends of framing elements acting as hinges, zero-length elements with a low stiffness value for the rotational degree of freedom along the *y*-axis have been used. Fully-fixed boundary constraints are assumed at the bottom corner nodes, as shown in Fig. 2, in order to assess the energy dissipation of the wall ensured only by the sheathing-to-framing connections.



**Figure 2. FE model implemented in OpenSees
(5 × 3 nodes are considered in this scheme).**

The reference wall configuration is the one considered in Ref. [23], which has the following geometrical features: width 1.8 m, height 2.6 m, nails spacing 50 mm, 4 vertical studs with internal releases (specimen PLS8). The mechanical characteristics used for the framing elements are referred to red spruce wood species with strength class C24, according to EN 14081-1 [24] and UNI EN 338 [25]. The SAWS mechanical model, originally proposed by Ref. [26] and developed in Ref. [27], has been adopted to simulate the behaviour of sheathing-to-framing connections. The corresponding model parameters are identified against the experimental result given for a single nail Φ 2.8 by Ref. [23]. In doing so, the non-classical identification methods presented in Ref. [28] have been used, adopting the following objective function:

$$f(\mathbf{x}) = \frac{1}{S \cdot \text{var}(F^{\text{exp}})} \sum_{s=1}^S (F_s^m(\mathbf{x}) - F_s^{\text{exp}})^2 \quad (1)$$

where \mathbf{x} is the vector collecting the model parameters whereas F_s^m and F_s^{exp} are predicted and experimental force values, respectively. Moreover, s is the generic sample (S denotes the total number of samples) and $\text{var}(F_s^{\text{exp}})$ is the variance of the experimental force values. The comparison between experimental and identified force-displacement curves of a single nail is shown in Fig. 3, together with the comparison between experimental and predicted load-displacement curves of the reference wall. It is possible to observe that racking capacity and hysteretic cycles evaluated using the proposed numerical FE model are in good agreement with the outcomes of the experimental tests shown in Ref. [23].

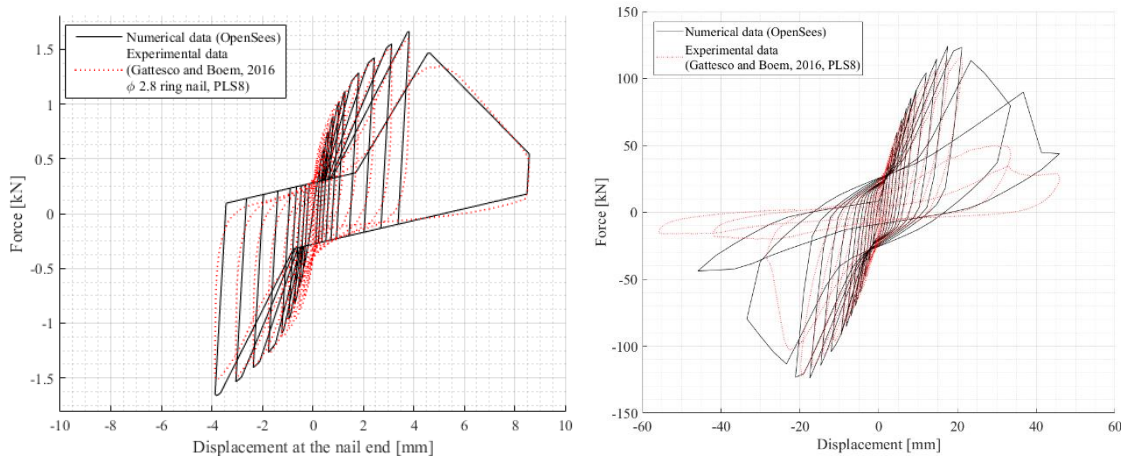


Figure 3. Identification of SAWS model parameters for the sheathing-to-framing connections and validation of the FE model: comparison between experimental and identified force-displacement curves of a single nail (left), comparison between experimental and predicted load-displacement curves for the reference wall (right).

Sensitivity analysis

Once identified the constitutive law of the sheathing-to-framing connections, the overall response in terms of hysteretic damping and racking capacity of different timber light-frame shear walls has been evaluated. A horizontal cyclic loading under displacement-controlled conditions has been applied. Aspect ratio (i.e., height-to-width ratio), horizontal and vertical nails spacing and number of vertical studs have been varied in order to quantify their influence. As pointed out in Ref. [8], the aspect ratio strongly influences the response of partially- and non-anchored walls. The contribution of shear deformation to storey displacements increases if

the base of the shear wall is significantly larger than its height, as pointed out in Ref. [29]. Conversely, if the base is about 30% of the height, then the flexural behaviour is dominant. By observing the overall behaviour of the wall in parallel with the local behaviour of each nail, the following definitions are given:

- 1) the Life Safety Limit State is recognized to occur in correspondence of the racking strength peak, when all nails along the perimeter framing elements are yielded;
- 2) the Collapse Limit State is recognized to occur when the most stressed nail, usually at the bottom corner, reaches its failure displacement.

As a consequence, the following criterion was adopted: the amount of dissipated energy is evaluated from the force-displacement curve of a certain configuration of shear wall once the first nail reached a resistance decrement equal to 65%, according to the experimental data in Ref. [23].

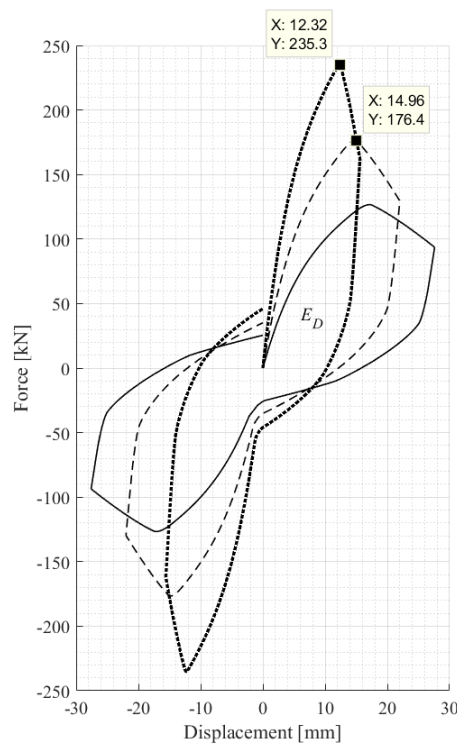


Figure 4. Influence of aspect ratio on the overall response of the wall:
1.4 (solid line, reference configuration),
1.0 (dashed line, square wall),
0.7 (dotted line, squat wall).

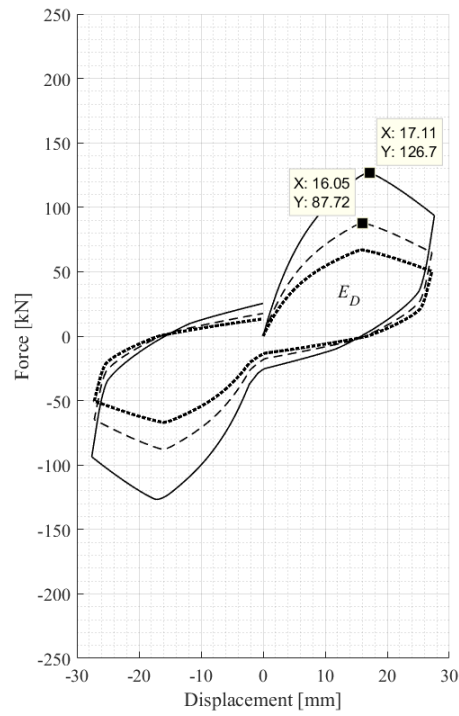


Figure 5. Influence of nails spacing on the overall response of the wall:
50 mm spacing (solid line, reference configuration),
75 mm spacing (dashed line),
100 mm spacing (dotted line).

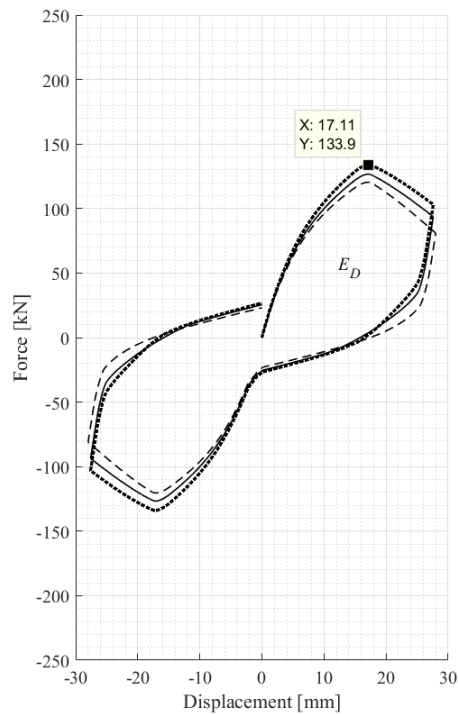


Figure 6. Influence of number of vertical studs on the overall response of the wall:
3 studs (dashed line),
4 studs (solid line, reference configuration),
5 studs (dotted line).

As shown in Fig. 4, the lower the aspect ratio, the higher the hysteretic damping and racking capacity. The relative rigid rotation of the sheathing panel with respect to the frame mostly stresses the nails near the corners, whereas the others remain in the elastic range. By varying horizontal and vertical nails spacing from 50 mm to 100 mm, it is possible to observe a reduction of the racking capacity (Fig. 5). Particularly, the reduced number of nails on the perimeter studs makes the overall system more flexible but leaving the ductility unchanged. Finally, the influence of the studs' number is shown in Fig. 6.

Equivalent viscous damping

A simplified way to take into account ductility or dissipative capacity of a structure in modern seismic design codes is based on the reduction of the elastic spectrum demand. The scaling of the elastic spectrum is function of an additional equivalent viscous damping ξ_{eq} , which is computed as follows:

$$\xi_{eq} = \frac{E_D}{4\pi E_{s0}} \quad (2)$$

where E_D is the dissipated energy in a single cycle, normalized to the elastic strain energy in a half cycle, E_{s0} . The total equivalent viscous damping ξ_{tot} is obtained by adding the inherent viscous damping $\xi_{0.05}$ (equal to 5%):

$$\xi_{tot} = \xi_{0.05} + \xi_{eq} \quad (3)$$

The reduced spectrum is then obtained by using the damping correction factor η , which is computed as follows [3]:

$$\eta = \sqrt{\frac{10}{5 + \xi_{tot}}} \quad (4)$$

The equivalent viscous damping as function of the drift value (defined as the ratio between horizontal displacement and wall height) for different wall configurations is given in Fig. 7.

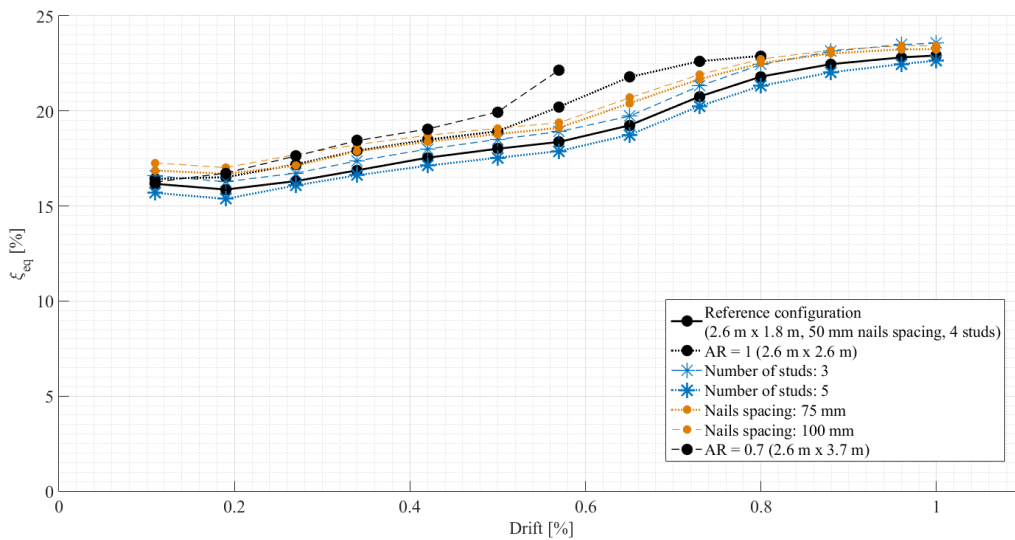


Figure 7. Equivalent viscous damping as function of drift for different wall configurations. The solid black line indicates the reference configuration: aspect ratio equal to 1.4, nails spacing equal to 50 mm and 4 vertical studs.

As shown in Fig. 7, the linear variation observed by Ref. [7] for wood framed buildings is fairly confirmed for timber walls as well. It can also be inferred that the energy dissipation strongly depends on the aspect ratio, thereby confirming the results in Ref. [8]. However, it is worth highlighting that the larger the wall size, the larger the number of vertical studs and the overall number of vertical nails: the resultant global system is stiffened. This is due to the fact that a lower plasticity level is reached by increasing the number of vertical studs and, consequently, the overall number of nails. Conversely, the number of yielded nails grows by reducing the number of vertical studs, and thus a higher amount of dissipated energy is achieved. Particularly, a higher amount of nails, especially on the intermediate studs, makes the overall system more resistant, preventing buckling of the sheathing panel, without providing a contribution in terms of plastic deformation and energy dissipation. By reducing the nails spacing, a stiffer wall is observed but a slightly lower value of the equivalent viscous damping is obtained. For the reference configuration (aspect ratio equal to 1.4, nails spacing equal to 50 mm, 4 vertical studs), the equivalent viscous damping is about 23%. This means that the value of the total equivalent viscous damping required to estimate the reduction of the elastic demand spectrum is about of 28% (assuming an inherent viscous damping equal to 5%). Hence, the resulting η factor is about 0.55. A summary of the results related to the variation of racking capacity, total equivalent viscous damping and damping factor, with respect to the geometric input parameters used in the parametric analyses, is shown in Fig. 8.

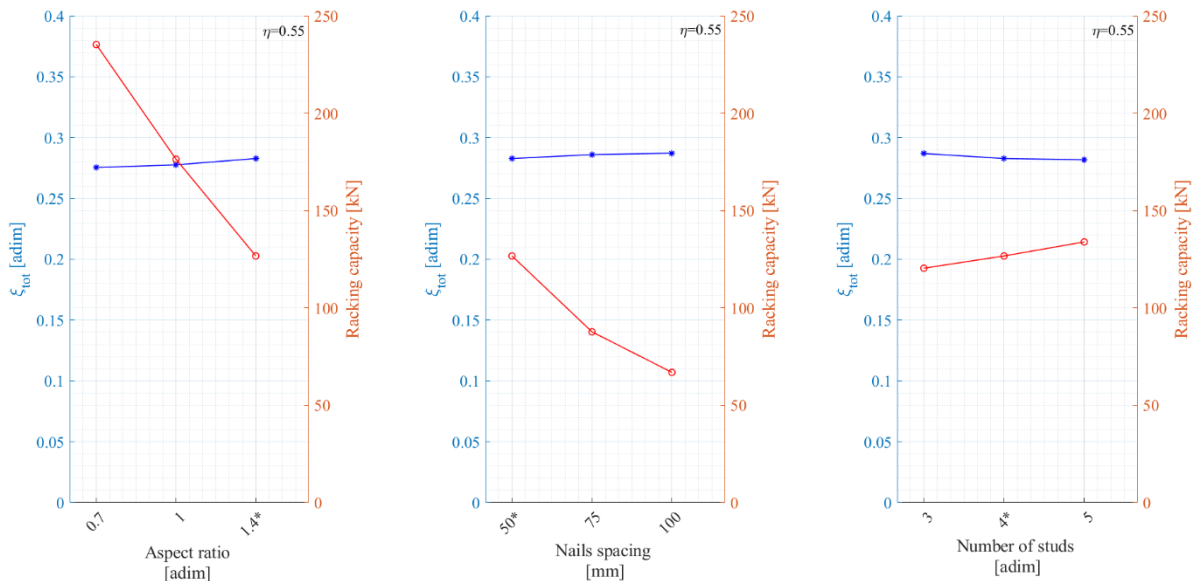


Figure 8. Variation of racking capacity (red), total equivalent viscous damping (blue) and damping factor η for different values of aspect ratio, nails spacing and number of vertical studs.
The reference configuration has aspect ratio equal to 1.4, nails spacing equal to 50 mm and 4 vertical studs
(the corresponding values are marked with the symbol *).

Conclusions

In this work an original parametric FE model for timber light-frame shear walls implemented in OpenSees is presented. It includes the following elements: *i*) elastic beam column for the framing elements, *ii*) non-linear coupled zero-length link elements for the sheathing-to-framing

connections, *iii*) zero-length elements for joints connecting framing elements, and *iv*) ShellMITC4 elements for the sheathing panels.

The numerical model has been validated considering the data carried out from experimental tests presented in Ref. [23]. The SAWS mechanical model, originally proposed by Ref. [26] and developed in Ref. [27], has been calibrated using the available experimental data and then implemented in order to simulate the behaviour of a single nail. Parametric analyses have been used to assess the influence of some basic design variables on racking capacity and equivalent viscous damping. Final results have demonstrated that the proposed model can be effectively used to carry out non-linear analyses and to calibrate the damping factor η in use within the Capacity Spectrum Method.

Acknowledgments

The authors would like to thank Prof. Natalino Gattesco (University of Trieste, Italy) and Dr. Ingrid Boem (University of Trieste, Italy) for sharing the data of the experimental tests reported in Ref. [23].

References

- [1] Peterson, J. (1983) Bibliography on lumber and wood panel diaphragms, *Journal of Structural Engineering*, **109**(12), 2838-2852.
- [2] Jorissen, A., Fragiaco, M. (2011) General notes on ductility in timber structures, *Engineering structures*, **33**(11), 2987-2997.
- [3] EN1998-1 (2004) Eurocode 8: Design of structures for earthquake resistance-Part 1.1: General rules, seismic actions and rules for buildings, CEN – European Committee for Standardization.
- [4] Fajfar, P. (1999) Capacity spectrum method based on inelastic demand spectra, *Earthquake Engineering & Structural Dynamics*, **28**(9), 979-993.
- [5] Freeman, S.A. (1978) Prediction of response of concrete buildings to severe earthquake motion, in *Douglas McHenry International Symposium on Concrete and Concrete Structures*, SP-55, 589-606, American Concrete Institute, Detroit, Michigan.
- [6] Freeman S.A., The Capacity Spectrum Method as a tool for seismic design, *Proceedings of the 11th European Conference on Earthquake Engineering*, Paris, September 6–11th, 1998.
- [7] Filiatrault, A., Isoda, H., Folz, B. (2003) Hysteretic damping of wood framed buildings, *Engineering structures*, **25**(4), 461-471.
- [8] Salenikovich, A., Dolan, J. D., The racking performance of light-frame shear walls with various tie-down restraints, *Proceedings of the 6th World Conference on Timber Engineering*, Canada, 2000.
- [9] Salenikovich, A., Dolan, J. D. (2003) The racking performance of shear walls with various aspect ratios. Part I. Monotonic tests of fully anchored walls, *Forest Products Journal*, **53**(10), 65-73.
- [10] Dhonju, R., D'Amico, B., Kermani, A., Porteous, J., Zhang, B. (2017) Parametric evaluation of racking performance of platform timber framed walls, *Structures*, **12**, 75-87.
- [11] McKenna, F. and Fenves, G. (2001) The OpenSees Command Language Manual: version 1.2, Pacific Earthquake Engineering Center, Univ. of Calif., Berkeley, USA.
- [12] Porteous, J., Kermani, A. (2013) *Structural timber design to Eurocode 5*, Blackwell Publishing Ltd.
- [13] Kolb J. (2008) *Systems in Timber Engineering, Loadbearing Structures and Component Layers*, Birkhauser, Basel – Boston – Berlin.
- [14] Mehta, M., Scarborough, W. Armpriest, D. (2009) *Building Construction: principles, materials and system*, Prentice Hall – Pearson, New Jersey.
- [15] Wacker, J. P. (2010) *Use of Wood in Buildings and Bridges*, in Wood Handbook – Wood as an Engineering Material, chapt. 17, Forest Products Laboratory – United States Department of Agriculture Forest Service – Madison, Wisconsin.

- [16] Varela, S., Correal, J., Yamin, L., Ramirez, F. (2012) Cyclic performance of glued laminated Guadua bamboo-sheathed shear walls, *Journal of structural engineering*, **139**(11), 2028-2037.
- [17] Wang, J., Demartino, C., Xiao, Y., Li, Y. (2018) Thermal insulation performance of bamboo- and wood-based shear walls in light-frame buildings, *Energy and Buildings*, **168**, 167-179.
- [18] Källsner, B. and Girhammar, U. A. (2009) Analysis of fully anchored light-frame timber shear walls-Elastic model, *Materials and structures*, **42**(3), 301-320.
- [19] Sartori, T. and Tomasi, R. (2013) Experimental investigation on sheathing-to-framing connections in wood shear walls, *Engineering Structures*, **56**, 2197-2205.
- [20] Humbert, J., Boudaud, C., Baroth, J., Hameury S., Daudeville, L. (2014) Joints and wood shear walls modelling I: constitutive law, experimental tests and FE model under quasi-static loading, *Engineering Structures*, **65**, 52-61.
- [21] Germano, F., Metelli, G., Giuriani, E., (2015) Experimental results on the role of sheathing-to-frame and base connections of a European timber framed shear wall, *Construction and Building Materials*, **80**, 315-328.
- [22] Gavric I, Fragiaco M, Ceccotti A. (2015) Cyclic behaviour of typical metal connectors for cross-laminated (CLT) structures, *Materials and structures*, **48**(6), 1841-1857.
- [23] Gattesco, N. and Boem, I. (2016) Stress distribution among sheathing-to-frame nails of timber shear walls related to different base connections: experimental tests and numerical modelling, *Construction and Building Materials*, **122**, 149-162.
- [24] UNI EN 14081-1 (2016) Timber structures–Strength graded structural timber with rectangular cross section–Part 1: General requirements.
- [25] UNI EN 338 (2016) Structural wood – strength classes.
- [26] Foschi, R. O. (1977) Analysis of wood diaphragms and trusses, Part 1:Diaphragms, *Canadian Journal of Civil Engineering*, Ottawa, **4**(3), 345–352.
- [27] Folz, B. and Filiatrault, A. (2001) Cyclic analysis of wood shear walls, *Journal of Structural Engineering*, **127**(4), 433-441.
- [28] Quaranta, G., Monti, G., Marano, G. C. (2010) Parameters identification of Van der Pol-Duffing oscillators via particle swarm optimization and differential evolution, *Mechanical Systems and Signal Processing*, **24**(7), 2076-2095.
- [29] Anil, Ö., Togay, A., İşleyen, Ü. K., Döngel, N., Söğütü, C. (2017) Effect of timber type and nail spacing on the hysteretic behavior of timber-framed shear walls with openings, *International Journal of Civil Engineering*, **16**(6), 629-646.

Computation of Deformable Image Registration by Meshless Kernel-based Collocation Method

*†S. M. Wong¹, K. S. NG¹ and T. S. Li¹

¹The school of Science and Technology, the Open University of Hong Kong

*Presenting and †Corresponding author: anwong@ouhk.edu.hk

Abstract

The deformable image registration (DIR) is a class of optimization methods and is commonly used to compute the spatial displacements between different regions of interest. When the process of an anatomical change whose mechanism is not comprehended analytically, the change of such system can be studied through tracing the spatial displacements between the static images and the deformed images. The diffusion of spatial transformation between the profiles of target object between n^{th} treatment and $(n + 1)^{th}$ treatment can be estimated. The volumetric analysis and the positional changes of the target object can be monitored and adaptive medical strategies can be applied during the course of treatment.

The present paper aims to set up a kernel-based collocation scheme in order to obtain a possible optimum deformable image registration. The proposed computational scheme will incorporate the meshless reciprocal multiquadric function. Its meshless configuration provides with a simple formulation and also significantly reduces the time used and the cost in setting up the computational algorithm. The computational result of the proposed scheme is verified with a real liver cancer medical imaging dataset. The influence of collocating different number of data points on the performance of the numerical results will also be explored and analyzed in this paper.

Keywords: Deformable image, Meshless RBF, Kernel collocation.

Introduction

Deformable image registration (DIR) is a common process in the medical image analysis. The DIR helps to trace the tumor growth, confine the treatment to a well-defined target region and avoid damage to healthy organs. All these factors are very important for medical treatment decision. For instance, the intended outcome of the medical treatment is to inhibit the tumor growth and minimize the side effects to the patient. In order to optimize the image-guided radiation therapy (IGRT) as discussed in [1], a partition deformable image registration (DIR) is used to delineate the region on the computed tomography (CT) images obtained during the radiation treatment.

The purpose of the deformable image registration is to find out the displacements of the deformed target region. The precision of the measurement of the region of interest depends tremendously on the accuracy of the registered image. Sotira et al. [2] had carried out a detailed survey and reviewed of the recent developments on the field of DIR. Over the last decade, DIR has been recognized a significant influence on the medical field, thus this topic still remains a challenging topic. A number of researchers from different disciplines contributed efforts to find optimal performance of deformable image registration.

The basic idea of the image registration problem is formulated as an optimization model. Suppose the target object of interest has undergone some changes from times t_0 to t_1 . Given the static image $s(x, y)$ and deformed image $m(x, y)$ at time t_0 and t_1 respectively, the spatial displacements between the static image $s(x_i, y_i)$ and the deformed image $m(x_i, y_i)$ can be predicted using an appropriate DIR scheme.

The proposed scheme presented in this study is applied to solve the classical Demon's DIR model which is a well established DIR model. The system of Demon's model has taken account of those dynamic and physical effects that occurred from the process of transformation.

The computational scheme is constructed based on the kernel radial basis function collocation method and use it to approximate the spatial displacements between the two deformable images. Radial basis function (RBF) is a novel interpolation method and has been proved to be effective in solving various kinds of differential equations. The meshless configuration of RBF also offers us with an ease and flexible design of the computational algorithm.

The Classical Deformable Image Registration Model

Several forms of deformable image registration models are established in different numerical methods. One of the classical DIR models is called Demon's algorithm developed by [3] in 1996. The basic deformation model are derived from Navier nonlinear elastic model. Let $\mathbf{D}(x, y) = (u_x, u_y)^T$ be the displacements matrix between the static and deformed images. The results of \mathbf{D} can be generated by the differential equations equations

$$\mathbf{D}(x, y) = \frac{(\mathbf{m} - \mathbf{s})\nabla\mathbf{s}(x, y)}{|\nabla\mathbf{s}(\mathbf{x}, \mathbf{y})|^2 + (\mathbf{m} - \mathbf{s})^2}, \quad (1)$$

where \mathbf{m} is moving image and \mathbf{s} is the static image, $(\mathbf{m} - \mathbf{s})$ is the differential forces between the moving image and deformed images. $\nabla\mathbf{s}$ is the gradient operator of the static image defined by

$$\nabla\mathbf{s}(x, y) = \left(\frac{\partial\mathbf{s}(x, y)}{\partial x}, \frac{\partial\mathbf{s}(x, y)}{\partial y} \right).$$

The equation in (1) can be rearranged to a homogeneous partial differential equation as

$$\mathbf{D} (|\nabla\mathbf{s}|^2 + (\mathbf{m} - \mathbf{s})^2) - (\mathbf{m} - \mathbf{s})\nabla\mathbf{s} = \mathbf{0}$$

subject to the given initial conditions $\mathbf{s}^0 = \mathbf{0}$ and $\mathbf{m}^0 = \mathbf{0}$.

Cachier *et al* [4] in 1999 revised model to improves the registration convergence rate and stability. The improved model is defined by

$$\mathbf{D} = \frac{(\mathbf{m} - \mathbf{s})\nabla\mathbf{s}(x, y)}{|\nabla\mathbf{s}(x, y)|^2 + \xi^2 |(\mathbf{m} - \mathbf{s})|^2} + \frac{(\mathbf{m} - \mathbf{s})\nabla\mathbf{m}(x, y)}{|\nabla\mathbf{s}(x, y)|^2 + \xi^2 |(\mathbf{m} - \mathbf{s})|^2}. \quad (2)$$

This revised model includes the image edge forces of the deformed image. The normalization factor ξ is added to adjust the force strengths. This attempts to normalize the relations between the moving and static image so as to improve the image registration.

Kernel-based Collocation Method

This paper discusses a class of kernel-based approximation methods in the form of radial basis function (RBF). The proposed scheme will be used to solve the deformable image registration model. The kernel-based approximation method has been refined and diversified for facilitating the needs of various types of differential equations. The RBF approximation was originally devised for scattered geographical data interpolation by Hardy [5], who introduced a class of RBF called multiquadric function in the early 1970's.

This study aims to determine an optimum spatial displacements between the deformable images. Let the displacements of DIR model be $\mathbf{D}(x, y) = (u_x, u_y)^T$ between the static and deformed images defined in the equation (1). The basic idea of the kernel-basis RBF interpolation is used to approximate this unknown displacements $\{\mathbf{D}(\mathbf{x}) : \mathbf{x} \in \Omega\}$ by a RBF interpolant, say $\{\mathbf{f}(\mathbf{x}) : \mathbf{x} \in \Omega\}$ at a given set of N distinct points $X = \{\mathbf{x}_i \in \Omega : i = 1, 2, \dots, N\}$.

The general form of the kernel-based RBF interpolation is a finite linear combination by the following equation

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= \alpha_1 \phi_1(\|\mathbf{x} - \tilde{\mathbf{x}}_1\|) + \dots + \alpha_N \phi_N(\|\mathbf{x} - \tilde{\mathbf{x}}_N\|) \\ &= \sum_{i=1}^N \alpha_i \phi(\|\mathbf{x} - \tilde{\mathbf{x}}_i\|), \quad \mathbf{x} \text{ and } \tilde{\mathbf{x}}_i \in \Omega, \end{aligned} \quad (3)$$

where ϕ refers to a specific choice of RBF functions that is solely dependent on the Euclidean distance $r = \|\mathbf{x} - \tilde{\mathbf{x}}_i\|$ between \mathbf{x} and a fixed centre $\tilde{\mathbf{x}}_i \in \Omega$. The unknown coefficients $\{\alpha_i : i = 1, 2, \dots, N\}$ can be determined by collocating

$$\mathbf{f}(\mathbf{x}_i) = \mathbf{D}(\mathbf{x}_i), \quad \text{for } i = 1, 2, \dots, N, \quad (4)$$

at a set of N distinct data points $\{\mathbf{x}_i \in \Omega, i = 1, 2, \dots, N\}$. This yields a system of linear equations which can be expressed in the following matrix form

$$[\mathbf{A}_\phi] [\boldsymbol{\alpha}] = [\mathbf{D}], \quad (5)$$

where

$$[\boldsymbol{\alpha}] = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} \quad \text{and} \quad [\mathbf{D}] = \begin{bmatrix} D(\mathbf{x}_1) \\ D(\mathbf{x}_2) \\ \vdots \\ D(\mathbf{x}_N) \end{bmatrix}$$

are $N \times 1$ column matrices. The $[\mathbf{A}_\phi] = [\phi(\|\mathbf{x}_i - \mathbf{x}_j\|)]_{i,j=1}^N$ is an $N \times N$ matrix given by

$$[\mathbf{A}_\phi] = \begin{bmatrix} \phi(\|\mathbf{x}_1 - \mathbf{x}_1\|) & \phi(\|\mathbf{x}_1 - \mathbf{x}_2\|) & \dots & \phi(\|\mathbf{x}_1 - \mathbf{x}_N\|) \\ \phi(\|\mathbf{x}_2 - \mathbf{x}_1\|) & \phi(\|\mathbf{x}_2 - \mathbf{x}_2\|) & \dots & \phi(\|\mathbf{x}_2 - \mathbf{x}_N\|) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\|\mathbf{x}_N - \mathbf{x}_1\|) & \phi(\|\mathbf{x}_N - \mathbf{x}_2\|) & \dots & \phi(\|\mathbf{x}_N - \mathbf{x}_N\|) \end{bmatrix}. \quad (6)$$

Generally, the interpolation points in interior and boundaries are distinct and a specific choice of RBF $\phi(\|\mathbf{x} - \mathbf{x}_j\|)$ is positive definite, the matrix $[\mathbf{A}_\phi]$ is always non-singular, so

the linear system in (5) has a unique solution as proved by [6]. The unknown coefficients $[\boldsymbol{\alpha}]$ can then be determined uniquely by solving the system of linear equations by

$$[\boldsymbol{\alpha}] = [\mathbf{A}_\phi]^{-1} [\mathbf{D}].$$

The approximated displacement matrix $[\mathbf{D}]$ can be evaluated once the unknown coefficients $[\boldsymbol{\alpha}]$ are found.

The most popular types of radial basis functions are listed below:

$$\phi(r_j) = \left\{ \begin{array}{lll} r_j^3 & \text{Cubic} & (a) \\ (r_j^2) \log r_j, & \text{Thin plate splines in } R^2 & (b) \\ e^{-\sigma r_j^2}, & \text{Gaussian, } \sigma > 0 & (c) \\ (r_j^2 + c^2)^{\frac{1}{2}}, & \text{Multiquadric } \mathbb{R} & (d) \\ (r_j^2 + c^2)^{-\frac{1}{2}}, & \text{Reciprocal multiquadric } \mathbb{R} & (e) \end{array} \right\}, \quad (7)$$

where $\{r_j = \|\mathbf{x} - \mathbf{x}_j\|, j = 1, 2, \dots, N\}$ is the Euclidean distance between \mathbf{x} and $\mathbf{x}_j \in \mathbb{R}^d$ and c^2 is called the shape parameter of the functions in (d) & (e). This shape parameter uses to control the fitting of a smooth surface to the data and could be greatly influence of the intended results. A recent study by Luh [7] in 2012 has developed a concrete function which could help to determine an optimal shape parameter.

The RBF collocation scheme is a well defined efficient scheme, however, there are two crucial issues of RBF always be questioned.

- (1) Is there always a unique solution to the system of equations in (5)?
- (2) What type of conditions on $\{\phi(r_{i,j}) : i, j = 1, \dots, N\}$ can guarantee the invertibility of the interpolation matrix $[\mathbf{A}_\phi]$?

Answering these questions is not simple. A full description can involve a number of properties and conditions. In this present study, the kernel-based RBF model is formulated by adding a finite polynomials $\{q_k(\mathbf{x}), k = 1, 2, \dots, M\}$ into the interpolation system in (3). The purpose is to avoid having singularity in solving the system matrix. The RBFs interpolant in (3) is now extended to the sum of finite series equation as below:

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^N \alpha_i \phi(\|\mathbf{x} - \tilde{\mathbf{x}}_i\|) + \sum_{k=1}^M b_k q_k(\mathbf{x}), \quad \mathbf{x} \in R^2, \mathbb{R}^2, \quad 0 \leq M < N, \quad (8)$$

where $\{\alpha_i\}$ and $\{b_k\}$ are the unknown coefficients to be determined. Given a set of N distinct nodes $X = \{\mathbf{x}_i \in \Omega, i = 1, 2, \dots, N\} \subseteq \mathbb{R}^d$, the approximation function in (8) will definitely produce a unique solution if the system satisfies the following condition

$$\mathbf{f}(\mathbf{x}_i) = \mathbf{D}(\mathbf{x}_i), \quad i = 1, 2, \dots, N \quad (9)$$

and the constraints

$$\sum_{i=i}^N \alpha_i q_k(\mathbf{x}) = 0, \quad k = 1, 2, \dots, M \text{ and } i = 1, 2, \dots, N.$$

The resulting system can be written in matrix form as,

$$\begin{bmatrix} \mathbf{A}_\phi & \mathbf{Q} \\ \mathbf{Q}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{D} \\ \mathbf{0} \end{bmatrix}, \quad (10)$$

where $[\mathbf{A}_\phi]$ is the same coefficient matrix and $[\boldsymbol{\alpha}]$ is the unknowns as defined in (5). The matrix $[\mathbf{Q}] = [q_k(\mathbf{x}_i)]$ is an $N \times M$ matrix and $[\mathbf{b}]$ is an $M \times 1$ column matrix for the finite polynomials. They are defined as

$$[\mathbf{Q}] = \begin{bmatrix} q_1(\mathbf{x}_1) & q_2(\mathbf{x}_1) & \cdots & q_M(\mathbf{x}_1) \\ q_1(\mathbf{x}_2) & q_2(\mathbf{x}_2) & \cdots & q_M(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ q_1(\mathbf{x}_N) & q_2(\mathbf{x}_N) & \cdots & q_M(\mathbf{x}_N) \end{bmatrix}, \quad [\mathbf{b}] = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix}.$$

The interpolation problem in equation (9) is solvable if the matrix of this system is

$$[\tilde{\Phi}] = \begin{bmatrix} \mathbf{A}_\phi & \mathbf{Q} \\ \mathbf{Q}^T & 0 \end{bmatrix}$$

is non-singular. The unknowns $\begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{b} \end{bmatrix}$ can then be determined by inverting the coefficient matrix as

$$\begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_\phi & \mathbf{Q} \\ \mathbf{Q}^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{D} \\ \mathbf{0} \end{bmatrix}.$$

The values of the displacement function (\mathbf{D}) can be determined uniquely once the unknown vectors $[\boldsymbol{\alpha}]$ and $[\mathbf{b}]$ are determined.

The Computational Algorithm

The computational scheme of this paper is established according to the basic idea of Demon's algorithm. The equations of Demon's model are discretized by using finite difference scheme. The incremental displacements $\mathbf{D}^j(x, y)$ for the j^{th} iteration is simulated iteratively by the following equation:

$$\mathbf{D}^j(x, y) = \mathbf{D}^{j-1}(x, y) - \frac{(\mathbf{m}^{j-1}(x, y) - \mathbf{s}^0(x, y)) \nabla \mathbf{s}^0}{\|\nabla \mathbf{s}^0(x, y)\|^2 + [\mathbf{m}^{j-1}(x, y) - \mathbf{s}^0(x, y)]^2} \quad (11)$$

for $j = 1, 2, \dots, N$. In the present algorithm, where the result of displacement $\mathbf{D}^{j-1}(x, y)$ is generated by using the kernel RBF approximation by

$$\mathbf{D}^{j-1}(x, y) = \left(\sum_{i=1}^N \alpha_i^{j-1} \phi(\|\mathbf{r}\|) + \sum_{k=1}^M b_k q_k^{j-1}(\mathbf{x}) \right),$$

where $\|\mathbf{r}\| = \|\mathbf{x} - \tilde{\mathbf{x}}_j\|$ is the Euclidean distance between \mathbf{x} and $\mathbf{x}_j \in \mathbb{R}^2$. The computation requires the input of the following initial values

$$\begin{aligned} \mathbf{D}^0(x, y) &= \mathbf{0}, \\ \mathbf{m}^0(x, y) &= \tilde{\mathbf{m}}^0(x, y), \\ \mathbf{s}^0(x, y) &= \tilde{\mathbf{s}}^0(x, y). \end{aligned}$$

The spatial derivatives $\nabla \mathbf{s}^0 = (\nabla \mathbf{s}_x^0, \nabla \mathbf{s}_y^0)$ of the static image are the distance between two neighbourhood pixels is taken to be 1. Therefore, the spatial gradients $\nabla \mathbf{s}_x^0$ and $\nabla \mathbf{s}_y^0$ can be determined according to the following $n \times n$ matrices.

$$\nabla \mathbf{s}_x^0 = \begin{bmatrix} s_{1,2}^0 - s_{1,1}^0 & s_{1,3}^0 - s_{1,2}^0 & \cdots & s_{1,n}^0 - s_{1,n-1}^0 & s_{1,n}^0 - s_{1,n-1}^0 \\ \vdots & \cdots & \cdots & \vdots & \vdots \\ s_{n,2}^0 - s_{n,1}^0 & s_{n,3}^0 - s_{n,2}^0 & \cdots & s_{n,n}^0 - s_{n,n-1}^0 & s_{n,n}^0 - s_{n,n-1}^0 \end{bmatrix},$$

and

$$\nabla \mathbf{s}_y^0 = \begin{bmatrix} s_{2,1}^0 - s_{1,1}^0 & s_{2,2}^0 - s_{1,2}^0 & \cdots & s_{2,n}^0 - s_{1,n}^0 \\ \vdots & \cdots & \cdots & \vdots \\ s_{n,1}^0 - s_{n-1,1}^0 & s_{n,2}^0 - s_{n-1,2}^0 & \cdots & s_{n,n}^0 - s_{n-1,n}^0 \\ s_{n,1}^0 - s_{n-1,1}^0 & s_{n,2}^0 - s_{n-1,2}^0 & \cdots & s_{n,n}^0 - s_{n-1,n}^0 \end{bmatrix}.$$

The deformed image \mathbf{m}^{*j} at the j^{th} iteration can be updated by substituting \mathbf{D}^{j-1} and \mathbf{m}^{*j-1} into the following forward iterative scheme

$$\begin{aligned} \mathbf{m}^{*j}(x, y) &= \mathbf{m}^{*j-1}(x, y) + \mathbf{D}^{j-1}(x, y) \nabla \mathbf{s}^0, \\ &= \mathbf{m}^{*j-1}(x, y) + \left(\sum_{i=1}^N \alpha_i^{j-1} \phi(\|\mathbf{r}\|) + \sum_{k=1}^M b_k q_k^{j-1}(\mathbf{x}) \right) \nabla \mathbf{s}^0. \end{aligned} \quad (12)$$

In the numerical experiments, a normalized global reciprocal multiquadric function is used. The numerical results could be improved and more stable by adding a weighted factor λ^2 . The normalized reciprocal MQ is defined by

$$\phi(\|x - x_j\|) = \frac{1}{\sqrt{c^2 + \lambda^2 \left[\left(\frac{x - x_j}{511} \right)^2 + \left(\frac{y - y_j}{511} \right)^2 \right]}}.$$

Our previous study reported in paper [8], the chosen of different factor λ^2 had been considered in the numerical experiments, it was found that the optimal result reached at choose $\lambda = 100$.

In the computation, the required stopping criteria is set by

$$\|\mathbf{D}^j(x, y) - \mathbf{D}^{j-1}(x, y)\| < \varepsilon \leq 0.002,$$

where ε is the preset upper bound of iterative error. The quality of the computed image $\mathbf{m}^*(x, y)$ by RBFs is examined and compared with those results by Thirion and Cachier Demon's algorithm. For comparison, the root mean square errors of the deformed image \mathbf{m}^{*j} are analyzed by the equation

$$RMSE = \sqrt{\frac{1}{n} \sum_{x,y} \left(\frac{\mathbf{m}^*(x, y) - \mathbf{m}(x, y)}{\mathbf{m}(x, y)} \right)^2}. \quad (13)$$

To illustrate the computational algorithm, the MATLAB code of the deformable image registration kernel-based RBF scheme is outlined below:

Algorithm: Demon's algorithm using kernel based RBF
1. Input data : static image \mathbf{s}^0 , deformed image \mathbf{m} , error tolerance ε
2. Result : \mathbf{D}^N , \mathbf{m}^N , RMSE
3. Initialization: $\mathbf{D}^j \leftarrow \mathbf{0}$, $\mathbf{m}^0 \leftarrow \mathbf{m}$, $N \leftarrow 1$
4. Compute $\nabla \mathbf{s}^0$
while $ \nabla \mathbf{s}^0 ^2 + (\mathbf{m}^{N-1} - \mathbf{s}^0)^2 \neq \mathbf{0}$ and $\ \mathbf{D}^N - \mathbf{D}^{N-1}\ \geq \varepsilon$ and $N < 500$ do
$\mathbf{D}^j := \sum_{i=1}^n \left(\alpha_i^j \phi(\ \mathbf{r}\) \right) + \sum_{k=1}^M b_k q_k^j(\mathbf{x})$
Compute $\mathbf{D}^N \leftarrow \mathbf{D}^{N-1} - \frac{(\mathbf{m}^{N-1} - \mathbf{s}^0) \nabla \mathbf{s}^0}{ \nabla \mathbf{s}^0 ^2 + (\mathbf{m}^{N-1} - \mathbf{s}^0)^2}$
Solve α_i^N from \mathbf{D}^N
Compute $\mathbf{m}^N \leftarrow \mathbf{m}^{N-1} + \mathbf{D}^{N-1} (\nabla \mathbf{s}^0)$
$N = N + 1$
END

Case Study: Tracing Liver Cancer Growth

A real-life deformable image registration from a patient with liver cancer is used as a reference case study. One of the original static registered images is depicted in *Figure 1* at time t_1 and the deformed images are obtained from two different treatment periods at time t_2 and t_3 .

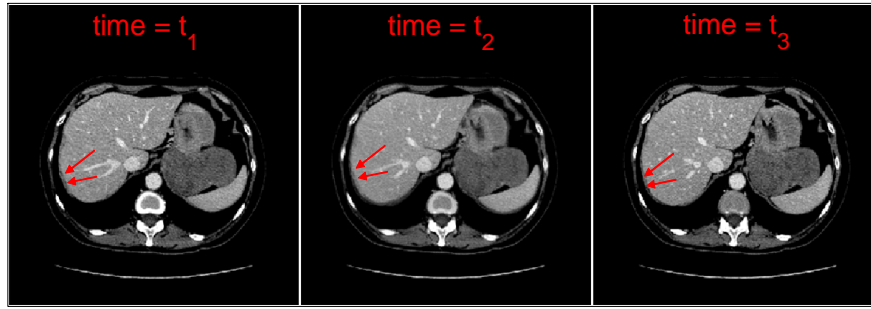


Figure 1 The growth of liver tumor as indicated by an arrow from different time t_1 , t_2 and t_3 , for $t_1 < t_2 < t_3$

The local ethics committee approval was obtained for a waiver of informed consent for retrospective analysis. The CT images were collected from collaborative hospital in year 2010.

The computational region is set up according to the original registered image. The geometrical structure of the original static image contains 512×512 pixels. The computational region is simplified by removing the insignificant backgrounds. These include the air with intensity equals to 0, and the bone with intensity equals to 1. The remaining valid pixels in the computational region are 82,633.

The objective of the present study is to use a kernel-based RBF collocation method to measure the tumor changes in time t_3 by using the known information given by the deformable image registration obtained from t_1 and t_2 . Similarly, the subsequent changes of tumor can be predicted by using the deformed relationship obtained in t_2 and t_3 .

Our model aims to give an optimum estimation of object change so that an appropriate medical treatment can be applied according to the stage of the cancer. For example, the cancer is at early stage, say at stage A when diagnosed, a complete medical treatment may be possible by means of using radiotherapy treatment.

Numerical Results

The effectiveness of RBF scheme is very dependent on the choice of number of collocation points on the studied region. It is generally known that the global RBF scheme will be very computational expensive if a high number of data points is collocated in the study region. In the numerical experiments, the level of accuracy and the efficiency are also considered. The computed image results based on $n = 128$ and on $n = 256$ collocation points are illustrated in *Table 1*. According to the experimental results, the RMSE of using 256 collocation points is clearly lower than that of using 128 collocation points. This indicates that a high degree of accuracy can also be obtained by using an intensive collocation points in the computational region.

Table 1: Reciprocal global MQ function with Finite Polynomial Collocation Scheme

Root Mean Square Errors (RMSE) for the image at time t_4			
Collocation points	Shape parameter:	RMSE	CPU times required
Case 1: 256 points	$c = 0.79$	0.0162	317 seconds
Case 2 : 128 points	$c = 0.65$	0.07273	87 seconds

To compare the performance of the proposed kernel collocation scheme, the Thirion and Cachier Demon's iterative finite difference schemes are set up to simulate the DIR result of the same problem. The analysis of RMSE of the Thirion and Cachier Demon's model is summarized in *Table 2*.

Table 2: Classical Demon's Algorithm with Finite Different Scheme

Root Mean Square Errors (RMSE) for the image at time t_4	
	RMSE
Thirion's Demon Model, 1996	0.0738
Cachier's Demon Model, 1999	0.0733

The rate of convergence of Thirion and Cachier Demon's algorithm are compared graphically in *Figure 2*.

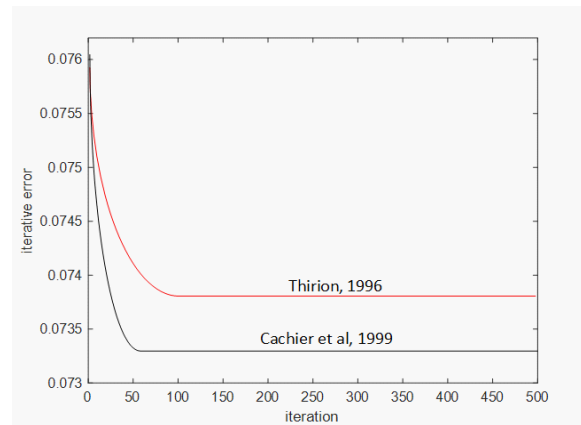


Figure. 2: The convergence of the two Demons algorithms.

The error analysis showed the optimum results of Thirion Demon's algorithm could be achieved in about 100 iterations, while Cachier Demon's algorithm produced an even faster convergent rate, the stopping criteria can be reached by less than 70 iterations. From the

comparison of the numerical results, we convinced that the kernel RBF collocation scheme could achieve a higher level of accuracy and stability than the Demon's finite difference scheme.

The simulation for tracing the propagation of tumor, the computed (registered) image at time t_3 and at time t_4 of the region of interest are illustrated in *Figure 3* and *Figure 4*.

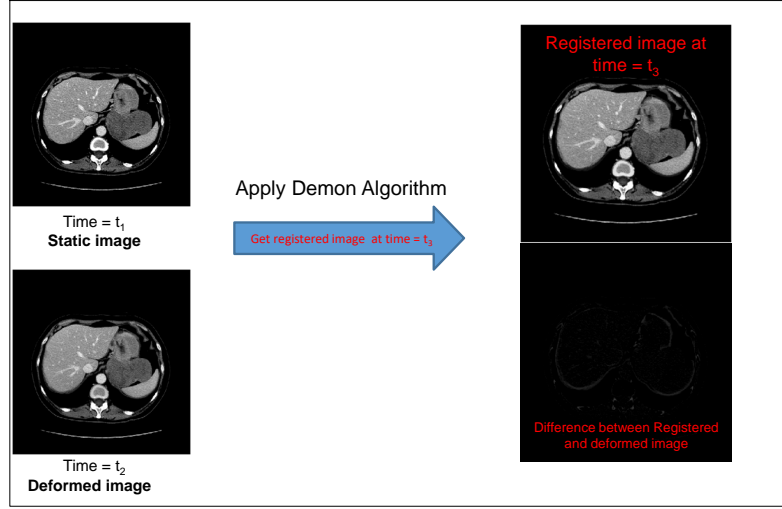


Figure 3 The computed (registered) image at time t_3

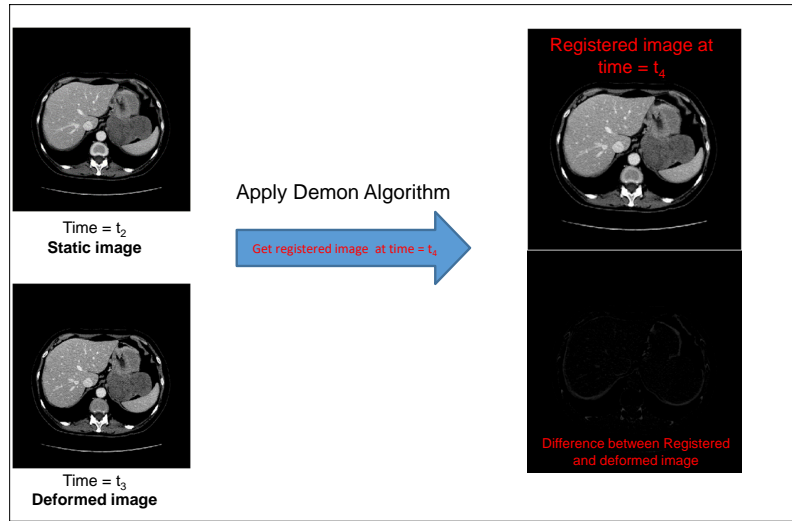


Figure 4 The computed (registered) image at time t_4

The displacements of these images can be identified through the deformation grid as shown in *Figure 4*. The difference between the computed registered image and the deformed CT image is also shown in the above *Figures*. As observed from the diagrams the overall difference is insignificant.

Conclusions

We have developed a kernel RBF collocation scheme to approximate the deformation image registration. In order to prevent the singularity problem, a finite polynomial term has been added into the RBF interpolant. The proposed algorithm based on a global reciprocal multiquadric function is applied to solve the classical Demon's DIR model. The computational region was set up based on a real-life deformable image registration in tracing the growth of a liver tumor. The computational efficiency and accuracy as well as the number of supporting points of the kernel approximation were explored. The computed images were calibrated with the real set of image registration.

In order to compare the capability of the proposed kernel RBF scheme, the two reference DIR models formulated by Thirion's and Cachier's Demon model were set up using finite difference scheme. The error analysis has shown that the proposed kernel RBF scheme produced a high level of accuracy. Regarding the issue of computational efficiency, our experiments showed that the computation of the kernel RBF scheme was clearly less efficient than the classical Demon's finite difference scheme. This was due to the defeat of the global supports of the kernel function setting. Indeed the numerical results lead an important role in the overall simulation of the DIR, thus the trade-off between numerical accuracy of the simulation results and computational efficiency has to be measured when using these schemes.

References

- [1] David Sarrut, Deformable registration for image-guided radiation therapy, *Zeitschrift für Medizinische Physik*, Volume 16, Issue 4 (2006) pp. 285-297.
- [2] A. Sotiras, C. Davatzikos, and N. Paragios, Deformable Medical Image Registration: A Survey, *IEEE Trans Med Imaging*, 32(7), Jul 2013 pp. 1153-90.
- [3] Thirion J. P. Non-rigid matching using demons. *Proc. Conf. Computer Vision and Pattern Recognition*, June 1996 pp. 245-251.
- [4] P. Cachier, X. Pennec, and N. Ayache, Fast Non Rigid Matching by Gradient Descent: Study and Improvements of the "Demons" Algorithm. RR-3706, INRIA 1999.
- [5] R. L. Hardy, Multiquadric equations of topography and other irregular surfaces, *J. Geophys, Res*, 176 (1971) pp. 1905-1915.
- [6] M. J. D. Powell, The theory of radial basis functions approximations in 1990, Chapter 3, *Wavelets, subdivision algorithms and radial basis functions*, ed. Will Light, Vol. II, Oxford University Press, (1990) pp. 105-210.
- [7] Lin-Tian Luh, "The shape parameter in the Gaussian functions", *Computers and Mathematics with Applications* 63 (2012) pp 687-694.
- [8] S. M. Wong, K. S. NG and T. S. Li, "Numerical Solution for Deformable Image Registration Using Radial Basis Functions", *J. Dynamics of Continuous, Discrete and Impulsive Systems Series B: Applications & Algorithms* 23 (2016) pp. 405-419.

Approximation of the parallel robot working area using the method of nonuniform covering

Mikhail Posypkin^{1, 1}, *Larisa Rybak², †Dmitry Malyshev², and Elena Gaponenko²

¹ Dorodnicyn Computing Centre, FRC CSC RAS, Russian Federation

²Belgorod State Technological University named after V.G. Shukhov, Belgorod, Russian Federation

*Presenting author: rl_bgtu@intbel.ru

†Corresponding author: malyshev.d.i@yandex.ru

Abstract

The work is devoted to optimization of the joint working area of 6-DOF relative manipulation mechanism, which include flat 3-RRR mechanism that rotates around the z axis and translates along the x and y axes of the lower platforms, and a tripod that provides movement of the tool along the z axis and its rotation around the x and y axes. The method of nonuniform covering is used to build the working area. As a result of the work of the method, external and internal approximations are obtained, given as a set of parallelepipeds. The paper presents the results of a computational experiment. The results of modeling the working areas of each of the relative manipulation modules and other parallel mechanisms, such as the flat 2-DOF DexTAR robot, are presented. The joint work area at moveable coordinate system is obtained. Moveable coordinate system is located in center of lower mechanism. The algorithms are implemented in C++ using the Snowgoose library.

Keywords: **Approximation, relative manipulation, working area, parallel robot.**

Introduction

The paper considers a parallel robot consisting of two mechanisms of relative manipulation (Fig. 1):

1. Upper mechanism is a tripod, which provide move along the z axis and rotate around the x and y axes.
2. Lower planar 3-RRR mechanism, which provide a rotation around the z axis and translational movement along the x and y axes.

Thus, the mechanism has 6 degrees of freedom. The mechanism will be designed to perform machining operations of details and other operations if the working tool is fixed to the mobile platform.

The main problem for parallel mechanisms is the small size of the working area and the presence of the singularity zones.

The task is approximation of the working area on the basis of optimization algorithms.

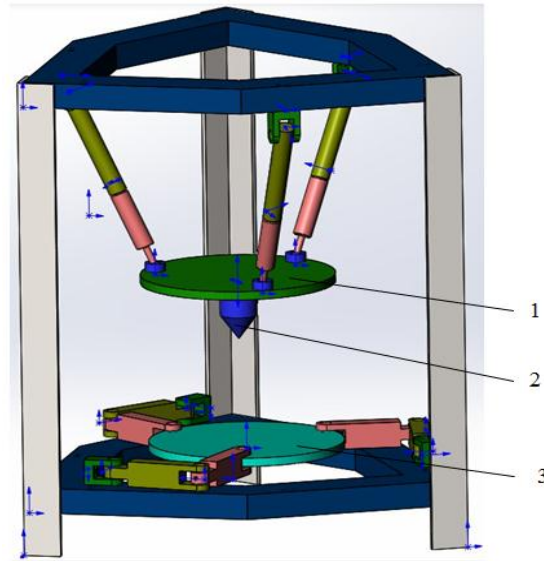


Figure 1. A 3D-model of the robot: 1 – the tool installation module, 2 –the tool, 3 – the detail machining module.

DexTAR working area approximation based on the method of nonuniform covering

The paper suggests the use of the method of nonuniform covering (Fig. 2) to solve the problem of approximation of the working area. The method is based on works [1,2]. This method allows approximating the solutions set of the equalities or inequalities systems describing the robot working area. The external and internal approximation sets are constructed. The internal approximation set is included in the set of solutions of the inequalities system. Both sets are represented as unions of n-dimensional parallelepipeds. Mathematical transformation of the coupling equations of some robots allows you to reduce the dimension, and hence the computation time. Using the parallelepipeds of large dimension avoids significant mathematical transformations, and then project them onto the coordinate axes necessary for the imaging.

Interval estimates using the developed Snowgoose library in C++ can be used to find the maximum and minimum of functions in parallelepipeds. The grid approximation method is used for cases of multiple occurrence of variables where errors can affect the result.

Method is development and tested on models of robots with 2 and 3 degrees of freedom within the framework of the Russian Science Foundation project, the agreement number 16-19-00148.

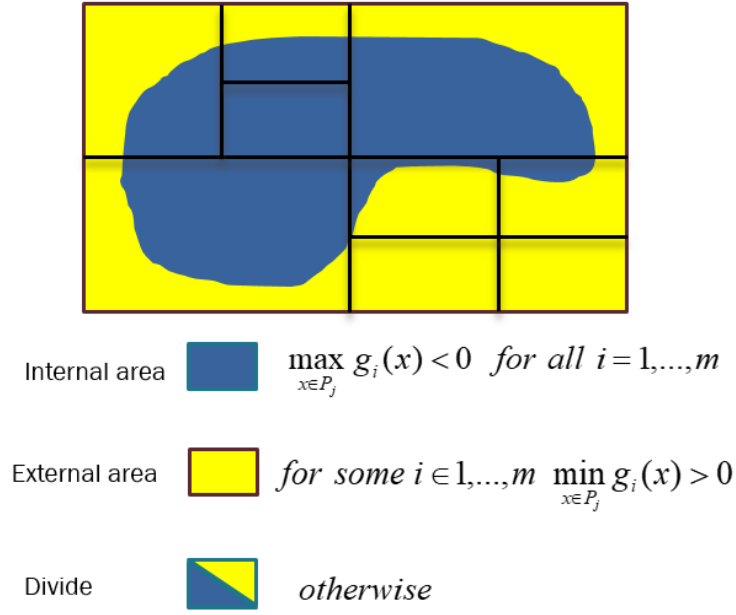


Figure 2. Illustration of the method.

The method was successfully applied to the 2-DOF robot DexTAR (Fig. 3). It is a planar parallel four-link mechanism, controlled by two engines. Result is presented in [3].

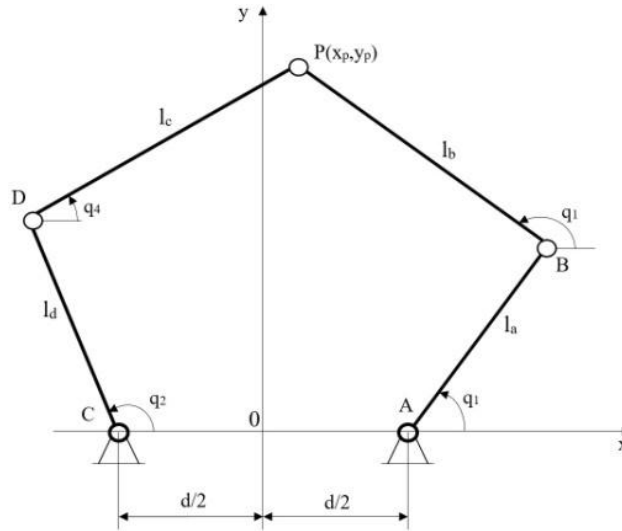


Figure 3. DexTAR robot scheme

Equations describing the motion:

$$\begin{cases} x_p = l_b \cdot \cos q_3 + l_a \cdot \cos q_1 + \frac{d}{2}, \\ x_p = -\frac{d}{2} + l_d \cdot \cos q_2 + l_c \cdot \cos q_4, \\ y_p = l_a \cdot \sin q_1 + l_b \cdot \sin q_3, \\ y_p = l_d \cdot \sin q_2 + l_c \cdot \sin q_4. \end{cases} \quad (1)$$

Equations have 6 variables:

- Input coordinates: q_1, q_2
- Output Coordinates: x_p, y_p

- Intermediate coordinates: q_3, q_4

Interval estimates are used to search for extrema. Under conditions of a single occurrence of variables in the expression, the interval estimates coincide with the extrema of the function on the parallelepiped, i.e. cannot be improved. However, the calculation time was 39 minutes 40 seconds, which is due to the large dimensionality of the problem (6-dimensional parallelepiped). Results of simulation is presented on Fig. 4, 5.

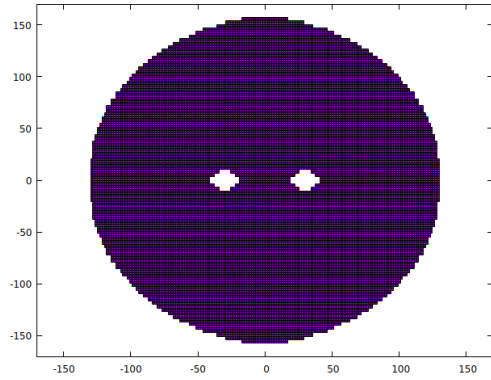


Figure 4. Results of simulation DexTAR working area

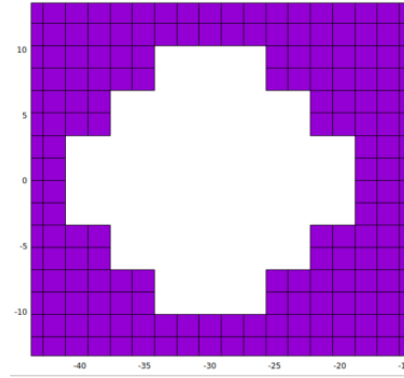


Figure 5. Singularity zone on an enlarged scale

Dimensionality of the problem was lowered and the grid approximation method was applied, due to the multiple occurrence of variables in expressions. The Fig. 6 shows the results of modeling the working area, Fig. 7 is the singularity zone on an enlarged scale. The calculation time, decreased by 10.35 times, while the accuracy of calculations increased by 10 times.

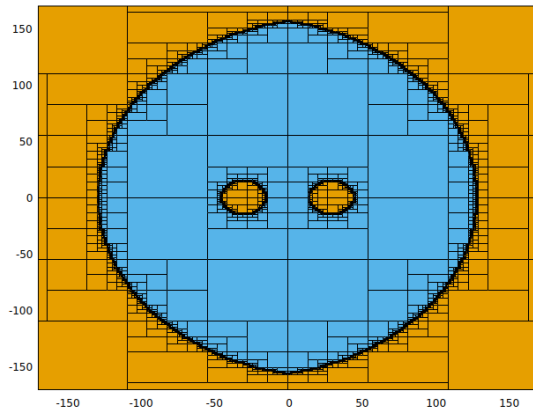


Figure 6. Results of simulation DexTAR working area

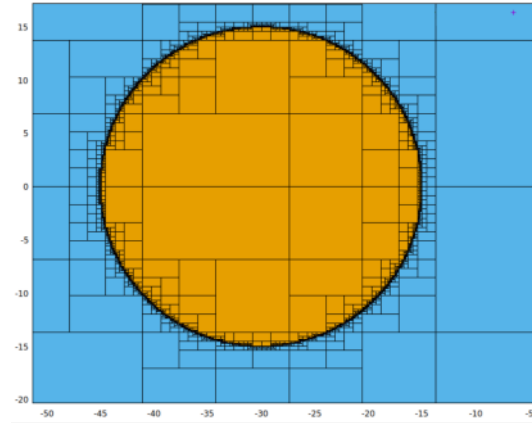


Figure 7. Singularity zone on an enlarged scale

Modeling the working area of the relative manipulation mechanism

Consider the construction of the working area of the robot with relative manipulation mechanisms (Fig. 1). Definition the working area of the upper mechanism - the tripod (Fig. 8) was developed within the framework of the Russian Science Foundation project, the agreement number 16-19-00148. The tripod consists of three link of variable length, which are connected by rotational hinges to the base and spherical hinges to the working platform. The base and the working platform are equilateral triangles. As a result of changing the

lengths of the rods, the working platform moves along the z axis and rotates along the x and y axes.

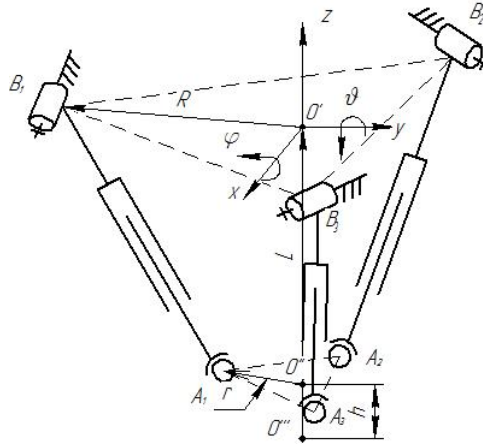


Figure 8. Tripod scheme

The center of the tripod mobile platform, in addition to the basic three basic degrees of freedom, has small displacements, which can be found by the following formulas [4]:

$\varphi = \tan^{-1}\left(\frac{\sin\psi\sin\theta}{\cos\psi+\cos\theta}\right)$, where φ – rotation angle around z axis, θ – around y axis, ψ – around x axis

$y = -r\cos\psi\sin\varphi$, where y – displacement on y axis, r – radius of moveable platform

$x = \frac{r}{2}(\cos\theta\cos\varphi + \sin\psi\sin\theta\sin\varphi - \cos\psi\cos\varphi)$, where x – displacement on x axis.

The rotation matrix for the transition from the coordinate system located in the center of the fixed platform to the coordinate system located in the center of the moving platform:

$$R_p = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi \\ 0 & \sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$$R_p = \begin{bmatrix} \cos\theta\cos\varphi + \sin\psi\sin\theta\sin\varphi & -\cos\theta\sin\varphi + \sin\psi\sin\theta\cos\varphi & \sin\theta\cos\psi \\ \cos\psi\sin\varphi & \cos\psi\cos\varphi & -\sin\psi \\ -\sin\theta\cos\varphi + \sin\psi\cos\theta\sin\varphi & \sin\theta\sin\varphi + \sin\psi\cos\theta\cos\varphi & \cos\theta\cos\psi \end{bmatrix} \quad (3)$$

$$= \begin{bmatrix} Rp11 & Rp12 & Rp13 \\ Rp21 & Rp22 & Rp23 \\ Rp31 & Rp32 & Rp33 \end{bmatrix}$$

The main criterion for determining the occurrence of a point in the working area is the occurrence of each of the link lengths in the permissible range. The generalized formula for link lengths:

$$L_i = \sqrt{(x_{Ai} - x_{Bi})^2 + (y_{Ai} - y_{Bi})^2 + (z_{Ai} - z_{Bi})^2} \quad (4)$$

The links lengths can be determined as follows:

$$L_1 = \sqrt{(x + r \cdot Rp11 - R)^2 + (z + r \cdot Rp31)^2}, \quad (5)$$

$$L_2 = \left((x + 0,5r(Rp11 + \sqrt{3}Rp12) - 0,5R)^2 + (y - 0,5r(Rp21 + \sqrt{3}Rp22) + \frac{\sqrt{3}}{2}R)^2 + (z - 0,5r(Rp31 + \sqrt{3}Rp32))^2 \right)^{0,5}, \quad (6)$$

$$L_3 = \left((x - 0,5r(Rp11 - \sqrt{3}Rp12) + 0,5R)^2 + (y - 0,5r(Rp21 - \sqrt{3}Rp22) - \frac{\sqrt{3}}{2}R)^2 + (z - 0,5r(Rp31 - \sqrt{3}Rp32))^2 \right)^{0,5}, \quad (7)$$

where z – moving along the z axis.

Given a list of parallelepipeds \mathbb{P}_1 , which initially includes one parallelepiped Q_1 in the moving coordinate system, which is guaranteed to include the work area. During the execution of the algorithm (Fig. 9), the link length functions are calculated. If at least one of the calculated lengths is out of the allowable range at all points of the parallelepiped, then this parallelepiped is not included in the working area and added in the list of external approximations. If in all points of a parallelepiped all the functions of the link lengths are in the permissible range or the diameter of the parallelepiped $d(Q_i)$ is smaller than the accuracy of the approximation δ , then this parallelepiped enters the working area and added in the list of inner approximation. In other cases, the parallelepiped is divided in half along the maximal length of the edge and 2 parallelepipeds are added to the end of the list \mathbb{P}_1 for further consideration. The following geometric parameters were chosen for simulation:

$$L_{i,min} = 70 \text{ mm}$$

$$L_{i,max} = 130 \text{ mm}$$

$$R = r = 50 \text{ mm}$$

Results of simulation is presented on Fig. 10, 11.

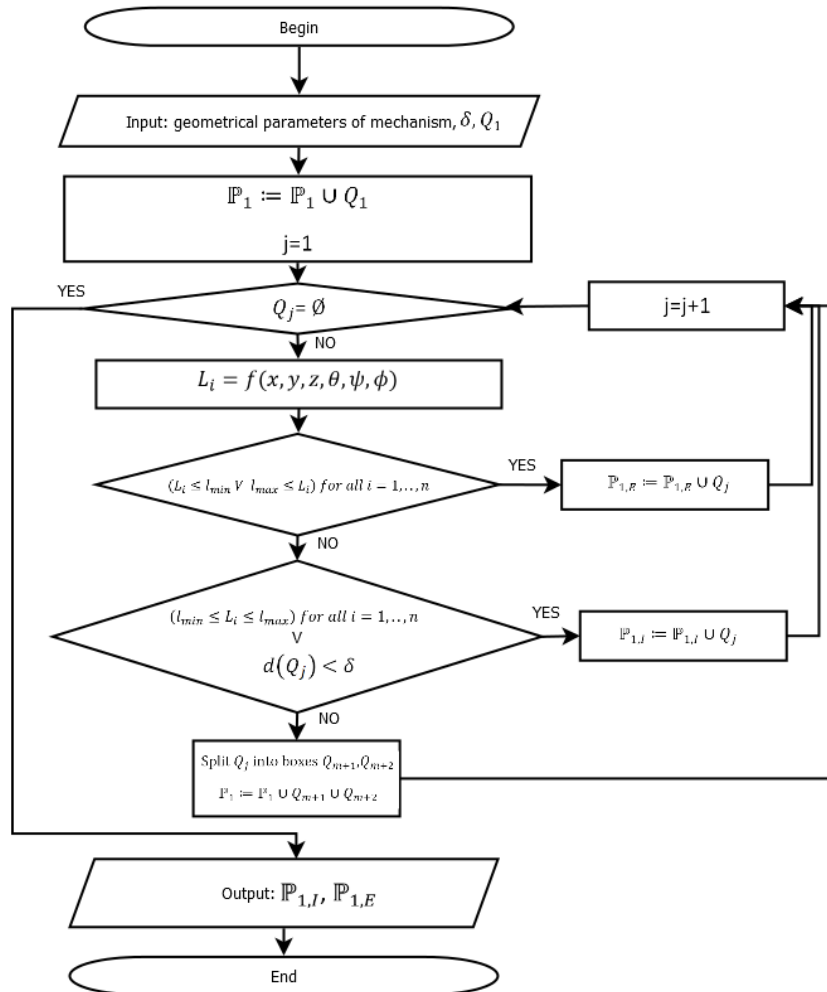


Figure 9. Algorithm for tripod

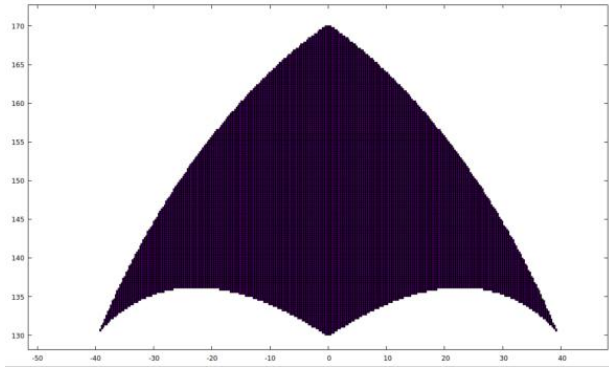


Figure 10. The working area in projection on yz plane

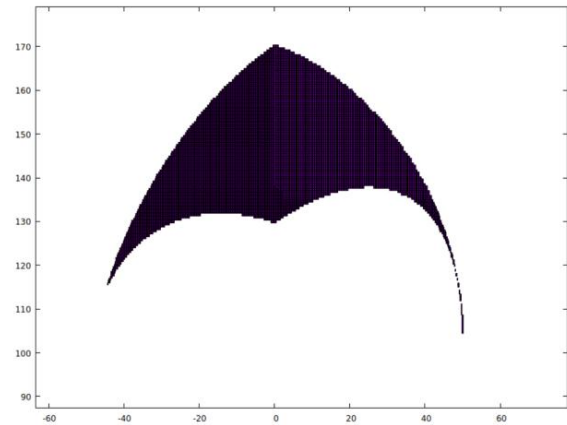


Figure 11. The working area in projection on xz plane

Definition the working area of the lower mechanism - a planar robot (Fig. 12) was developed within the framework of the Russian Science Foundation project, the agreement number 17-79-10512. This mechanism consists of three chains containing three rotational kinematic pairs O_i , A_i , B_i ($i=1,2,3$). The axes of rotation of all pairs are parallel to each other and perpendicular to the plane in which the mechanism moves. Rotary pairs A_i are fixed on a base, and their position is given by the coordinates x_i , y_i in a fixed rectangular coordinate system. The position of the output link of the mechanism is given by the position of the point D and is described by the coordinates x and y as well as the angle of rotation φ of this link with respect to some initial position. The displacement of the output link is carried out due to the rotation of the driving (input) pairs A_i . The angles of rotation θ_i of these pairs are generalized coordinates for the given mechanism. R and r – the radii of the circles described the triangles $O_1O_2O_3$ and $B_1B_2B_3$ respectively.

The geometry of the output link, i.e. the mutual arrangement of points C_1 , C_2 , C_3 and D is given by angles γ_i as well as distances C_iD .

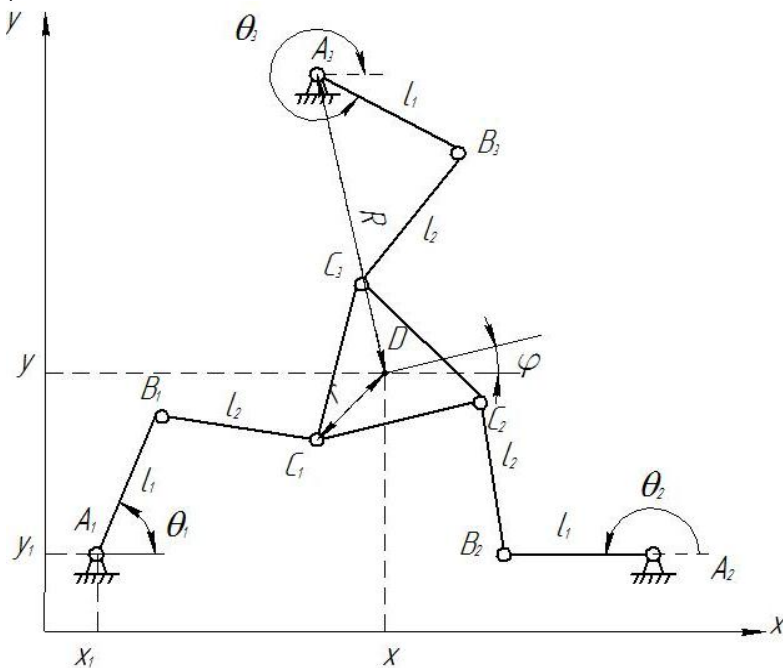


Figure 12. Scheme of a planar 3-RRR mechanism

The criterion of a planar mechanism is the permissible range of rotation angles of drive pairs. The formulas to calculate the angles θ_i are:

$$\theta_i = \arcsin \left(\frac{l_{2,i}^2 - l_{1,i}^2 - [b_{i,\cos}] - [b_{i,\sin}]}{\sqrt{[a_{i,\sin}]^2 + [a_{i,\cos}]^2}} \right) - \phi, \quad (8)$$

$$[a_{i,\sin}] = 2[y_i l_{1,i} - y l_{1,i} - l_{1,i} l_{3,i} \sin(\gamma_i + \varphi)], \quad (9)$$

$$[b_{i,\cos}] = \begin{bmatrix} x^2 + 2x l_{3,i} \cos(\gamma_i + \varphi) - 2x x_i + l_{3,i}^2 \cos^2(\gamma_i + \varphi) - \\ -2x_i l_{3,i} \cos(\gamma_i + \varphi) + x_i^2 \end{bmatrix}, \quad (10)$$

$$[a_{i,\cos}] = 2[x_i l_{1,i} - x l_{1,i} - l_{1,i} l_{3,i} \cos(\gamma_i + \varphi)], \quad (11)$$

$$[b_{i,\sin}] = \begin{bmatrix} y^2 + 2y l_{3,i} \sin(\gamma_i + \varphi) - 2y y_i + l_{3,i}^2 \sin^2(\gamma_i + \varphi) - \\ -2y_i l_{3,i} \sin(\gamma_i + \varphi) + y_i^2 \end{bmatrix}. \quad (12)$$

Planar mechanism algorithm (Fig. 13) is developed. It is similar to the tripod algorithm. The following geometric parameters were chosen for simulation:

$$\begin{aligned} l_{1i} = l_{2i} = l_{3i} = r &= 50 \text{ mm}, \\ R &= 100 \text{ mm}, \\ \theta_{O,min} &= -45^\circ, \\ \theta_{O,max} &= 45^\circ. \end{aligned}$$

Results of simulation is presented on Fig. 14.

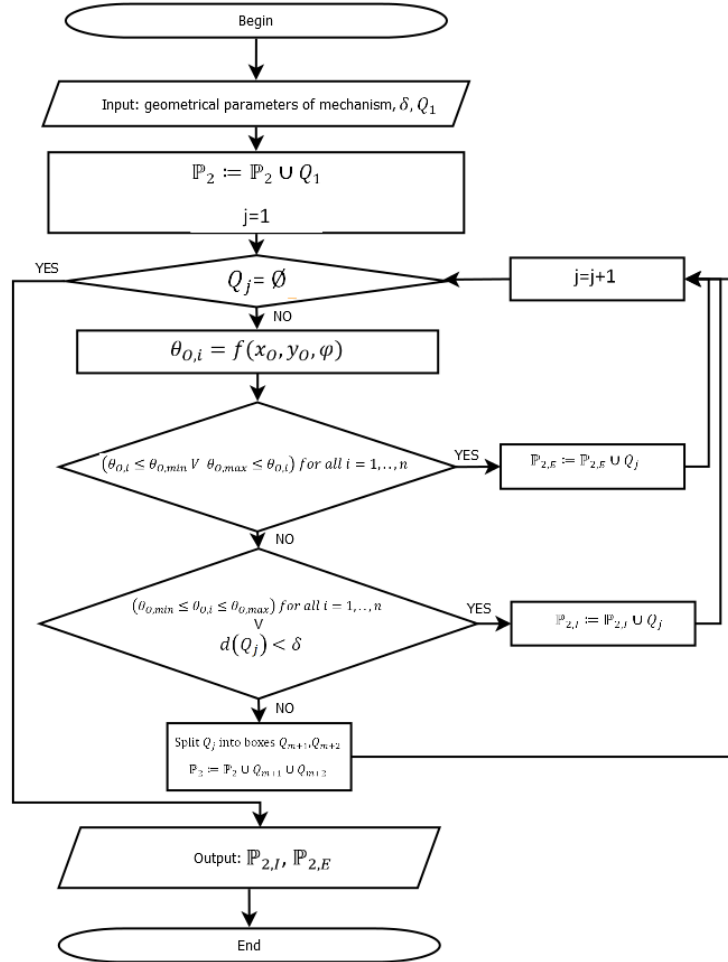


Figure 13. Algorithm for planar 3-RRR.

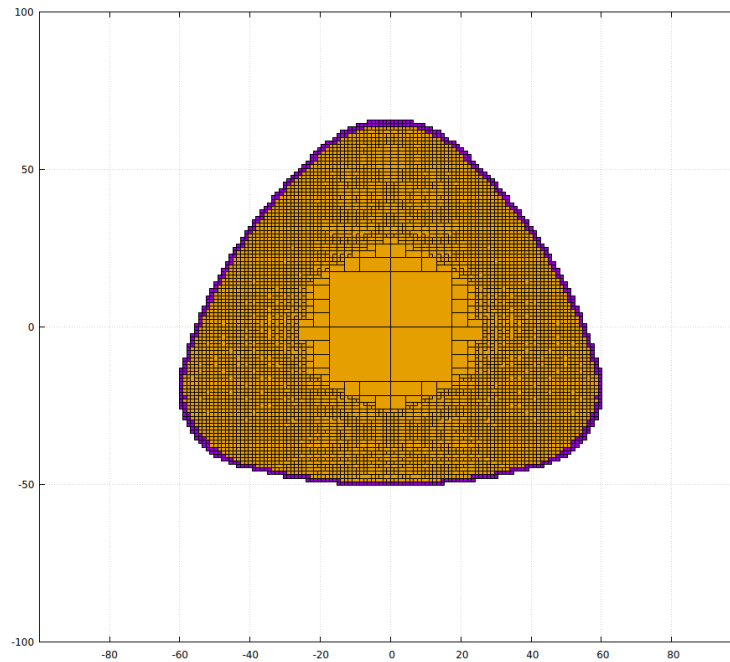


Figure 14. Results of simulation

Consider the scheme of a complicated mechanism which include the tripod and flat 3-RRR mechanism.

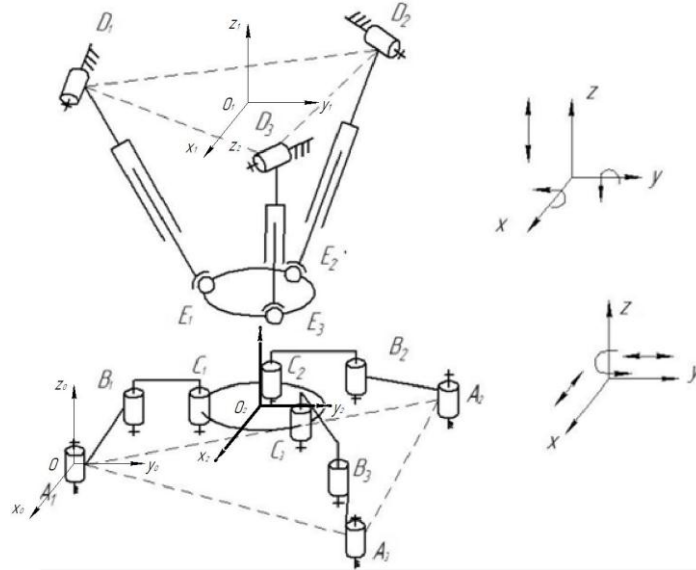


Figure 15. Kinematic scheme of the robot with relative handling modules

Development of an algorithm for approximating a joint working area

Algorithm for approximating a joint working area in a moving coordinate system is developed within the framework of the Russian Science Foundation project, the agreement number 17-79-10512. The algorithm is described below:

1. 2 fixed coordinate systems: $x_0y_0z_0$ is located in the center of the joint A_1 , $x_1y_1z_1$ in the center of the upper fixed platform O_1 .
2. The moving coordinate system $x_2y_2z_2$ of the planar mechanism is placed in the center of the planar mechanism platform O_2 .
3. Let's use the working area parallelepiped lists of the upper mechanism $\mathbb{P}_{1,I}$ and the lower mechanism $\mathbb{P}_{2,I}$.
4. Given a list of parallelepipeds \mathbb{P} , which initially includes one parallelepiped Q_1 in the moving coordinate system, which is guaranteed to include the work area.
5. Apply a grid A_1 to this parallelepiped.
6. Apply the grid A_2 to the parallelepipeds from list $\mathbb{P}_{2,I}$ of the flat mechanism.
7. For each of the grid A_1 nodes of the parallelepiped Q_i , calculate the rotation and displacement matrices of the transition from the coordinate system $x_2y_2z_2$ to the coordinate system $x_1y_1z_1$. These matrices have variable values equal to the coordinates of the grid A_2 nodes in the coordinate system $x_0y_0z_0$.
9. If all nodes of the grid A_1 after the transition do not belong to the working area $\mathbb{P}_{1,I}$, then the parallelepiped Q_i is added to the \mathbb{P}_E list of parallelepipeds of the external approximation.
10. If all nodes of the grid A_1 after the transition belong to the working area $\mathbb{P}_{1,I}$ or the diameter of the parallelepiped $d(Q_i)$ is smaller than the accuracy of the approximation δ , then the parallelepiped Q_i is added to the list \mathbb{P}_I of parallelepipeds of internal approximation.
11. In other cases, the parallelepiped is divided in half along a larger edge and added to the end of the list \mathbb{P} .
12. Repeating items 5-11 is performed before check the last parallelepiped in list \mathbb{P} .

Results of simulation with projection on different plane is presented on Fig. 17.

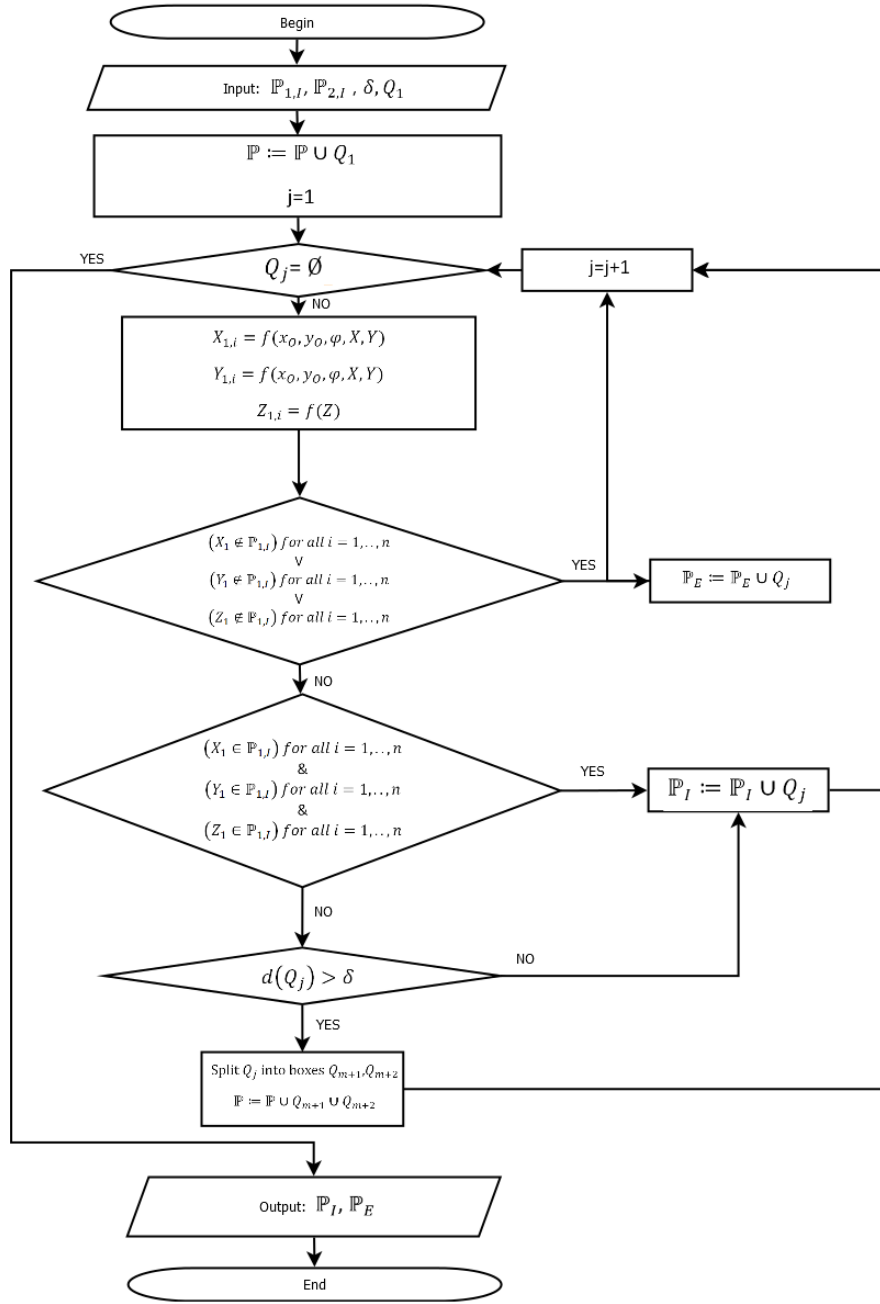


Figure 16. Algorithm for joint working area

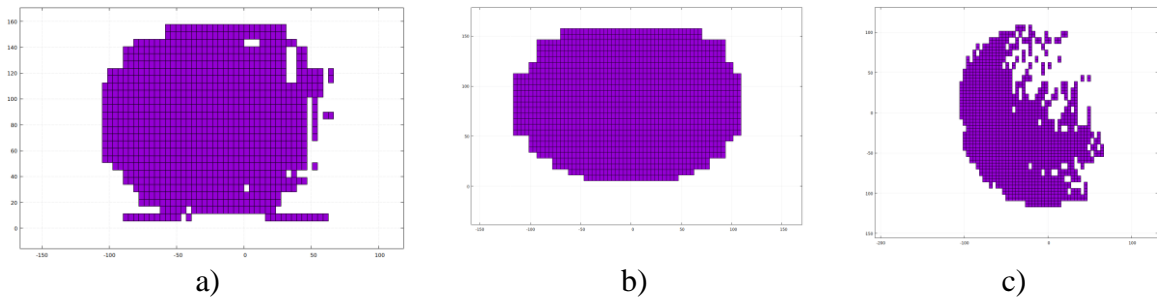


Figure 17. The working area in projection on a) xz plane, b) yz plane, c) xy plane

Conclusions

The applied algorithm showed its efficiency. The computation time for accuracy of approximation $\delta = 6\text{mm}$ and the dimension of the grid for enumerating the calculation of functions $64 \times 64 \times 64$ on a personal computer was 2 hours and 45 minutes. The accuracy of the approximation ($\delta = 6\text{ mm}$) of given complicated robot with 6 degrees of freedom in comparison with planar mechanisms with three degrees of freedom ($\delta = 0.006\text{-}0.06\text{ mm}$) is reduced in 100-1000 times, which is due to increase in the dimensionality of the problem and significant computing resources are required to increase accuracy. The resulting work area has fuzzy boundaries, which may be due to insufficient computational power and the presence of singularity zones.

Conclusions

This work was supported by the Russian Science Foundation, the agreement number 16-19-00148, the agreement number 17-79-10512

References

- [1] Evtushenko Y. G., Posypkin M. A. An application of the nonuniform covering method to global optimization of mixed integer nonlinear problems // Computational Mathematics and Mathematical Physics. 2011. T. 51. № 8. C. 1286-1298.
- [2] Yuri G. Evtushenko, Mikhail A. Posypkin, Larisa A. Rybak, Andrei V. Turkin. Numerical method for approximating the solution set of a system of non-linear inequalities // International Journal of Open Information Technologies ISSN: 2307-8162 vol. 4, no. 12, 2016
- [3] Dmitry I. Malyshev, Mikhail A. Posypkin, Larisa A. Rybak, Alexander L. Usov. Analysis of the working area of the robot DexTAR - dexterous twin-arm robot // International Journal of Open Information Technologies vol. 6, no.7, 2018. pp. 15-20.
- [4] Pundru Srinivasa Rao and Nalluri Mohan Rao. Position Analysis of Spatial 3-RPS Parallel Manipulator // International Journal of Mechanical Engineering and Robotics Research. 2013. Vol. 2. No. 2. P. 80-90.

Numerical Investigation of Vortex-induced Vibration (VIV) of a Flexible Cylinder in Combined Flow

Di Deng¹, Zhe Wang¹, Decheng Wan¹

¹State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

Corresponding author: dcwan@sjtu.edu.cn

Abstract

In actual oil exploration process, vortex-induced vibration (VIV) is the main source of structural fatigue damage of the risers. In this paper, VIV of a flexible cylinder experiencing combined uniform and oscillatory flow is investigated numerically. All investigations are carried out by the in-house CFD code viv-FOAM-SJTU, which is developed basing on the pimplyDyMFOAM solver attached to the open source OpenFOAM. The effects of flow ratio on VIV are concentrated, while the flow ratio α is defined as the proportion of the uniform flow velocity in the total velocity. Main parameters of the cylinder are as follows: the mass ratio $m^* = 1.53$, the diameter $D = 0.024$, the length $L = 4$ and the Keulegan–Carpenter (KC) number $KC = 178$. The flow ratio varies from 0 to 1 with an interval of 0.2. The modal analysis method and the wavelet analysis method are used to study the effect of flow ratio to VIV response of the cylinder in combined flow.

Keywords: vortex-induced vibration (VIV); viv-FOAM-SJTU solver; strip method; flow ratio

Introduction

Vortex-induced vibration (VIV) of a flexible cylinder in steady flow has been investigated extensively during the past decades through experimental and numerical methods, such as Chaplin et al[3][4], Lie and Kaasen[9], Willden and Graham [13][14] and Yamamoto et al[16]. Chaplin et al[3] carried out benchmark experiments of VIV of a long flexible vertical tension cylinder in stepped flow. The modal analysis method is used to obtain modal weights of each vibration mode and determine the dominant vibration mode of the cylinder. Lie and Kaasen[9] also used the modal analysis method to analyze the vibration feature of a flexible cylinder in sheared flow. And they chose to solve modal amplitudes through the least-square sense with the existence of some modal amplitudes that are not physical with regard to the frequency content. Willden and Graham[13] and Yamamoto et al[16] carried out numerical studies of VIV of a flexible cylinder in uniform flow adopting strip method. Numerical results were in good agreements with experiments and the strip method was appropriate for solving VIV problems.

In order to study the vibration features of a circular cylinder around oscillatory flow, comprehensive researches have been carried out by Bearman[1][2], Kozakiewicz et al[8], Sarpkaya[11][12], Williamson[15], and Zhao et al[17]-[19]. Williamson[15] and Sarpkaya[11][12] conducted a series of experiments to investigate motions of vortices around a single cylinder in relative oscillatory flow. And several vortex regimes were identified within particular ranges of Keulegan-Carpenter (KC) Numbers: the attached vortices regime

($0 < KC < 7$), where no major vortices shed during a cycle; the single pair regime ($7 < KC < 15$); the double pairs regime ($15 < KC < 24$); the three pairs regime ($24 < KC < 32$) and ect. For further KC regimes, the number of vortices pairs shed in each oscillating period would be increased by one each time the KC regime changed to a higher one. Kozakiewicz et al[8] and Zhao et al[17] carried out experiments and numerical simulations of a cylinder exposed to oscillatory flow for two KC numbers of 10 and 20 respectively. Kozakiewicz et al[8] found that the cross-flow vibration of the cylinder changed the vortex shedding trail and the number of vortices generated over one oscillating cycle comparing with the fixed cylinder. Zhao et al[17] found that the reduced velocity had significant effects to the XY- trajectory mode of the cylinder and the VIV frequency decreased with the increase of reduced velocity. And when the reduced velocity was extremely large, the vibration amplitude in the cross-flow direction was negligible smaller than that of the inline direction. Basing on the previous simulations, Zhao et al[19] carried out simulations of a circular cylinder experiencing combined oscillatory flow and steady flow at $KC=10$. They found that the lock-in regime was widened due to the combination of oscillatory and steady flow and the widest lock-in regime were twice as wide as that in the pure oscillatory or pure steady flow. For flexible cylinder condition, Fu[7] carried out a series of experiments of a flexible cylinder in relative oscillatory flow at $KC=178$. They proposed the VIV development process of “Build up—Lock in—Die out” in each half oscillating period. And Moreau and Huang[10] conducted experiments of cross-flow vortex-induced vibration in combined in-line current and oscillatory flow, including 12 different combinations of flow and cylinder conditions in total. He found that the VIV amplitude response was much reduced in the combined flow comparing with the pure steady flow at a given reduced velocity.

In this paper, VIV of a flexible cylinder experiencing combined oscillatory and uniform flow is investigated. All numerical simulations are carried out by the in-house CFD code viv-FOAM-SJTU, which is developed basing on the strip theory method and the pimpleDyMFOAM solver attached to the open source code OpenFOAM. The effect of flow ratio is concerned and the numerical model refers to the experiments of Fu et al[7]. The flow ratio α is defined as the proportion of the uniform flow velocity in the total velocity. The flow ratio varies from 0 to 1 with an interval of 0.2. Firstly, the validation is conducted at the pure oscillatory flow condition comparing results of cross-flow vibration history and dominant vibration frequency with Fu et al[7]. Then modal analysis and wavelet analysis methods are used to study the effect of flow ratio to VIV response.

This paper is organized as follows: the first section gives a brief introduction to the referenced experiments and the numerical methodology. The second section presents the results and the final section concludes the paper.

Method

Hydrodynamics Governing Equations

The flow field is supposed to be incompressible, with constant dynamic viscosity μ and constant density ρ . The Reynolds-averaged Navier-Stokes equations are used as the hydrodynamics governing equations as follow:

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0 \quad (1)$$

$$\frac{\partial}{\partial t}(\rho \bar{u}_i) + \frac{\partial}{\partial x_j}(\rho \bar{u}_i \bar{u}_j) = -\frac{\partial \bar{p}}{\partial x_i} + \frac{\partial}{\partial x_j}(2\mu \bar{S}_{ij} - \rho \overline{u_j u_i}) \quad (2)$$

where $\bar{S}_{ij} = \frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right)$ is the mean rate of strain tensor, $-\rho \overline{u_j u_i}$ is referred as Reynolds stress τ_{ij} computed by $\tau_{ij} = -\rho \overline{u_j' u_i'} = 2\mu_t \bar{S}_{ij} - \frac{2}{3} \rho k \delta_{ij}$, where μ_t is the turbulent viscosity and $k = (1/2) \overline{u_i' u_i'}$ is the turbulent energy, computing from the fluctuating velocity field.

Structural Dynamic Governing Equations

In order to form the relatively oscillatory flow, the supporting frame is forced to oscillate harmonically during the investigations. The oscillation can be expressed as:

$$x_s = A_m \cdot \sin\left(\frac{2\pi}{T}t\right) \quad (3)$$

$$U_s = \frac{2\pi \cdot A_m}{T} \cdot \cos\left(\frac{2\pi}{T}t\right) = U_m \cdot \cos\left(\frac{2\pi}{T}t\right) \quad (4)$$

$$KC = \frac{2\pi \cdot A_m}{D} = \frac{U_m \cdot T}{D} \quad (5)$$

where A is the oscillating amplitude, T is the oscillating period, x_s is the oscillating displacement, U_s is the oscillating velocity, U_m is the amplitude of the oscillating velocity, D is the diameter of the cylinder.

Fu[6] uses the support excitation method combined with the Bernoulli–Euler bending beam theory to obtain the structural response of the cylinder. The in-line displacement of the cylinder is the sum of support frame motion and the relative in-line vibration of the cylinder:

$$x_t = x_s + x \quad (6)$$

where x_t is the in-line displacement, x_s is the support displacement and x is the relative in-line displacement.

The equilibrium of forces for this system can be written as follow:

$$f_I + f_D + f_S = f_H \quad (7)$$

where f_I , f_D , f_S , f_H are the inertial, the damping, the spring, and the hydrodynamic force respectively.

Then the equilibrium of forces for the system can be written as:

$$m\ddot{x}_t + c\dot{x} + kx = f_H \quad (8)$$

$$m\ddot{x} + c\dot{x} + kx = f_H - m\ddot{x}_s \quad (9)$$

where m , c , k are the mass, the damping and the stiffness of the system.

Adopting the finite element method(FEM), the equations can be discretized as:

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{F}_{Hx} - \mathbf{M}\ddot{\mathbf{x}}_s \quad (10)$$

$$\mathbf{M}\ddot{\mathbf{y}} + \mathbf{C}\dot{\mathbf{y}} + \mathbf{K}\mathbf{y} = \mathbf{F}_{Hy} \quad (11)$$

where \mathbf{M} , \mathbf{C} , \mathbf{K} are the mass, the damping and the stiffness matrices, while \mathbf{x} , \mathbf{x}_s and \mathbf{y} are the relative in-line, the support and the cross-flow nodal displacement vectors. \mathbf{F}_{Hx} and \mathbf{F}_{Hy} are the hydrodynamic force in the in-line and cross-flow direction respectively.

Problem Description

The numerical model used in this paper follows experiments of Fu[7] and the layout of the experiments is shown in Fig. 1. Detailed information about main parameters of the cylinder is shown in Table 1. 20 strips located equidistantly along the cylinder totally. Fig. 2 shows the distribution of flow field strips along the span of the cylinder and the entire computational domain and meshes of strips.

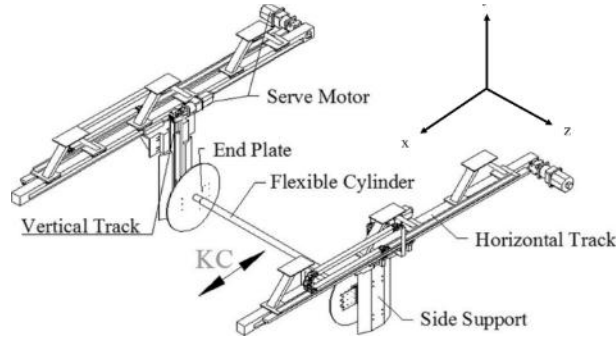


Figure 1. Layout of the experiments of Fu et al

Table 1: Main parameters of the cylinder

	Symbols	Values	Units
Mass ratio	m^*	1.53	—
Diameter	D	0.024	m
Length	L	4	m
Bending stiffness	EI	10.5	$N \cdot m^2$
Top tension	T_t	500	N
First natural frequency	f_n^1	2.68	Hz
Second natural frequency	f_n^2	5.46	Hz

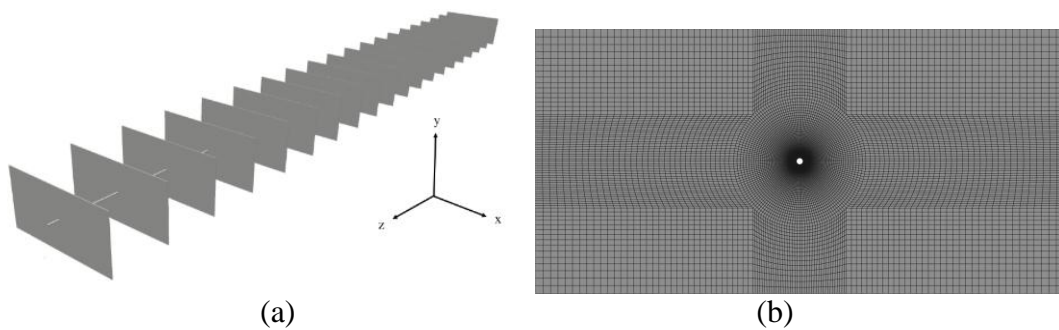


Figure 2. Illustration of multi-strip model and computational domain of a strip

In this paper, VIV of a cylinder in combined uniform and oscillatory flow are investigated. The flow ratio α represents the proportion of uniform flow velocity in the total flow velocity. According to equations (3) and (4), the total velocity and the flow ratio can be written as equations (12) and (13). Detailed computational conditions are shown in Table 2.

$$U_c(t) = U_s + U_m \cos\left(\frac{2\pi}{T}t\right) = U_s + A_m \frac{2\pi}{T} \cos\left(\frac{2\pi}{T}t\right) \quad (12)$$

$$\alpha = \frac{U_s}{U_s + U_m} = 1 - \frac{A_m}{U_c} \cdot \frac{2\pi}{T} \quad (13)$$

where U_s is the uniform flow velocity, U_c is the total velocity, A_m is the amplitude of the oscillation, T is the oscillating period.

Table 2 Computational conditions

	U_c	α	U_s	A_m	T	KC
Case1	0.2589	0	0	0.68	16.5	178
Case2	0.2589	0.2	0.05178	0.68	20.6	178
Case3	0.2589	0.4	0.10356	0.68	27.5	178
Case4	0.2589	0.6	0.15534	0.68	41.3	178
Case5	0.2589	0.8	0.20712	0.68	82.5	178
Case6	0.2589	1	0.2589	0.68	-	178

Strip Theory

In this paper, numerical investigations are carried out by the viv-FOAM-SJTU solver basing on the strip method and the pimpDyMFOAM solver attached to the open source code OpenFOAM. The strip method is very appropriate for solving CFD investigations of supramaximal computational domain. It owns high computational efficiency and the computational accuracy is reliable, The reliability of the viv-FOAM-SJTU solver has been testified by Duan[2] in which the benchmark case has been verified in detail.

For a long flexible cylinder, the direct computation of the three dimensional flow field will cost too much resources. Instead of this, we simplify CFD model and obtain the two dimensional flow field on strips distributed equably along the cylinder. The hydrodynamic force is obtained from each strip, which is then applied to the structural field. The structural displacements of all nodes are interpolated to get the boundary motion of dynamic mesh of flow field. The strip theory is shown as Fig. 1.

During the numerical investigations, the RANS equations and SST k- ω turbulence model are adopted to solve the flow field in each strip, while the whole structure filed is solved through Bernoulli–Euler bending beam theory with the finite element method. The fluid-structure interaction is carried out by loose coupling strategy.

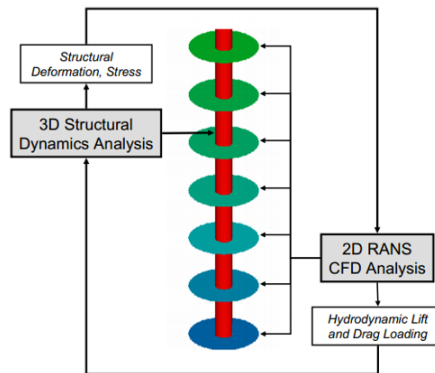


Figure. 3 Schematic diagram of strip theory

Results

Validation

Fig. 4 shows subplots of non-dimensional cross-flow amplitude of the intermediate node of the cylinder between experiment and simulation. From comparison, it can be concluded: (i) the development process of “Building-up—Lock-in—Dying-out” of vortex-induced vibration is observed in both experiment and numerical simulation; (ii) the lock-in region is 17.3% of the half oscillating period in numerical simulation, which is close to the experiment result of 17%; (iii) the non-dimensional cross-flow amplitude is 0.37D in half oscillating period, which is close to the experiment result of 0.36D.

Fig. 5 are subplots of power spectral density and modal weight of each vibration mode in an oscillating period respectively. From these figures, it can be known that the dominant vibration frequency is 2.2Hz, which is close to the result of experiments of 2.1Hz. While the dominant vibration mode of the cylinder is the 1st mode.

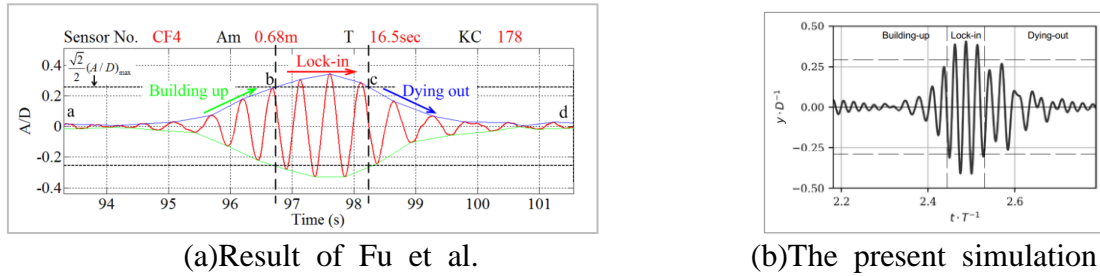


Figure 4. Non-dimensional cross-flow vibration amplitude of the intermediate node in half an oscillating period

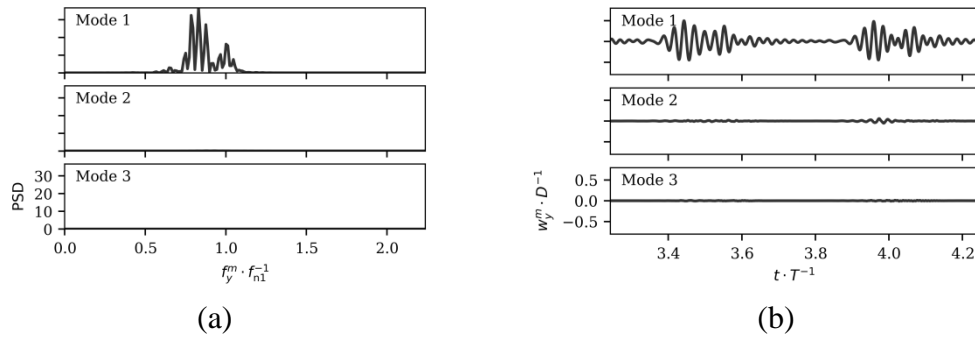


Figure 5. Cross-flow power spectral density and modal weight of each vibration mode of the intermediate node: (a) power spectral density; (b) modal weight

Modal Analysis

Fig. 6 are subplots of non-dimensional cross-flow vibration amplitude of the intermediate node of the cylinder ranging from $\alpha=0$ to $\alpha=1.0$ in an oscillating period. Two VIV development process of “Building-up—Lock-in—Dying-out” can be observed both in Fig. 6(a) and 6(b), which shows that the oscillatory flow plays the dominant role in the VIV phenomenon of the cylinder. There are two obvious lock-in region in Fig. 7(b) at $\alpha=0.2$ in an oscillating period. When flow velocities are in the same direction, the vibration amplitude is 0.27D and the lock-in region is 31.4% of the half oscillating period. When flow velocities are in the opposite direction, the vibration amplitude is 0.05D and the lock-in region is 13.6% of

the half oscillating period. As shown in Fig. 6(c) and 6(d), the obvious VIV phenomenon is observed in the half oscillating period where the oscillatory flow velocity and the uniform flow velocity are in the same direction. While no obvious VIV phenomenon happens when two flow velocities are in opposite direction. Both oscillatory flow and uniform flow have non-negligible influence to the VIV of the cylinder. With the increase of flow ratio, the proportion of uniform flow velocity in the total flow velocity increases and the dominant effect of the uniform flow to VIV of the cylinder becomes obvious. From Fig. 6(e) and 6(f), it can be seen that the obvious VIV phenomenon is observed in the whole oscillating period. It can be concluded that the vibration feature of the cylinder in combined flow is similar to that in pure oscillatory flow ($\alpha=0$) when flow ration $\alpha \leq 0.2$ and similar to that in pure uniform flow ($\alpha=1.0$) when $\alpha \geq 0.8$.

Comparing Fig. 6(b) with 6(c), it can be known that the dominant effect of oscillatory flow becomes weak with the increase of flow ratio. When flow velocities are in the same direction at $\alpha=0.4$, the reduced velocity is large enough to generate VIV phenomenon in the whole half period. When oscillatory flow velocity reverses, the reduced velocity decreases and no VIV phenomenon generated in the whole half period. When flow velocity increases to $\alpha=0.6$ as shown in Fig. 7(d), VIV phenomenon occurs in the whole region when flow velocities are in the same direction as that of $\alpha=0.4$ and in the preliminary stage and final stage of the half period when flow velocities are in the opposite direction. In these two stages, the oscillatory flow velocity is small and the uniform flow still owns dominant effect to the vibration of the cylinder. However in the intermediate stage, the increase of oscillatory flow velocity leads to the decrease of total flow velocity, then VIV phenomenon becomes weaker and disappears finally. When flow velocity increases to $\alpha=0.6$ as shown in Fig. 7(e), it is found that the uniform flow plays dominant role in the cross-flow vibration of the cylinder. And the cross-flow vibration amplitude reaches its peak or valley value when the oscillatory velocity reaches its peak value.

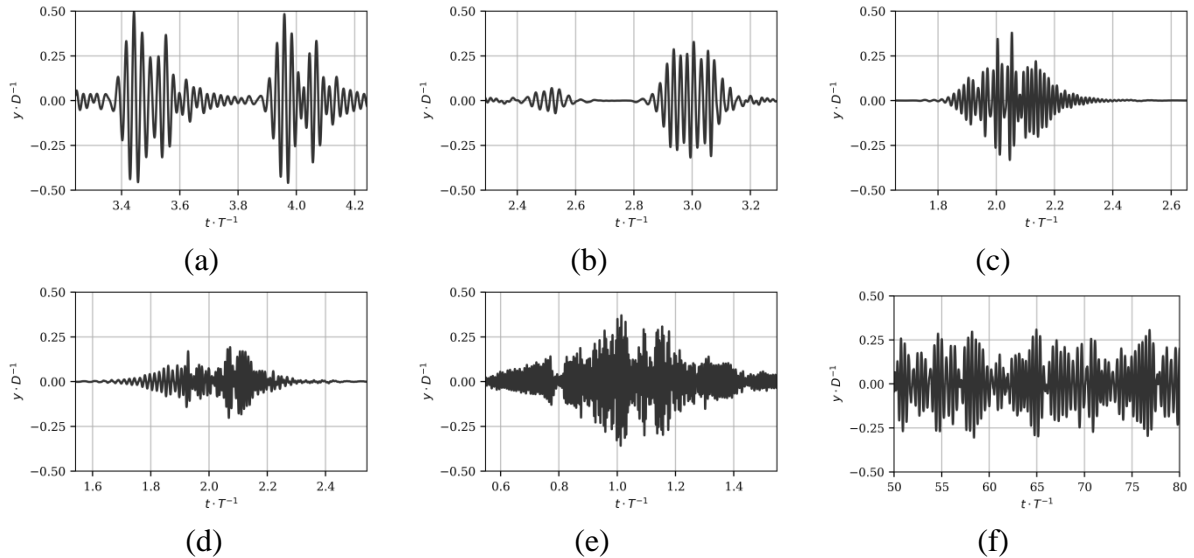


Figure 6. Non-dimensional cross-flow vibration amplitude of the intermediate node in an oscillating period: (a) $\alpha=0$; (b) $\alpha=0.2$; (c) $\alpha=0.4$; (d) $\alpha=0.6$; (e) $\alpha=0.8$; (f) $\alpha=1.0$

Fig. 7 are subplots of cross-flow modal weight of each vibration mode of the intermediate node of the cylinder ranging from $\alpha=0$ to $\alpha=1.0$ in an oscillating period. It can be found that the first mode is the dominant vibration mode when obvious VIV phenomenon occurs. And the disturbance of second mode is too small to change the dominant vibration mode of and

only can be observed in pure oscillatory flow and combined flow when flow velocities are in the same direction.

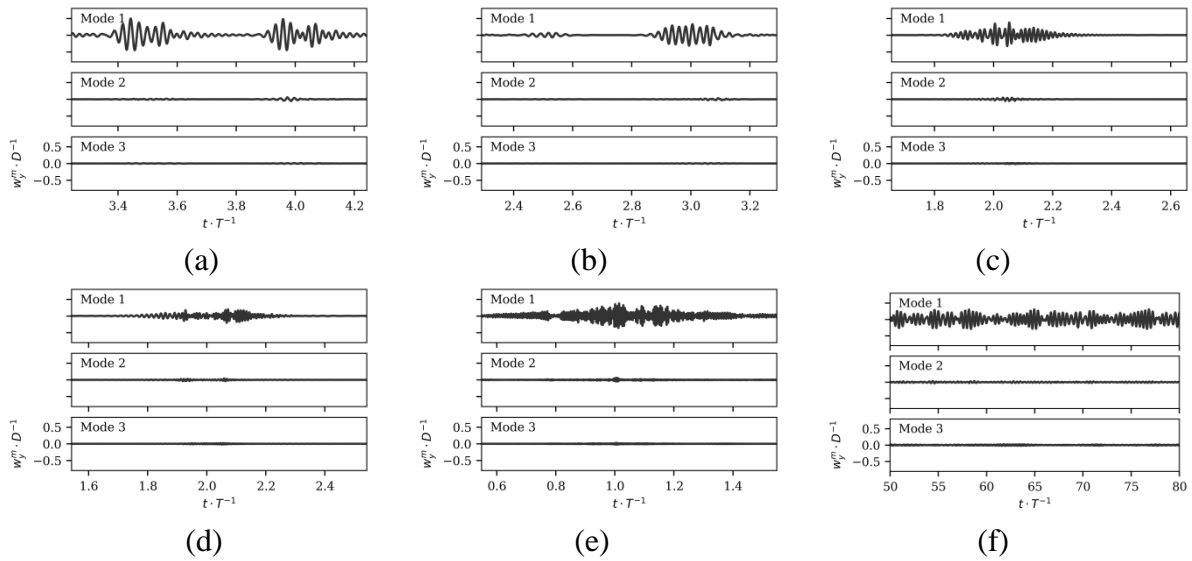


Figure 7. Cross-flow modal weight of each vibration mode of the intermediate node in an oscillating period: (a) $\alpha=0$; (b) $\alpha=0.2$; (c) $\alpha=0.4$; (d) $\alpha=0.6$; (e) $\alpha=0.8$; (f) $\alpha=1.0$

Wavelet Analysis

The wavelet analysis method is used to obtain the dominant cross-flow vibration frequency of the cylinder along the time. Fig. 8 shows cross-flow vibration wavelet of different nodes along the span of the cylinder. These subplots can be divided into three groups: (A) Fig. 8(a) and 8(b); (B) Fig. 8(c) and 8(d); (C) Fig. 8(e) and 8(f). The periodical variation of the dominant vibration frequency is obvious from Fig. 8(a) to 8(d).

In Fig. 8(a) and 8(b), the dominant vibration frequency is close to the first natural frequency of the cylinder in most of the oscillating period. In the flow ratio range of group A, the oscillatory flow has significant effect to the cross-flow vibration frequency. With the increase of flow ratio, the proportion of the cross-flow vibration frequency close to the first natural frequency decreases. As shown in Fig. 8(c) and 8(d), we can see that the increasing and decreasing process of the dominant vibration frequency is similar to the shape of sinusoidal function. During the half process of flow velocities in the same direction, the cross-flow vibration of the cylinder is drastic which leads to the generation of high vibration frequency region. During the half process of flow velocities in the opposite direction, the cross-flow vibration of the cylinder is relatively weak which leads to the generation of the low vibration frequency region. In the flow ratio range of group B, both oscillatory flow and uniform flow influence the cross-flow vibration frequency. From Fig. 8(e) and 8(f), it can be known that the variation of dominant vibration frequency is relatively small comparing with other cases. And the uniform flow plays the dominant role in this flow ratio range.

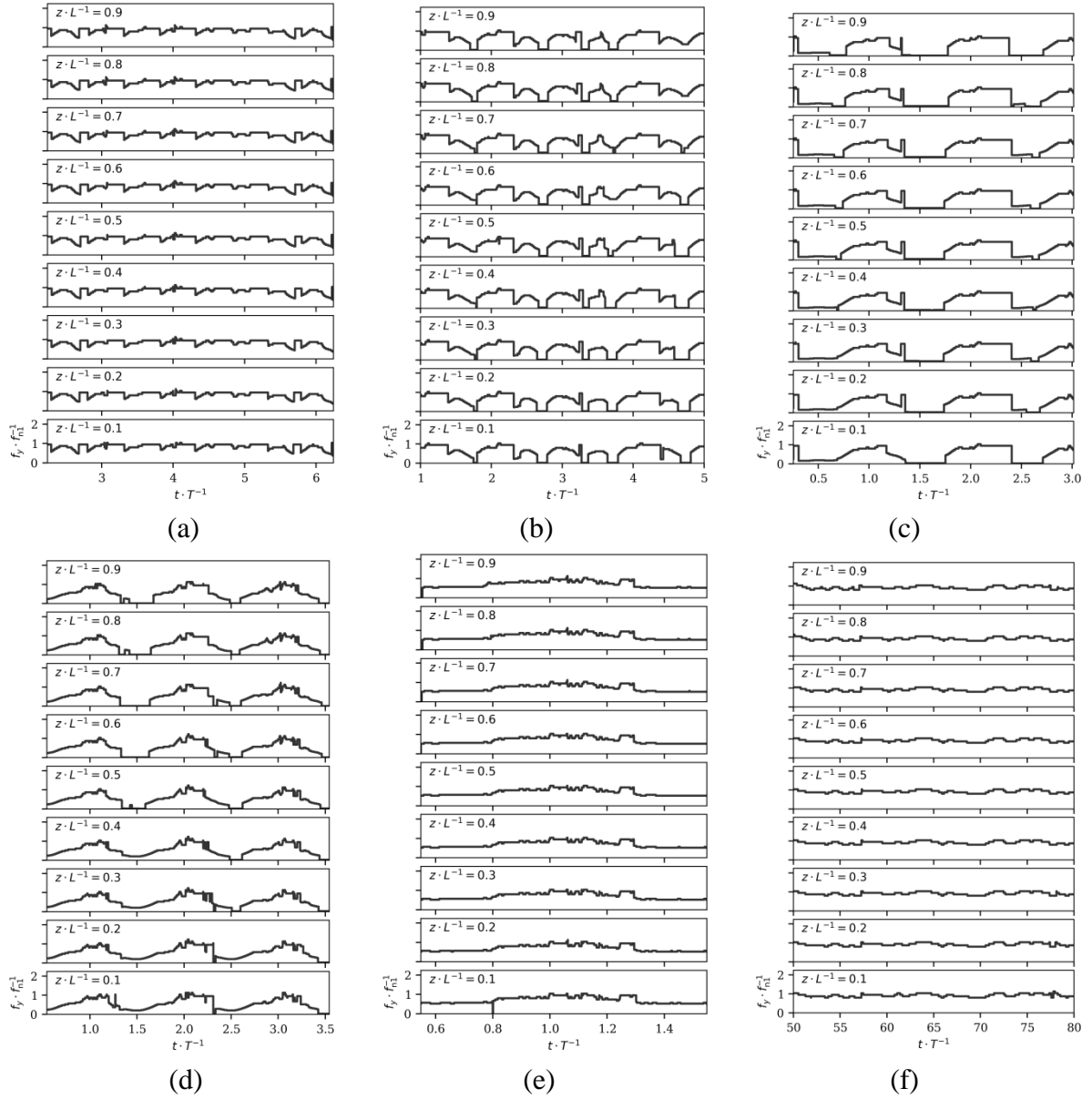


Figure 8. Cross-flow vibration wavelet of different nodes along the span of the cylinder: (a) $\alpha=0$; (b) $\alpha=0.2$; (c) $\alpha=0.4$; (d) $\alpha=0.6$; (e) $\alpha=0.8$; (f) $\alpha=1.0$

Cross-flow Vibration Trajectory

Fig. 9 are subplots of cross-flow vibration trajectory of three nodes along the span of the cylinder. With the increase of flow ratio, the trajectory shape of the intermediate node of the cylinder changes from the “H” type to the “1” type. For the pure oscillatory flow condition (Fig. 9(a)), the cross-flow vibration and the in-line deformation reaches its peak value when the cylinder moves across the center, which generate the two sides of the “H”. During the preliminary and the final stages, the oscillatory flow velocity reaches its valley value that leads to the small cross-flow vibration and in-line deformation of the cylinder, which generates the short transverse line of the “H” shape. When the flow ratio increases to 0.2, the flow velocity decreases during the process of oscillatory flow and uniform flow velocity in opposite direction. And the right side of the trajectory is generated due to the low cross-flow vibration amplitude and in-line deformation when the cylinder moves across the center during the process. With the flow ratio increasing, the total flow velocity in the process of opposite

flow velocities keeps decreasing. Then the cross-flow vibration amplitude and in-line deformation becomes smaller than those in flow ratio of 0.2. So the trajectory in the region of $x \geq 0$ is very small as shown in Fig. 9(c) and 9(d). Meanwhile, the proportion of the total flow velocity that generate relatively larger vibration amplitude increases in the process of the same flow velocities direction, which lead to the change of trajectory shape from triangle to rectangle in the region of $x \leq 0$. From Fig. 9(e), it can be concluded that the cross-flow vibration and in-line deformation of the cylinder are extremely small in the region of $x \geq 0$ and the trajectory shape is similar to that in Fig. 9(f), which means that the uniform flow plays the dominant role when flow ratio $\alpha=0.8$.

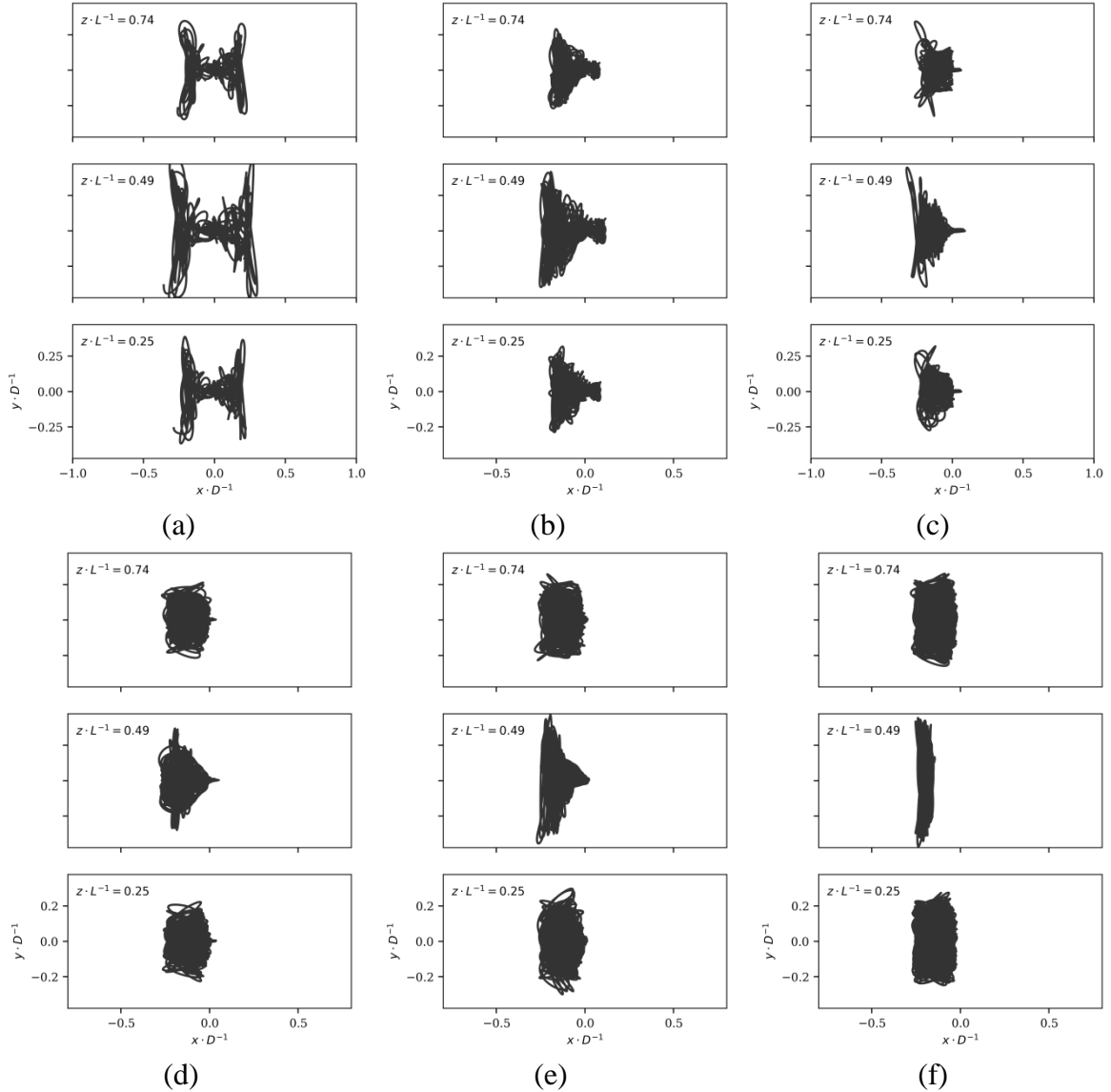


Figure 9. Cross-flow vibration trajectory of different nodes along the span of the cylinder: (a) $\alpha=0$; (b) $\alpha=0.2$; (c) $\alpha=0.4$; (d) $\alpha=0.6$; (e) $\alpha=0.8$; (f) $\alpha=1.0$

Conclusion

In this paper, numerical simulations of vortex-induced vibrations of a flexible cylinder experiencing combined oscillatory and uniform flow are carried out by the in-house CFD code viv-FOAM-SJTU solver. Results of cross-flow displacement history, modal weights,

wavelet and vibration trajectory are analyzed in detail.

Results of cross-flow displacement history and modal weights show that the first mode is the dominant vibration mode of the cylinder in all flow conditions. The dominant vibration frequency of the cylinder is approximately near the first natural frequency through results of cross-flow wavelet. With the flow ratio increasing, the vibration trajectory of the cylinder changes from the “H” type to the “1” type. The vibration responses of the cylinder in combined oscillatory and uniform flow are similar to that in pure oscillatory flow ($\alpha=0$) when flow ratio $\alpha \leq 0.2$ and similar to that in pure uniform flow ($\alpha=1.0$) when $\alpha \geq 0.8$.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd’s Register Foundation for doctoral student, to which the authors are most grateful.

Reference

- [1] Bearman P.W. Vortex shedding from oscillating bluff bodies. *Annual Review of Fluid Mechanics*, 1984,16, 195–222.
- [2] Bearman P.W., Downie M., Graham J. et al. Forces on cylinders in viscous oscillatory flow at low Keulegan-Carpenter numbers. *Journal of Fluid Mechanics*, 1985, 154, 337-356.
- [3] Chaplin J.R., Bearman P.W., Huera Huarte F.J., et al. Laboratory measurements of vortex-induced vibrations of a vertical tension riser in a stepped current. *Journal of Fluids Structures*, 2005, 21(1), 3-24.
- [4] Chaplin J.R., Bearman P.W., Cheng Y., et al. Blind Predictions of Laboratory Measurements of Vortex induced Vibrations of a Tension Riser. *Journal of Fluid Structures*, 2005, 21(1), 25-40.
- [5] Duanmu Y., Wan D.C. and Xue H.X. Prediction of Response for Vortex-Induced Vibrations of a Flexible Riser Pipe by using Multi-Strip Method. *Proceeding 26th International Offshore and Polar Engineering Conference, ISOPE, Rhodes, Greece, 26 June-2 July, 2016*, 65-1073.
- [6] Fu B.W., Zou L. and Wan D.C. Numerical study of vortex-induced vibrations of a flexible cylinder in an oscillatory flow. *Journal of Fluids Structures*, 2018, 77:170-181.
- [7] Fu S.X., Wang J.G., Baarholm R., et al. VIV of Flexible Cylinder in Oscillatory Flow. *International Conference on Offshore Mechanics and Arctic Engineering, OMAE, Nantes, France, June 9–14, 2013*, pp. V007T08A021.
- [8] Kozakiewicz A., Sumer B.M., Fredsøe J. et al. Vortex regimes around a freely vibrating cylinder in oscillatory flow. *Proceeding 6th International Offshore and Polar Engineering Conference, ISOPE, Los Angeles, US, 1996*, pp490-498.
- [9] Lie H. and Kaasen K.E. Modal analysis of measurements from a large-scale VIV model test of a riser in linearly sheared flow. *Journal of Fluids and Structures*, 2006, 22(4), 557-575.
- [10] Moreau F. and Huang S. Model Testing on Cross-Flow Vortex-Induced Vibration in Combined In-Line Current and Oscillatory Flows. *International Conference on Ocean, Offshore and Arctic Engineering, ASME 2010*, 2010:869-874.
- [11] Sarpkaya T. Force on a circular cylinder in viscous oscillatory flow at low Keulegan —Carpenter numbers. *Journal of Fluid Mechanics*, 1986, 165, 61-71.
- [12] Sarpkaya T. Hydrodynamic damping, flow-induced oscillations, and biharmonic response. *Journal of Offshore Mechanics and Arctic Engineering*, 1995, 117, 232–238.
- [13] Willden R.H.J. and Graham J.M.R. Multi-modal Vortex-Induced Vibrations of a vertical riser pipe subject to a uniform current profile. *European Journal of Mechanics*, 2004, 23(1):209-218.
- [14] Willden R.H.J. and Graham J.M.R. Numerical prediction of viv on long flexible circular cylinders. *Journal of Fluids and Structures*, 2001, 15(3-4):659-669.
- [15] Williamson C. Sinusoidal flow relative to circular cylinders. *Journal of Fluid Mechanics*, 1985, 155, 141-174.
- [16] Yamamoto C.T., Meneghini J.R., Saltara F. et al. Numerical simulations of vortex-induced vibration on flexible cylinders. *Journal of Fluids and Structures*, 2004, 19, 467-489.

- [17] Zhao M., Cheng. L. and Zhou. T. Numerical Investigation of Vortex-Induced Vibration (VIV) of a Circular Cylinder in Oscillatory Flow. ASME. International Conference on Offshore Mechanics and Arctic Engineering, OMAE, Rotterdam, The Netherlands, 2011, pp 597-603.
- [18] Zhao M. Numerical investigation of two-degree-of-freedom vortex-induced vibration of a circular cylinder in oscillatory flow. Journal of Fluids and Structures, 2013, 39(5):41-59.
- [19] Zhao M., Kaja K., Xiang Y. et al. Vortex-induced vibration (VIV) of a circular cylinder in combined steady and oscillatory flow. Ocean Engineering, 2013, 73:83-95.

Numerical Analysis on Two Floating Offshore Wind Turbines with Different Layouts

***Yang Huang, Ping Cheng and †Decheng Wan**

Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, China

*Presenting author: 958185119@qq.com

†Corresponding author: dcwan@sjtu.edu.cn

Abstract

The layout of wind turbines has significant effects on the power generating capacity and the economic efficiency of the wind farm. How to rationally arrange the wind turbines, minimize the mutual interference in the wake, and improve the efficiency of the whole wind farm have been paid enough attention. To investigate the aero-hydrodynamic characteristics of floating offshore wind turbines (FOWTs) with different layouts, fully coupled simulations for FOWTs under variable wind and wave conditions are necessary. In the present work, the unsteady actuator line model (UALM) is embedded into in-house CFD solver naoe-FOAM-SJTU to establish a fully coupled CFD analysis tool named FOWT-UALM-SJTU for numerical simulations of FOWTs. Coupled aero-hydrodynamic simulations of two OC3 Hywindspar FOWT models in both tandem and offset configurations under shear wind and regular wave conditions are performed. From the simulations, unsteady aerodynamic characteristics including the rotor power, thrust, and detailed wake flow information can be obtained, and hydrodynamic responses such as the six-degree-of-freedom motions and mooring tensions are also available. The coupled aero-hydrodynamic characteristics of FOWTs with different layouts are compared and analyzed. Strong wake interaction phenomena are observed and the blades of the downstream turbine are subject to a highly asymmetric interaction with the wake induced by the upstream turbine in offset configuration. The coupled aero-hydrodynamic performance of downstream FOWT is significantly influenced by the layout of FOWTs.

Keywords: Floating offshore wind turbines; Coupled aero-hydrodynamic simulations; Layout of wind turbines; FOWT-UALM-SJTU solver

Introduction

Along with the growing energy crisis and environmental crisis, the demands of renewable energy have become increasingly urgent. As one of the most promising non-polluting renewable energy, wind energy is developing rapidly in recent years. In addition, the wind turbines that are used to convert the wind energy into electricity have also achieved great development. To obtain more wind power, the rotor blade of wind turbines have become significantly larger. The wind turbine sizes have increased to multi-megawatt levels. Additionally, the wind turbine have experienced from onshore wind turbines to offshore wind turbines for gaining huge amount wind energy. Compared with onshore wind turbines, the offshore wind turbines have several advantages. The wind speed from the sea is much stronger and more uniform than it from the land, which means the offshore wind turbines can gain more wind energy than onshore wind turbines. The onshore wind turbines have constraints such as visual impact and noise emissions, while the offshore wind turbines can avoid these disadvantages and do not take up precious land resources^[1]. The using of offshore wind turbines has become a trend in the development of wind energy, and the offshore wind farms composed of multiple offshore

wind turbines for the development of huge amount offshore wind power have attracted a lot of attention.

In the offshore wind farms, the layout design has great influence on the initial investment cost, annual energy production, operation and maintenance cost during the service life time of wind turbines^[2]. An adequate wind farm layout design would lead to higher than expected wind power capture, decreased maintenance costs, longer service time, and so on. Many research works focusing on the layout design have been done. Bansal et al.^[3] improved the biogeography based optimization (BBO) and solved the wind farm layout optimization problem with non-uniform hub height and rotor radius based on fitness difference strategy (FD-BBO). Compared with numerical experiments on benchmark test problems, the proposed FD-BBO was proven to be an efficient optimization algorithm. Wang et al.^[4] presented a novel control strategy approach for the optimization of a simple square wind farm, and optimal wind farm design considering both the wind turbine placement and control were studied using three different optimization approaches. Rehman et al.^[5] proposed an optimization approach based on the cuckoo search (CS) algorithm for wind farm layout design. The proposed CS algorithms were compared with genetic and particle swarm optimization algorithms, and the comparative results including the yearly power output and efficiency showed the CS algorithms outperformed other optimization algorithms. Choi et al.^{[6][7]} performed numerical simulations of a 6 MW wind farm consisting of three sets of 2MW wind turbines in tandem configuration. The influence of the inter-turbine spacing on the aerodynamic power output, wake interaction and the dynamic responses of wind turbine was studied. They also investigated the aerodynamic performance of a wind farm with two sets of 2MW class wind turbines using a full 3-D wind turbine model. The effect of separation distance between two turbines on power output of the wind farm was studied. Fletcher and Brown^[8] studied the aerodynamic interaction between two wind turbines in both co-axial and offset configurations using vorticity transport model. The influence of horizontal space and longitudinal space on the aerodynamic interaction was discussed. Mikkelsen et al.^[9] analyzed the effect of wake interaction for three in-line model wind turbines in a wind farm based on the actuator line technique. Both full wake and half wake situations are considered, and detailed unsteady behavior of interacting wakes was captured.

With the progress in offshore wind energy, the floating wind farms are planned for huge amount of clean electricity recently. The floating offshore wind turbines (FOWTs) are usually clustered in the floating wind farms to decrease the overall installation and maintenance expenses, causing an adverse effect that the wind turbines generally experience a significant increased turbulence because of wake interaction from surrounding wind turbines^[10]. Considering the fact that the wake interaction between FOWTs has remarkable effects on the FOWT's power output, system dynamic responses and structural loadings, it should be paid enough attention. The wake interaction phenomena is observed and investigated originally in onshore wind farms, and many researches have been conducted to study the influence of complicated wake characteristics on wind turbines in onshore wind farms. Initially, different wake field models^{[11]-[14]} are developed for the wake calculations. But detailed turbulence characteristics in the wake flow, which have great influence on wake interaction, cannot be obtained. To better understand the complicated wake characteristics in wind farms, model tests^{[15][16]} are also conducted. The wake characteristics of wind turbines and the wake interaction phenomena are investigated based on wind tunnel tests. Considering the influence of scale effect on wake flow cannot be avoided in the model tests, CFD techniques that can consider turbulence characteristics in wake flow and eliminate the influence of scale effect become more and more popular in the study of wake interaction in wind farms. Churchfield et

al.^[17] investigated the influence of atmospheric stability and surface roughness on wind turbine dynamics. Numerical simulations for two wind turbines were conducted to study the wake effects under different surface roughness and atmosphere conditions. Troldborg et al.^[18] studied the wake interaction between two wind turbines based on the actuator line model. Different ambient turbulence intensities were taken into consideration in the simulations. The averaged velocity and turbulence fields as well as the development of wake generated vortex structure were extracted to understand the interacting wakes. For floating wind farms, the environment loads acting on the FOWTs are complex, and the coupling effects between wind turbine and floating platform make the wake interaction more complicated. Dörenkämper et al.^[19] studied the impact of the stratified atmospheric boundary layer on power production and wake effects in offshore wind farms by the means of measurements large-eddy simulations. Barthelmie et al.^[20] modelled the wake of large wind farms based on computational fluid dynamics models and analyzed the power losses due to wakes at offshore wind farms. Above all, far limited work have been done to investigate the complicated wake field characteristics in floating wind farms. It is necessary to study coupled aero-hydrodynamic responses of FOWTs and wake flow field in floating wind farms for optimal layout design.

In the present paper, the unsteady actuator line model (UALM) is embedded into in-house CFD solver naoe-FOAM-SJTU to establish a fully coupled CFD analysis tool named FOWT-UALM-SJTU for full-scale simulations of FOWTs. Coupled aero-hydrodynamic simulations of two OC3 Hywindspar FOWT models in both tandem and offset configurations under shear wind and regular wave conditions are performed. From the simulations, unsteady aerodynamic characteristics including the rotor power, thrust, and detailed wake flow information can be obtained, and hydrodynamic responses such as the six-degree-of-freedom motions and mooring tensions are also available. The coupled aero-hydrodynamic characteristics of FOWTs with different layouts are compared and analyzed to study the influence of layout of FOWTs.

Numerical Method

The Unsteady Actuator Line Model

The actuator line model (ALM) developed by Sørensen and Shen^[21] is a simplified method to study the aerodynamic performance of wind turbine. It is an effective way to replace the real tower surfaces with virtual actuator lines. In consequence, it acquires a benefit of not requiring to solve the blade geometry layer. The body forces distributed along the lines are calculated from the local attack angle and a look-up table of airfoil data. The main advantage of modeling the rotor of wind turbine using ALM is that the calculation resource can be greatly saved.

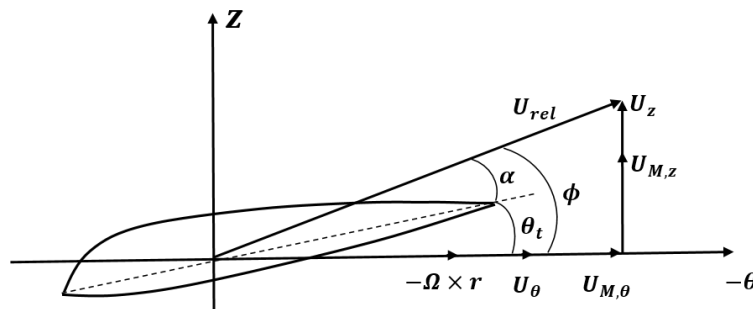


Fig. 1 Cross-sectional airfoil element

In the present work, modifications should be made to the initial ALM so that it can be used to simulate the FOWT. This is accomplished by accounting for the influence of the platform

motion (\mathbf{U}_M shown in Fig. 1) on the blades. Then the UALM used in this study is developed by modifying the initial ALM. To determine the body forces acting on the blades, a blade element approach combined with two-dimensional airfoil characteristics is used. As Fig. 1 shows, a cross-sectional element at radius r defines the airfoil at the (θ, z) plane. Denoting the tangential and axial velocity in the inertial frame of reference as \mathbf{U}_θ and \mathbf{U}_z , respectively. \mathbf{U}_M represents the added velocity vector induced by the motions of floating support platform, which will lead to complicated interactions between the rotor and its wake.

The integral velocity vector relationship can be described as:

$$\mathbf{U}_{rel} = \mathbf{U}_\theta - \boldsymbol{\Omega} \times \mathbf{r} + \mathbf{U}_z + \mathbf{U}_M \quad (1)$$

Where $\boldsymbol{\Omega}$ is the angular velocity of the rotor. The local velocity \mathbf{U}_{rel} relative to the rotating blade is calculated as:

$$|\mathbf{U}_{rel}| = \sqrt{(U_z + U_{M,z})^2 + (U_\theta - \Omega r + U_{M,\theta})^2} \quad (2)$$

Here $U_{M,\theta}$ and $U_{M,z}$ are the projections of \mathbf{U}_M on (θ, z) plane.

The attack angle is defined as:

$$\alpha = \phi - \theta_t \quad (3)$$

Where $\phi = \tan^{-1} \left(\frac{U_z + U_{M,z}}{U_\theta - \Omega r + U_{M,\theta}} \right)$ is the inflow angle. θ_t is the local twist angle. And the body force can be given by the following equation:

$$\mathbf{f} = (\mathbf{L}, \mathbf{D}) = \frac{\rho |\mathbf{U}_{rel}|^2 c N_b}{2r d \theta dz} (C_L \mathbf{e}_L + C_D \mathbf{e}_D) \quad (4)$$

Where c is the chord length; N_b is the number of blades; C_L and C_D are the lift and drag coefficient, respectively; \mathbf{e}_L and \mathbf{e}_D denote the unit vectors in the directions of the lift and the drag, respectively. The lift and drag coefficients are determined from measured or computed two-dimensional airfoil data that are corrected for three-dimensional effects.

The body force needs to be smoothed to avoid singular behavior before it is added into the momentum equations.

$$\mathbf{f}_\varepsilon = \mathbf{f} \otimes \eta_\varepsilon \quad (5)$$

where

$$\eta_\varepsilon(d) = \frac{1}{\varepsilon^3 \pi^{3/2}} \exp \left[-\left(\frac{d}{\varepsilon} \right)^2 \right] \quad (6)$$

Here d is the distance between the measured point and the initial force points on the rotor. ε is a constant which serves to adjust the strength of regularization function, and the influence of the parameter ε has been studied and some experienced conclusions have been obtained^[22].

The regularized force pre unit volume force can be written as:

$$\mathbf{f}_\varepsilon(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}) = \sum_{i=1}^N \mathbf{f}(\mathbf{x}_i, \mathbf{y}_i, \mathbf{z}_i, \mathbf{t}) \frac{1}{\varepsilon^3 \pi^{3/2}} \exp \left[-\left(\frac{d}{\varepsilon} \right)^2 \right] \quad (7)$$

Then \mathbf{f}_ε is added into the right hand of momentum equations as a source term.

Six-degree-of-freedom Motions

The six-degree-of-freedom motions of the floating support platform are predicted by in-house CFD solver naoe-FOAM-SJTU. Two coordinate systems (as shown in Fig. 2) are used in the procedure of solving six-degree-of-freedom motion equations. In each time step simulation, the motion equations are solved in platform-fixed coordinate system and the forces are calculated in earth-fixed coordinate system. And the added velocity induced by the motions of floating support platform is updated by the following equation:

$$\mathbf{U}_{M,i} = [\mathbf{J}](\mathbf{U}_c + \boldsymbol{\omega}_c \cdot (\mathbf{x}_i - \mathbf{x}_c)) \quad (8)$$

Where $[\mathbf{J}]$ is the transformation matrix defined from the platform-fixed coordinate to earth-fixed coordinate; \mathbf{U}_c and $\boldsymbol{\omega}_c$ donate the translation velocity and the angular velocity of the rotating center, respectively; \mathbf{x}_c is the position coordinate of the rotating center.

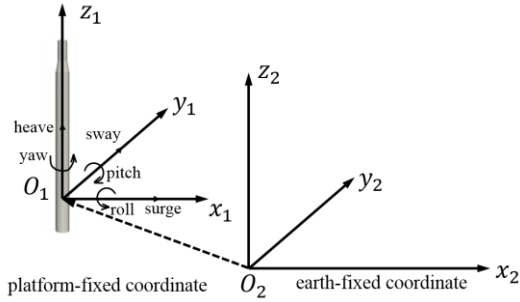


Fig. 2 Two coordinate systems

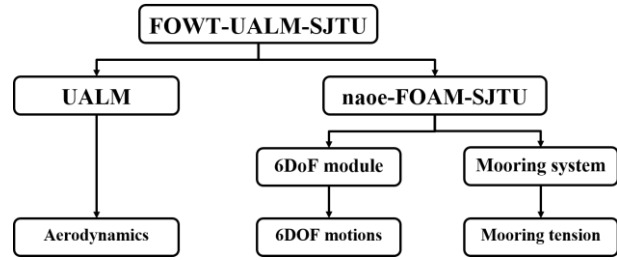


Fig. 3 Frame diagram of FOWT-UALM-SJTU

Coupled Aero-hydrodynamic Analysis Method

In the present work, the UALM is embedded into naoe-FOAM-SJTU to establish a fully coupled CFD analysis tool named FOWT-UALM-SJTU to study the coupled aero-hydrodynamic characteristics of FOWTs. As Fig. 3 shows, the aerodynamic forces can be got by the UALM, and the six-degree-of-freedom motions are predicted by the naoe-FOAM-SJTU. Moreover, the piecewise extrapolating method (PEM) is used to study the performance of the mooring system. It is static analysis method. The gravity and tensile deformation of mooring lines are both taken into consideration in the calculation of mooring tensions.

In FOWT-UALM-SJTU solver, VOF (Volume of Fluid) method with bounded compression technique is used to solve two-phase flow problem with free surface. The $k-\omega$ SST turbulence model is applied to solve the RANS equation. And the governing equations can be written as:

$$\nabla \cdot \mathbf{U} = 0 \quad (9)$$

$$\frac{\partial \rho \mathbf{U}}{\partial t} + \nabla \cdot (\rho (\mathbf{U} - \mathbf{U}_g)) \mathbf{U} = -\nabla p_d - \mathbf{g} \cdot \mathbf{x} \nabla \rho + \nabla \cdot (\mu_{eff} \nabla \mathbf{U}) + (\nabla \mathbf{U}) \cdot \nabla \mu_{eff} + \mathbf{f}_\sigma + \mathbf{f}_s + \mathbf{f}_\varepsilon \quad (10)$$

Where \mathbf{U} is velocity of field; \mathbf{U}_g is the velocity of mesh points; $p_d = p - \rho \mathbf{g} \cdot \mathbf{x}$ is the dynamic pressure, subtracting hydrostatic component from total pressure; \mathbf{g} is the gravity of acceleration vector; ρ is the mixture density with two phases; $\mu_{eff} = \rho(\nu + \nu_t)$ is effective dynamic viscosity, in which ν and ν_t are kinematic viscosity and eddy viscosity respectively; \mathbf{f}_σ is the surface tension term in two phases model and takes effect only on the liquid free surface; \mathbf{f}_s is the source term for sponge layer, which is set to avoid the wave reflection at the end of the tank and takes effect only in sponge layer; \mathbf{f}_ε is the body force calculated from UALM, representing the effects of turbine blades on the flow field.

The solving procedure of coupled aero-hydrodynamic simulations for the FOWTs is shown in Fig. 4. Coupling effects between the wind turbine, floating platform and mooring system are considered. It can be found that the calculated motion responses are inputs of UALM. The calculation of body force needs the information of the motions of the floating platform. In addition, the calculated aerodynamic forces are also inputs of the calculation of six-degree-of-freedom (6DoF) motions. It can be seen that the aerodynamic forces obtained from the UALM are added into the 6DoF motion equations.

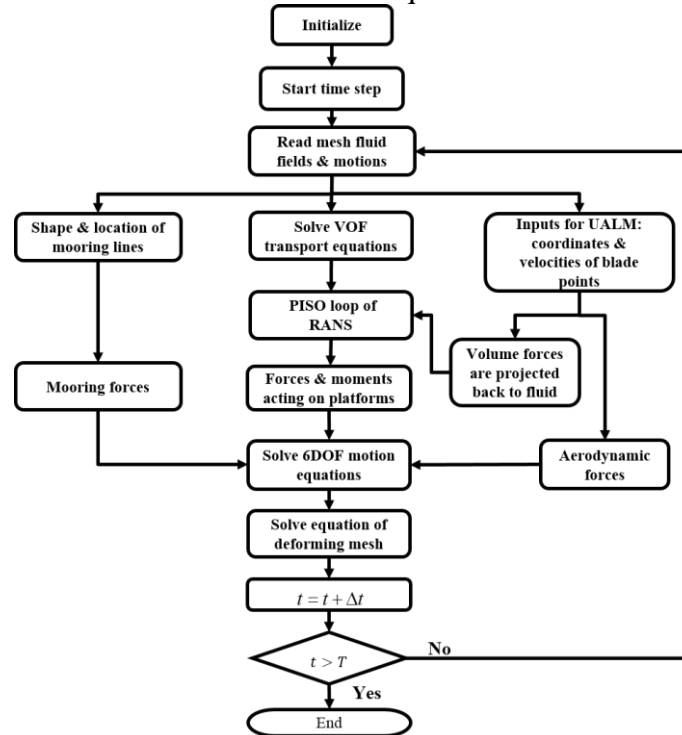


Fig. 4 Solving procedure of coupled aero-hydrodynamic simulations

Simulation Conditions

Geometric Model

The FOWT model used in the present work is OC3 Hywindspar FOWT model. General arrangement of the FWOT system is shown in Fig. 5. The wind turbine is NERL offshore 5-MV baseline wind turbine, which is a conventional three-bladed, upwind, variable-speed and variable blade-pitch-to-feather controlled turbine. The main specifications of the wind turbine are given in Table 1^[23].

Table 1 Specifications of NERL 5-MW turbine

Rating	5 MW
Rotor Orientation, Configuration	Upwind, 3 Blades
Control	Variable Speed, Collective Pitch
Drivetrain	High Speed, Multiple-Stage Gearbox
Rotor, Hub Diameter	126 m, 3 m
Hub Height	90 m
Cut-in, Rated, Cut-out Wind Speed	3 m/s, 11.4 m/s, 25 m/s
Cut-in, Rated Rotor Speed	6.9 rpm, 12.1 rpm
Rated Tip Speed	80 m/s
Overhang, Shaft Tilt, Precone Angle	5 m, 5°, 2.5°
Rotor Mass	110,000 kg

Nacelle Mass	240,000 kg
Tower Mass	347,460 kg
Coordinate Location of Overall Center of Mass	(-0.2 m, 0.0 m, 64.0 m)

The floating support platform is the spar-buoy concept platform called Hywindspar applied in OC3 project, and detailed information is listed in Table 2^[24].

Table 2 Specifications of Hywindspar platform

Depth to Platform Base Below SWL (Total Draft)	120 m
Elevation to Platform Top (Tower Base) Above SWL	10 m
Depth to Top of Taper Below SWL	4 m
Depth to Bottom of Taper Below SWL	12 m
Platform Diameter Above Taper	6.5 m
Platform Diameter Below Taper	9.4 m
Platform Mass, Including Ballast	7,466,330 kg
CML Location Below SWL Along Platform Center Line	89.9155 m
Platform Roll Inertia about CM	4,229,230,000 kg·m ²
Platform Pitch Inertia about CM	4,229,230,000 kg·m ²
Platform Yaw Inertia about Platform Centerline	164,230,000 kg·m ²

The mooring system consisting of three mooring lines is symmetrically distributed around the platform. Main characteristics of the mooring system are shown in Table 3. And the arrangement of the mooring lines is shown in Fig. 6.

Table 3 Parameters of mooring system for OC3-HywindSpar platform

Number of Mooring Lines	3
Angle Between Adjacent Lines	120°
Depth to Anchors Below SWL (water depth)	320 m
Depth to Fairleads Below SWL	70.0 m
Radius to Anchors From Platform Centerline	853.87 m
Radius to Fairleads From Platform Centerline	5.2 m
Unstretched Mooring line length	902.2 m
Mooring Line Diameter	0.09 m
Equivalent Mooring Line Mass Density	77.7066 kg/m
Equivalent Mooring Line Mass Weight in Water	689.094 N/m
Equivalent Mooring Line Extensional Stiffness	384,243,000 N
Additional Yaw Spring Stiffness	98,340,000 Nm/rad



Fig.5 Sketch of FOWT system

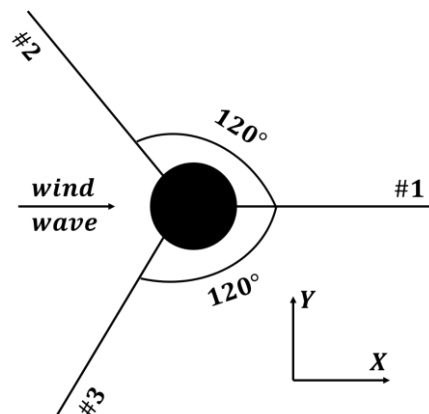


Fig. 6 Mooring system

Simulation Conditions

To investigate the aero-hydrodynamic characteristics of FOWTs with different layouts, fully coupled simulations for FOWTs under variable wind and wave conditions are performed. In this work, coupled aero-hydrodynamic simulations of two OC3 Hywindspar FOWT models in both tandem and offset configurations under shear wind and regular wave conditions are conducted. Detailed simulation cases are listed in Table 4.

Wind and wave conditions are kept in the same in these two cases. Wave period and wave length are $T = 10\text{s}$ and $\lambda = 156\text{m}$, respectively. And the wave height is $H = 4\text{m}$. Considering the characteristics of height-dependent wind speed, exponential model is used to describe wind shear.

$$u_z = u_0 \times \left(\frac{z}{90}\right)^{0.143} \quad (11)$$

Where u_z is the wind velocity at the height of z , u_0 is the wind velocity at the height of hub center. And the wind speed in these case are kept in the same at $u_0 = 5\text{ m/s}$.

Table 4 Simulation cases

Case Number	Distance along x direction	Distance along y direction
Case 1 (tandem case)	$2D$	0
Case 2 (offset case)	$2D$	$0.5D$

* $D = 126\text{m}$ is the diameter of the rotor.

Computation Domain and Boundary Condition

All cases adopt the same computation domain. The length and width of computation domain are 5λ and 2λ (λ is wave length), respectively. Considering the expansion effect of turbine wake, the height of air phase is $2D$ ($D = 126\text{m}$ is the diameter of the rotor). The depth of water phase is set to be 70% of the real water depth ($d = 320\text{m}$), for the effect of the water depth on the motion responses can be ignored at that water depth. The length of sponge layer before outlet boundary is 100m . The computational domain is shown in Fig. 7.

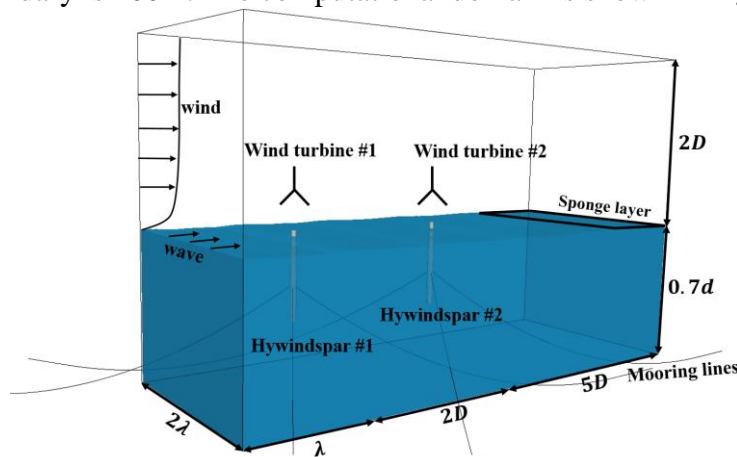


Fig. 7 Computational domain

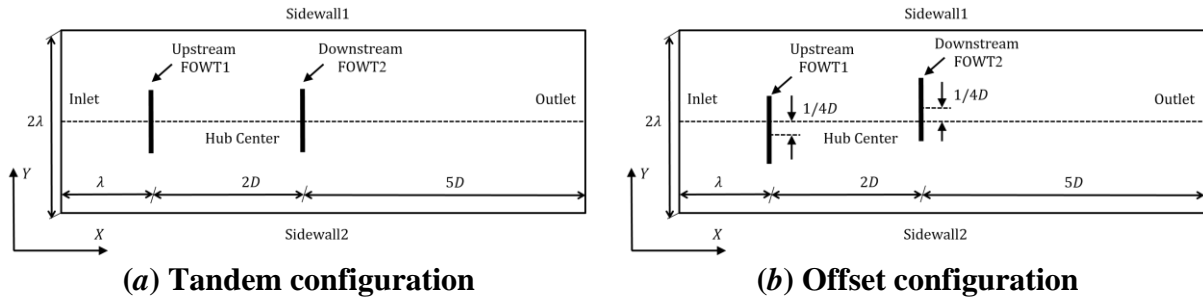
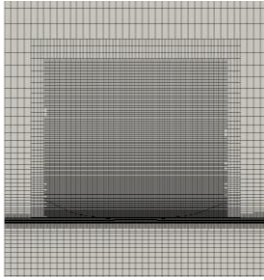


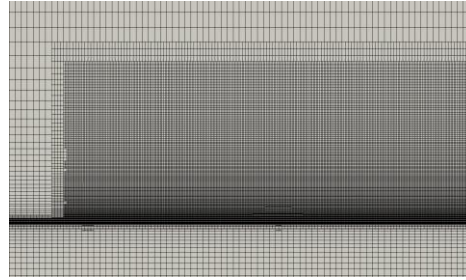
Fig. 8 Different arrangements of FOWTs

In these two cases, the distance from the inlet boundary to upstream FOWT is λ . The distance from the downstream FOWT to outlet boundary is $5D$. For FOWTs in tandem case, the longitudinal distance between these two FOWTs is $2D$, and the horizontal distance between two FOWTs is $0m$. For FOWTs in offset case, the longitudinal distance between these two FOWTs is also $2D$, while the horizontal distance between two FOWTs is $0.5D$. Different arrangement of FOWTs in shown in Fig. 8.

To capture the wake flow of wind turbine, the refined grids are utilized in the region behind the wind turbine. And the grids near the water surface are refined to capture the free surface. The grid distribution is shown in Fig. 9



(a) Grid in lengthwise section



(b) Grid in cross section

Fig. 9 Grid distribution

The boundary conditions are shown below:

- (1) Inlet boundary: velocity condition is wave inlet condition, and pressure condition is Neumann boundary condition that the normal gradient of pressure is equal to zero;
- (2) Outlet boundary: velocity condition is inletoutlet condition defined in OpenFOAM, and pressure condition is Dirichlet boundary condition that the pressure is constant;
- (3) Top boundary: both velocity condition and pressure condition are Dirichlet boundary conditions;
- (4) Bottom boundary: both velocity condition and pressure condition are slip conditions;
- (5) Left boundary and right boundary: boundary conditions are defined as symmetry plane that directional derivative perpendicular to the boundary is equal to zero;
- (6) Body surface: the moving wall boundary condition is adopted.

Results and Discussion

Aerodynamic Loads

Unsteady aerodynamic loads including the rotor power and thrust are presented here to analysis the influence of layout on the aerodynamic performance of FOWTs. The time history curves of rotor power and thrust of FOWTs in tandem and offset configurations are shown in Fig. 10 and Fig. 11, respectively.

The rotor power and thrust of FOWTs in two simulation cases both fluctuate greatly and change periodically, and the change period is approximately equal to the incident wave period. It indicates that this fluctuation and periodicity of the aerodynamic loads is in a large part due to the motions of floating support platform. The platform motions have significant effects on the unsteady aerodynamic performance of the FOWTs.

The time-averaged values of aerodynamic loads are listed in table 5. The rotor power and thrust of downstream FOWT are obviously smaller than those of upstream FOWT. In tandem case, the rotor power and thrust of downstream FOWT are only 30% and 62% compared to upstream FOWT. And the rotor power and the thrust of the downstream FOWTs are 78% and 85% compared to upstream FOWT in offset case, respectively. Due to the wake interaction between the FOWTs, the downstream FOWT experiences lower incoming wind velocity and higher turbulence intensity compared with the upstream FOWT, resulting in the aerodynamic loads decrease of downstream FOWT. It suggests the aerodynamic loads of downstream FOWT are affected significantly by the wake from upstream FOWT.

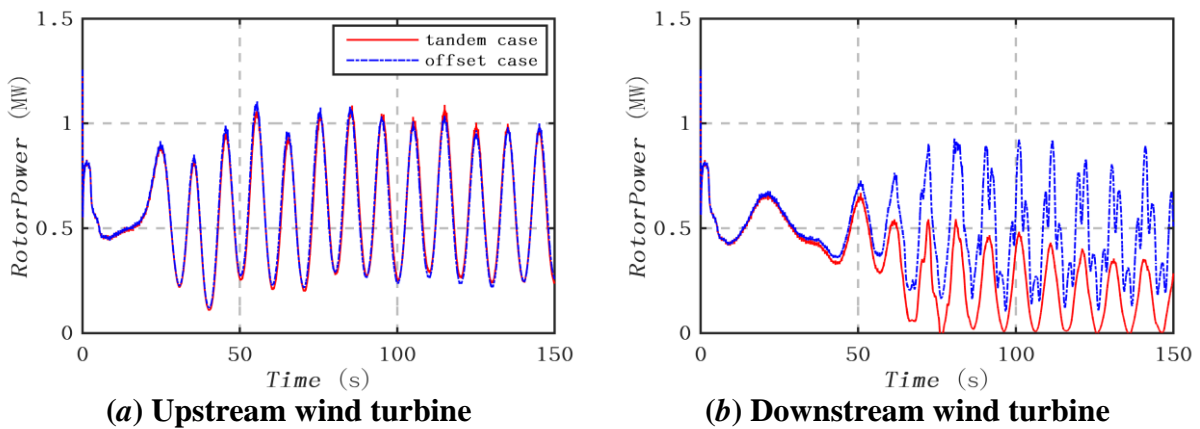


Fig. 10 Time history curves of rotor power in tandem and offset configurations

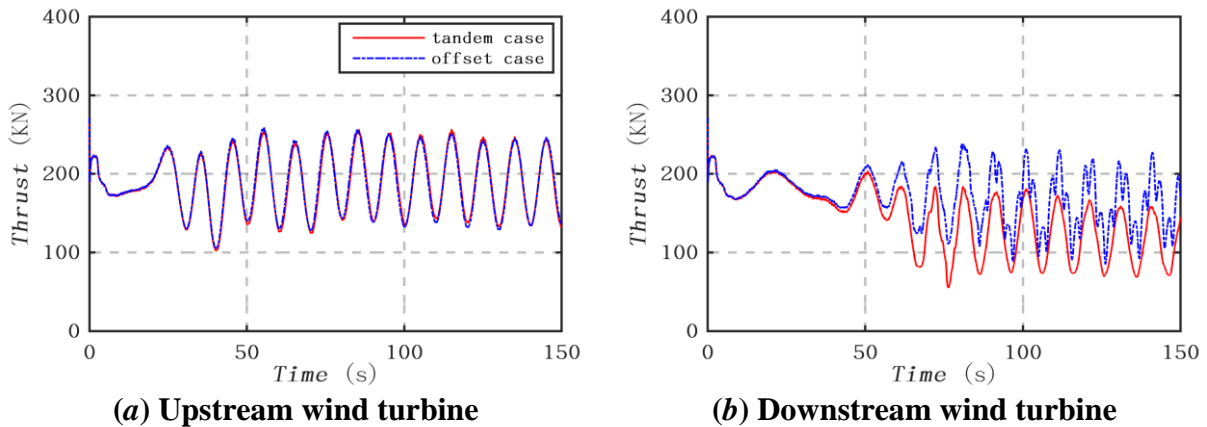


Fig. 11 Time history curves of thrust in tandem and offset configurations

Table 5 Rotor power and thrust of FOWTs in tandem and offset configurations

		Power (kW)	Power ratio (FOWT2/FOWT1)	Thrust (kN)	Thrust ratio (FOWT2/FOWT1)
Case 1	FOWT 1	610	30%	193	62%
	FOWT 2	182		119	
Case 2	FOWT 1	588	78%	189	85%
	FOWT 2	459		161	

*FOWT 1 and FOWT 2 represent the upstream FOWT and downstream FOWT, respectively.

Table 6 Comparisons between aerodynamic loads in tandem and offset configurations

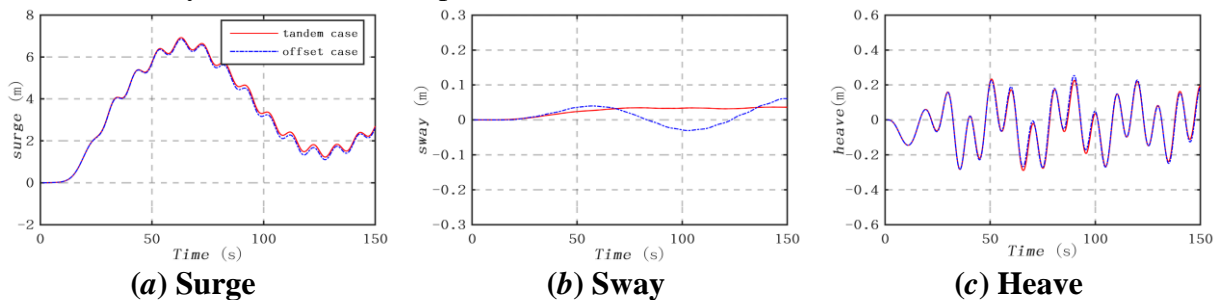
		Power (kW)	Power ratio (Case1/Case2)	Thrust (kN)	Thrust ratio (Case1/Case2)
FOWT 1	Case 1	610	96%	193	98%
	Case 2	588		189	
FOWT 2	Case 1	182	40%	119	74%
	Case 2	459		161	

The comparisons between rotor power and thrust in tandem and offset configurations are shown in table 6. For upstream FOWT, the time-average aerodynamic loads in tandem case are almost identical to those in offset case. And the aerodynamic loads of upstream FOWT in different layouts shown in Fig. 10(a) and Fig. 11(a) show little discrepancy. It means that the FOWT layouts have little effect on the aerodynamic loads of the upstream FOWT. While the situation is quite different for downstream FOWT. The time averaged value of the rotor power and thrust of downstream FOWT in tandem case are 40% and 74% compared to those in offset case. And notable discrepancy between the aerodynamic loads of downstream FOWT in tandem case and offset case can be found in Fig. 10(b) and Fig. 11(b). The variation progress of aerodynamic loads in offset case is more complex than that in tandem case. In tandem case, the downstream FOWT is in full wake, while the downstream FOWT is in half wake in offset case. So the average incoming wind velocity and turbulence intensity for downstream FOWT in offset case are larger than those in tandem case, which leads to the discrepancy in the aerodynamic loads of downstream FOWT in offset case and tandem case. It suggests that FOWT layouts have significant effects on the aerodynamic loads of the downstream FOWT.

Above all, it can be found that the aerodynamic loads are greatly influenced by the motions of floating support platform. The rotor power and thrust both fluctuate greatly and vary periodically. The aerodynamic loads of downstream FOWT are much smaller than those of upstream FOWT. The FOWT layouts have little influence on the aerodynamic loads of upstream FOWT, and the offset layout of the FOWTs have beneficial effects on the aerodynamic loads of downstream FOWT compared with the tandem layout.

Platform Motions

The floating support platform is an important part of the FOWT system. The platform motions have significant effects on the aerodynamic performance of wind turbine, and the aerodynamic forces will act on the floating platform and influence the motion responses in turn. There are complicated coupling effects between the floating support platform and wind turbine. The motion responses of floating support platforms in different layouts are shown in Fig. 12 and Fig. 13. The platform motions are compared and analyzed to investigate the influence of layouts on motion responses.



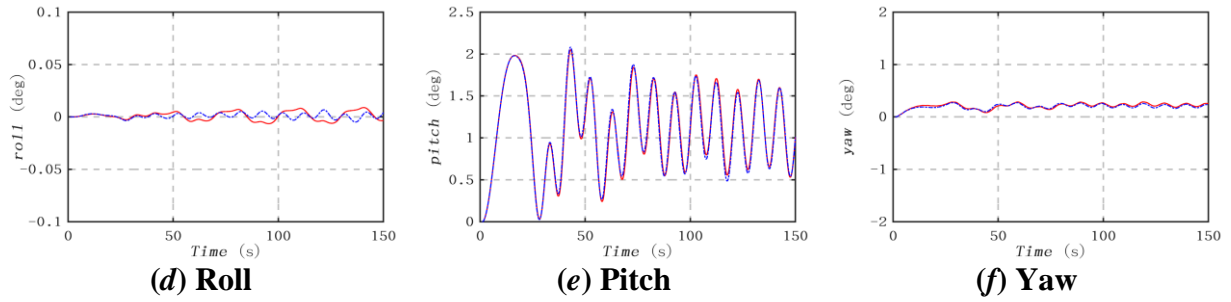


Fig. 12 Motion responses of upstream platform in tandem and offset configurations

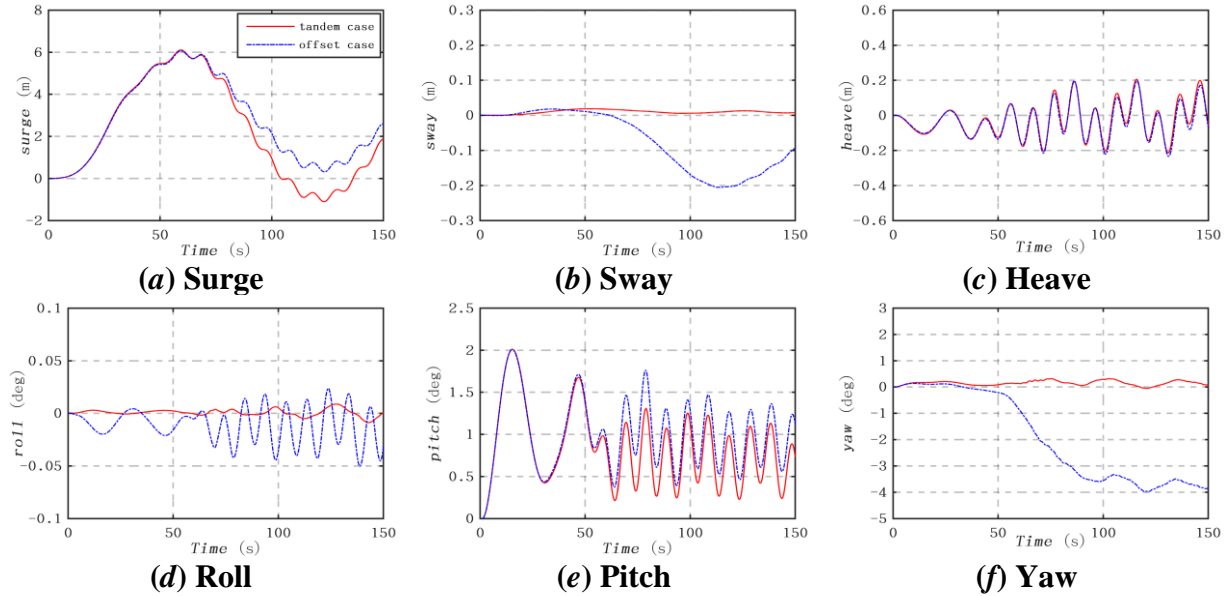


Fig. 13 Motion responses of downstream platform in tandem and offset configurations

For the upstream platform, the six-degree-of-freedom motions in tandem case are much the same with the motions in offset case. It is because the loads acting on the upstream platform in different cases are almost identical. The loads acting on the floating support platform include the aerodynamic forces, wave loads and mooring forces. And these forces acting the upstream platform are nearly the same in different cases. It means the motion responses of the upstream platform have little difference in different FOWT layouts.

For the downstream platform, there is large difference between the motions in tandem and offset configurations. It can be seen that the average value of pitch motion of downstream platform in offset case is larger than that in tandem case. This large amplitude of pitch motion will lead to serious interaction between the rotor and its wake. As motioned above, the aerodynamic forces of downstream FOWT in offset case are larger than those in tandem case, and the aerodynamic forces will act on the floating support platform. It indicates the loads acting on the downstream platform in offset case are larger than those in tandem case, which explains the discrepancy of pitch motions of downstream platform between the tandem case and offset case. In addition, the amplitudes of sway and yaw motions of the downstream platform in tandem case are much larger than those in offset case. When the FOWTs is in offset configuration, the downstream FOWT is in half wake of the upstream FOWT. Affected by the wake interaction, the incoming wind velocity of the part of rotor in the wake are much larger than that not in the wake. This causes the forces acting on the rotor plane are not uniform. So the yaw and sway motions of the downstream platform in offset case are much

larger than those in tandem case. For surge, roll and heave motions of the downstream platform, there is little discrepancy between the tandem case and the offset case.

Wake Field

The wake interaction between the FOWTs in both tandem and offset configurations are clearly observed. Detailed wake field characteristics are presented here to study the wake interaction between two FOWTs in different layouts. Fig. 14 shows the contours of the axial direction wind velocity in the horizontal plane at the reference height $z = 90\text{m}$ (the height of the center of rotor) for tandem case and offset case, respectively. The expansion of stream-tube is observed in both simulation cases. That the tangential velocity increases with decreasing axial direction wind velocity leads to this phenomena. In tandem case, the downstream FOWT is in full wake of the upstream FOWT. It can be seen the incoming wind velocity of the downstream FOWT decreases greatly compared with that of the upstream FOWT, which explains why the aerodynamic loads downstream FOWT are much smaller. In offset case, the downstream FOWT is in half wake of the upstream FOWT, so the incoming wind velocity of the part of downstream FOWT in the wake is smaller than that not in the wake. Furthermore, the asymmetric forces distributed on the rotor plane result in the yaw motion of the downstream FOWT. Compared with incoming wind velocity of the downstream FOWT in tandem case, the downstream FOWT in offset case experiences larger incoming wind velocity. So the aerodynamic loads of downstream FOWT in offset case are bigger than those in tandem case.

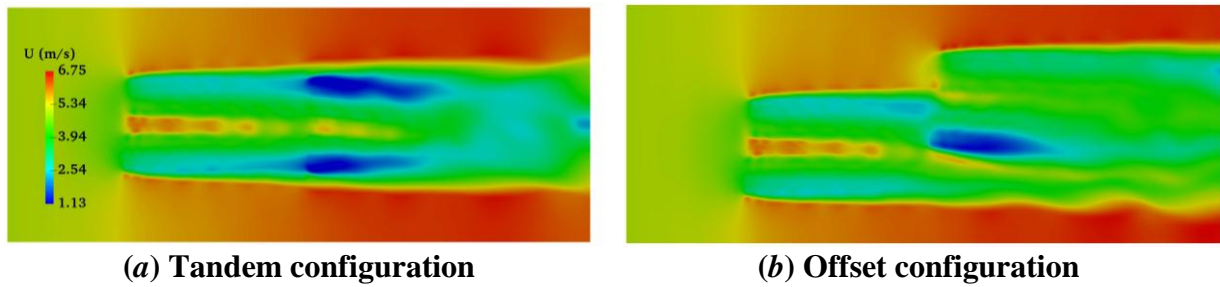


Fig. 14 Axial direction wind velocity counters in horizontal plane at the reference height $z = 90\text{m}$

Profiles of the streamwise velocity in horizontal plane through the center of wind turbine rotor for different cases are presented in Fig. 15. The velocity profiles before the downstream FOWT in different cases are almost the same, while the velocity profile behind the downstream FOWT in offset case is quite different with that in tandem case. The velocity field in offset case is more complicated, and the velocity deficit region in offset case is much larger.

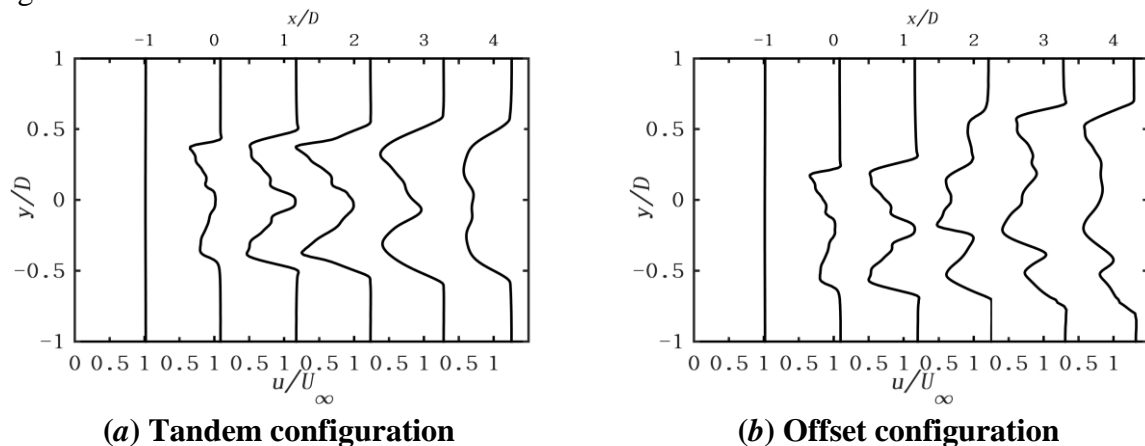


Fig. 15 Profiles of the streamwise velocity in horizontal plane through the center of wind

turbine rotor for different layouts

The evolution of wake vortex at different times of an entire wave circle in coupled case is illustrated in Fig. 16. The wave is contoured by wave height and the mooring lines are represented by black lines. The second-order invariant of velocity gradient tensor Q is used to visualize the wake vortex. Clearly spiral tip vortex from the upstream FOWT can be captured in both tandem and offset cases, while this vorticity is quickly diffused in the downstream. Affected by the wake of the upstream FOWT, the downstream FOWT experiences increased turbulence and the vortex structures become more unstable. In addition, the tip vortex from the downstream FOWT in tandem case are different from that in offset case. The tip vortex from downstream FOWT is much more affected by the wake of upstream FOWT in tandem case. Moreover, the platform motions lead to the interaction between the rotor and its wake and increase the instability of wake field.

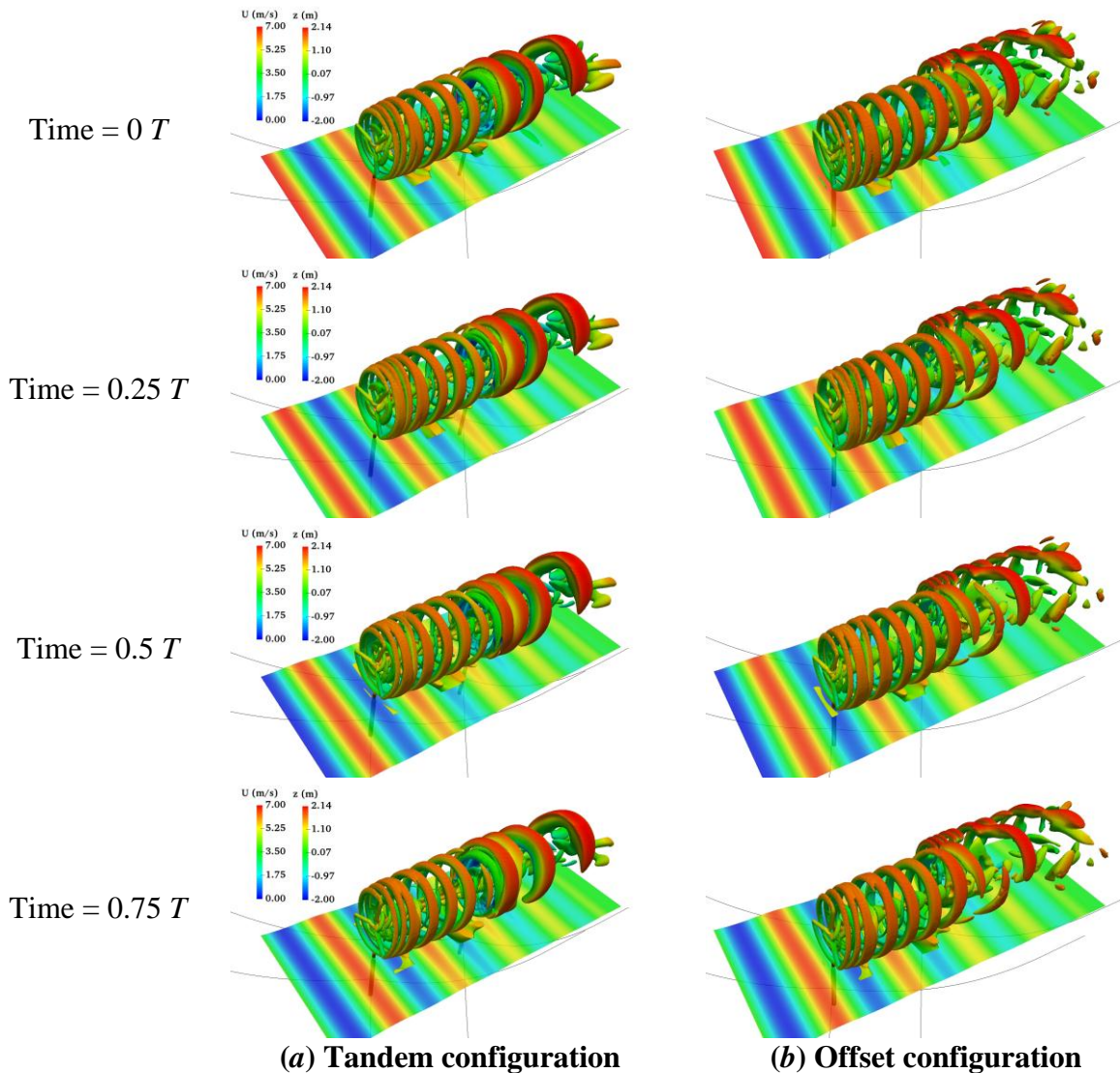


Fig. 16 Instantaneous vortex structure of the rotor in tandem and offset cases

Conclusions

In this study, the unsteady actuator line model (UALM) is embedded into in-house CFD solver naoe-FOAM-SJTU to establish a fully coupled CFD analysis tool named FOWT-UALM-SJTU for full-scale simulations of FOWTs. Coupled aero-hydrodynamic

simulations of two OC3 Hywindspar FOWT models in both tandem and offset configurations under shear wind and regular wave conditions are performed. The aerodynamic loads including the rotor power and thrust, the six-degree-of-freedom motions and detailed wake field characteristics are obtained and analyzed. It can be found that aerodynamic loads are greatly influenced by the motions of floating support platform, which causes the rotor power and thrust both fluctuate greatly and change periodically. Affected by the wake interaction, the aerodynamic loads of downstream FOWT are much smaller than those of upstream FOWT. The FOWT layouts have little influence on the aerodynamic loads of upstream FOWT, and the offset configuration of the FOWTs have beneficial effects on the aerodynamic loads of downstream FOWT compared with the tandem configuration. For platform motions, the FOWT layouts have significant effects on the pitch, sway and yaw motions of downstream platform. While the surge, roll and heave motions are little influenced. The amplitudes of pitch, sway and yaw motions of downstream platform in offset configuration are much larger than those in tandem configuration. The wake field becomes more complicated affected by the wake interaction between the FOWTs. The tip vortex of upstream FOWT is clear, while the vortex structure of downstream FOWT is highly unstable. In addition, the wake characteristics of downstream FOWT in tandem configuration are much influenced the wake of upstream FOWT. In the future, the influence of inter-turbine spacing between FOWTs on wake field characteristics will be studied for the optimal layout design of floating wind farms.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

- [1] Shen, X., Chen, J., Hu, P., Zhu, X., and Du, Z. (2018) Study of the unsteady aerodynamics of floating wind turbines, *Energy* 145, 793-809.
- [2] Troldborg, N., Larsen, G. C., Madsen, H. A., Hansen, K. S., Sørensen, J. N., and Mikkelsen, R. (2011) Numerical simulations of wake interaction between two wind turbines at various inflow conditions, *Wind Energy* 14(7), 859-876.
- [3] Bansal, J. C., Farswan, P., and Nagar, A. K. (2018) Design of wind farm layout with non-uniform turbines using fitness difference based BBO, *Engineering Applications of Artificial Intelligence* 71, 45-59.
- [4] Wang, L., Tan, A., and Gu, Y. (2016) A novel control strategy approach to optimally design a wind farm layout, *Renewable Energy* 95, 10-21.
- [5] Rehman, S., Ali, S. S., and Khan, S. A. (2017) Wind farm layout design using cuckoo search algorithms, *Applied Artificial Intelligence* 30(10), 899-922.
- [6] Choi, N. J., Sang, H. N., Jeong, J. H., and Kim, K. C. (2014) CFD study on aerodynamic power output changes with inter-turbine spacing variation for a 6 mw offshore wind farm, *Energies* 7(11), 7483-7498.
- [7] Choi, N. J., Sang, H. N., Jeong, J. H., and Kim, K. C. (2013) Numerical study on the horizontal axis turbines arrangement in a wind farm: effect of separation distance on the turbine aerodynamic power output, *Journal of Wind Engineering & Industrial Aerodynamics* 117(117), 11-17.
- [8] Fletcher, T. M., and Brown, R. E. (2010) Simulation of wind turbine wake interaction using the vorticity transport model, *Wind Energy* 13(7), 587-602.
- [9] Mikkelsen, R., Sørensen, J. N., Øye, Stig, and Troldborg, N. (2007) Analysis of power enhancement for a row of wind turbines using the actuator line technique, *Journal of Physics Conference Series* 75, 012044.
- [10] Kusiak, A., and Song, Z. (2010) Design of wind farm layout for maximum wind energy capture, *Renewable Energy* 35(3), 685-694.

- [11] Katic, I., Højstrup, J., and Jensen, N. O., A simple model for cluster efficiency, *Proceedings European Wind Energy Association Conference and Exhibition*, Rome, Italy, 1986, 407-409.
- [12] Crespo, A., Manuel, F., and Hernández, J., Numerical modelling of wind turbine wakes, *Proceedings of the 1990 European Communication Wind Energy Conference*, Madrid, Spain, 1990, 166.
- [13] Ainslie, J. F., Development of an eddy viscosity model for wind turbine wakes, *Proceedings 7th BWEA Wind Energy Conference*, Oxford, 1985, 61-66.
- [14] Larsen, G. C., Højstrup, J., and Madsen, H. A., Wind fields in wakes, *Proceedings 1996 European Union Wind Energy Conference*, Göteborg, 1996, 764-768.
- [15] Whale, J., Anderson, C. G., An experimental investigation of wind turbine wakes using particle image velocimetry, *1993 European Communication Wind Energy Conference and Exhibition*, Lübeck-Travemünde, Germany, 1993, 457.
- [16] Vermeer, L. J., Sørensen, J. N., and Crespo, A. (2003) Wind turbine wake aerodynamics, *Progress in Aerospace Science* 39(6–7), 467-510.
- [17] Churchfield, M., Sang Lee, John Michalakes, and Moriarty, P. (2012) A numerical study of the effects of atmospheric and wake turbulence on wind turbine dynamics, *Journal of Turbulence* 13(14), 1-32.
- [18] Troldborg, N., Larsen, G. C., Madsen, H. A., Hansen, K. S., Sørensen, J. N., and Mikkelsen, R. (2011) Numerical simulations of wake interaction between two wind turbines at various inflow conditions, *Wind Energy* 14(7), 859-876.
- [19] Dörenkämper, M., Witha, B., Steinfeld, G., Heinemann, D., and Kühn, M. (2015) The impact of stable atmospheric boundary layers on wind-turbine wakes within offshore wind farms, *Journal of Wind Engineering & Industrial Aerodynamics* 144, 146-153.
- [20] Barthelmie, R. J., Pryor, S. C., Frandsen, S. T., Hansen, K. S., Schepers, J. G., and Rados, K., et al. (2010) Quantifying the impact of wind turbine wakes on power output at offshore wind farms, *Journal of Atmospheric & Oceanic Technology* 27(8), 1302-1317.
- [21] Sørensen, J. N., and Shen, W. Z. (2002) Numerical modeling of wind turbine wakes, *Journal of Fluids Engineering* 124(2), 393.
- [22] Sørensen, J. N., Shen, W. Z., and Munduate, X. (2015) Analysis of wake states by a full - field actuator disc model, *Wind Energy* 1(2), 73-88.
- [23] Jonkman, J., Butterfield, S., Musial, W., and Scott, G. (2009) Definition of a 5-mw reference wind turbine for offshore system development, *Office of Scientific & Technical Information Technical Reports*, 1–75.
- [24] Jonkman, J., and Musial, W. (2010) Offshore code comparison collaboration (oc3) for IEA wind task 23 offshore wind technology and deployment, *Office of Scientific & Technical Information Technical Reports* 303, 275-3000.

Numerical Calculations for Smooth Circular Cylinder Flow at 3900 Reynolds Numbers with SST-IDDES Turbulence Model

Jiawei HE^{*}, Weiwen ZHAO, Decheng WAN[‡]

School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

^{*}Presenting author: jiaweihe@sjtu.edu.cn

[‡]Corresponding author: dcwan@sjtu.edu.cn
<http://dcwan.sjtu.edu.cn>

Abstract

The turbulent flow around a cylinder is simulated and analysed physically at $Re = 3900$, by means of hybrid RANS-LES turbulence modelling approaches. The numerical investigations are conducted by using naoe-FOAM-SJTU, a viscous flow solver (naoe-FOAM-SJTU) which is developed and based on the popular open source toolbox OpenFOAM. To simulate the large separated flow in high Reynolds Numbers, turbulence closure has been chosen the Shear Stress Transport (SST) based Improved Delay Detached eddy simulation (IDDES). In the present study, the benefits of these hybrid approaches have been discussed for capturing the vortex dynamics and frequency modes. The coefficients of drag C_d and Strouhal number (St) were analyzed and compared with experimental results. Meanwhile, the boundary layer transition phenomenon and the flow characteristics were more accurately predicted. The good agreement validated the CFD solver naoe-FOAM-SJTU is applicable and reliable to study such problems.

Keywords: Circular Cylinder; Reynolds number; SST-IDDES; CFD.

Introduction

The fixed smooth cylinder-flow problem is considered to be a prominent research subject in computational fluid dynamics (CFD). Despite the reasonably simple geometry and grid-making, the very complex nature of this flow makes it extremely difficult to calculate. In this problem, the boundary layer, free-shear layers, and wake interact and the laminar-turbulent transition and boundary-layer separation move as Reynolds number varies. Also, several instabilities associated with the shear layers play an important role influencing the flow behavior in the different regimes. This particular issue is not only an academic challenge, but mainly an important engineering problem, being present in many fields, such as the vortex induced motion in the offshore industry. In this field, the vortex-induced vibrations of risers and vortex-induced motions of floating units stand out as a very important and current demand for which no complete and general solution has been developed.

As well known, the behavior of flow over a circular cylinder changes dramatically with the changes of Reynolds number. At low Reynolds number, the flow is laminar. While in region with higher Re , the turbulence appears. With increasing Reynolds number, the onset of transition moves upstream towards the separation point and, simultaneously, the large-scale vortices are formed closer to the base of the cylinder. The flow over a cylinder has been the subject of various experimental^[1-3] and numerical studies^[4-11]. Though, Kravchenko et al^[5]

argued that the experiment of Lourenco and Shih^[3] might have been contaminated by external disturbances which led to premature transition of the shear layers and thus to a shorter recirculation bubble. Nevertheless, the majority of the computational studies with CFD for $Re = 3900$, mainly chosen the experimental data of Lourenco and Shih and Ong and Wallace. Dong^[7] investigated the near wake flow of a circular cylinder with PIV and direct numerical simulation (DNS) at $Re = 3900/4000$ and $10\,000$. They present a PIV/DNS comparison of mean and turbulent contour maps and focus on shear layer instability with their DNS data. WANG^[12] has developed a finite-volume TVD numerical model to simulate an unsteady incompressible flow around a circular cylinder for subcritical Reynolds numbers of 1000 , 3900 and 1×10^4 . The method presented by WANG is validated by comparing with the available experimental data and numerical results. Charles Mockett^[8] presented for the detached-eddy simulation (DES) of the flow around a circular cylinder at a high sub-critical Reynolds number (1.4×10^5). Good comparability with PIV experimental field data is facilitated by the confined and clearly-defined geometry, although some uncertainty remains regarding the free stream turbulence intensity. The combination of DES with an appropriate low-dissipative hybrid numerical convection scheme and high temporal resolution delivers excellent agreement with the experiment for the time and phase- averaged fields as well as the spectral content. Krishnan^[13] used DES97 and DDES to predict the massively separated flow around a circular at Reynolds numbers based on the cylinder diameter and freestream velocity of 1.4×10^5 and 8×10^6 . The DES97 and DDES predictions are in reasonable agreement with experimental measurements of the pressure coefficient. In a numerical experiment the possibility to influence the shear layer dynamics comes through the span-wise domain size chosen, the boundary conditions. Enforced the numerical errors inherent to the schemes used, the underlying grid.

The main objective of the present paper is to help building confidence of SST-IDDES turbulence model for massively separated flows. The recent publications by several authors of numerical simulations of the cylinder flow at $Re = 3900$ led us to choose this case for the validation of SSTIDDES method in our solver. Flow around a cylinder at subcritical Reynolds number ($Re = 3900$) was extensively studied numerically. It has become a benchmarking case and good agreements for mean stream-wise velocity obtained at both near wake and far downstream compared with the experimental data. To systematically validate the performance of SST-IDDES, the widely used hybrid RANS/LES model DDES is also presented in this paper for comparison.

Turbulence Modeling

Spalart^[14] proposed the DES Detached Eddy Simulation approach to bridge the gap between RANS and LES. DES is a promising concept that could enable full scale engineering applications at high Reynolds number to be computed within the resources available today or in the near future^[15]. To address this drawback, delayed-detached eddy simulation (DDES) modifies the character length scale to protect the RANS region from being prematurely switched into LES region. However, this modification is still incapable of completely preventing the occurrence of MSD problem^[16]. In addition, several variants of the DES model, like Delayed DES (DDES) and Improved DDES (IDDES) have been proposed with rather different characteristics, making model selection and interpretation of results challenging.

SST-DDES

SST-DDES is a hybrid Reynolds-Averaged Navier-Stokes (RANS) - Large Eddy Simulation (LES) method. It utilizes sub-grid scale model to handle the flow in the free shear flow area far away from the wall, and RANS's SST model is used to solve the flow in the boundary layer near the wall and other areas. This can guarantee the accuracy of LES solution, but also reduce the amount of calculation in the near-wall region of the boundary layer. The SST-DDES equations are briefly described at first. The governing equations of the SST DDES model are as follows:

$$\frac{\partial \rho k}{\partial t} + \frac{\partial(\rho u_j k)}{\partial x_j} = P_k - \rho k^{\frac{2}{3}}/l_{DDES} + \frac{\partial}{\partial x_j} \left[(\mu + \sigma_k \mu_t) + \frac{\partial k}{\partial x_j} \right] \quad (1)$$

$$\frac{\partial \rho \omega}{\partial t} + \frac{\partial(\rho u_j \omega)}{\partial x_j} = \alpha \frac{\rho}{\mu_t} P_k - \beta \rho \omega^2 + \frac{\partial}{\partial x_j} \left[(\mu + \sigma_\omega \mu_t) + \frac{\partial \omega}{\partial x_j} \right] + 2(1 - F_1) \rho \sigma_{\omega 2} \frac{\nabla k \cdot \nabla \omega}{\omega} \quad (2)$$

where F_1 and F_2 is a blending function which is defined as:

$$F_1 = \tanh(\arg_1^4) \quad (3)$$

$$\arg_1 = \min \left[\max \left(\frac{\sqrt{k}}{C_\mu \omega d_w}, \frac{500\nu}{d_w^2 \omega} \right), \frac{4\rho \sigma_{\omega 2} k}{CD_{k\omega} d_w^2} \right] \quad (4)$$

$$CD_{k\omega} = ax \left(2\rho \sigma_{\omega 2} \frac{\nabla k \cdot \nabla \omega}{\omega}, 10^{-10} \right) \quad (5)$$

$$F_2 = \tanh(\arg_2^2) \quad (6)$$

$$\arg_2 = \max \left(\frac{2\sqrt{k}}{C_\mu \omega d_w}, \frac{500\nu}{d_w^2 \omega} \right) \quad (7)$$

Here d_w is the distance to the nearest wall. The production term P_k in Equation (1) reads as follows:

$$P_k = \min(\mu_t S^2, 10 \cdot C_\mu \rho k \omega) \quad (8)$$

$$\mu_t = \rho \frac{a_1 \cdot k}{\max(a_1 \cdot \omega, F_2 \cdot S)} \quad (9)$$

The DDES length scale in Equation (1) reads as follows:

$$l_{DES} = \min\{l_{RANS}, C_{DES} h_{max}\} \quad (10)$$

$$l_{DDES} = l_{RANS} - f_d \max\{0, l_{RANS} - C_{DES} h_{max}\} \quad (11)$$

$$l_{RANS} = \frac{\sqrt{k}}{C_\mu \omega} \quad (12)$$

$$C_{DES} = C_{DES1} \cdot F_1 + C_{DES2} \cdot (1 - F_1) \quad (13)$$

Here h_{max} is the maximum edge length of the cell. Finally, the empiric blending function f_d in is computed with the use of the following relations:

$$f_d = 1 - \tanh[(C_{d1} r_d)^{C_{d2}}] \quad (14)$$

$$r_d = \frac{\nu + \nu_t}{\kappa^2 d_w^2 \sqrt{0.5 \cdot (S^2 + \Omega^2)}} \quad (15)$$

Here S is the magnitude of the strain rate tensor and Ω is the magnitude of vorticity tensor.

The model constants read as follows:

$$C_\mu = 0.09, \kappa = 0.41, a_1 = 0.31, C_{DES1} = 0.78, C_{DES2} = 0.6, C_{d1} = 20, C_{d2} = 3.$$

All the constants are computed by a blend from the corresponding constants of the $k-\varepsilon$ and $k-\omega$ model via $\alpha = \alpha_1 \cdot F_1 + \alpha_2 \cdot (1 - F_1)$ etc. :

$$\alpha_1 = 5/9, \beta_1 = 0.075, \sigma_{k1} = 0.85, \sigma_{\omega 1} = 0.5$$

$$\alpha_2 = 0.44, \beta_1 = 0.0828, \sigma_{k2} = 1, \sigma_{\omega 2} = 0.856$$

SST-IDDES

The simplified version of IDDES length scale is as follows:

$$l_{IDDES} = \tilde{f}_d d + \tilde{f}_d (1 - \tilde{f}_d) \psi C_{DES} \Delta \quad (16)$$

$$\tilde{f}_d = \max\{f_B, 1 - f_{dt} \Delta\} \quad (17)$$

The LES length-scale Δ is defined as:

$$\Delta = \min\{C_w \max[d_w, h_{max}], h_{max}\} \quad (18)$$

$$f_{dt} = 1 - \tanh[(C_{dt1} r_{dt})^{C_{dt2}}] \quad (19)$$

$$r_{dt} = \frac{v_t}{\kappa^2 d_w^2 \sqrt{0.5 \cdot (S^2 + \Omega^2)}} \quad (20)$$

$$f_B = \min\{2e^{-9\alpha^2}, 1.0\} \quad (21)$$

$$\alpha = 0.25 - d_w/h_{max} \quad (22)$$

In addition to the model constants in Equations (18) the following constants are introduced in the model:

$$C_w = 0.15, C_{dt1} = 20, C_{dt2} = 3$$

The full version of SST-IDDES model equations is shown in paper of Gritskevich^[17] (2012).

Computational Domain and Mesh

The sketch of the computational domain can be seen in Fig.1. The distance spans $20D$ in stream-wise direction, $10D$ in lateral and $3.14D$ in vertical direction, respectively. Structured grid generation is chosen in this case due to the simple geometry of the cylinder. As shown in Fig. 2, the mesh domain of $10D$ around the cylinder is generated with the O block grids. While the rest of mesh domain is generated with orthogonal hexahedral grids. The thickness of the first grid near the wall of the cylinder is set as $\Delta = 0.005D$, corresponding to approximately $y^+ = 1.5$. As indicated by Zhao^[18] (2016) and Travin^[19] (2000), grid is not convergent on DES and LES simulations, i.e., the finest grid didn't get the best agreement for all quantities compared with experiments. Finally, a mesh is generated with the total number of grid units being 1.4 million.

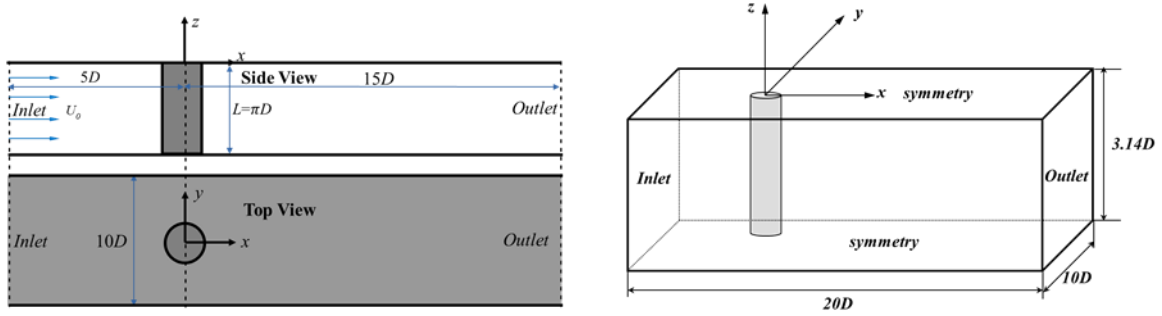


Fig. 1 Sketch of the computational domain

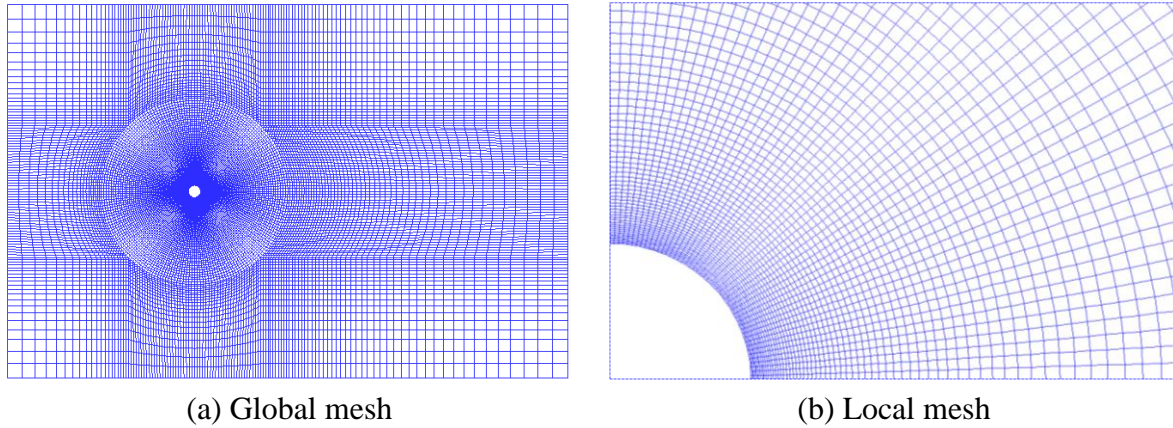


Fig. 2 Global and local mesh for circular cylinder. (a) Global mesh. (b) Local mesh

Boundary Conditions and Other Computational Details

According to the physics feature of the computation domain, the boundary is marked as the inlet, the outlet, the sides, the bottom and top. The surface of the cylinder is considered as a no-slip wall. At the inlet boundary, a uniform incoming flow with velocity equal to the free stream velocity $U_{\infty} = 0.039 \text{ m} \cdot \text{s}^{-1}$ is defined. At the outlet boundary, the pressure gradient is set equal to 0. The rest of the boundaries is defined as symmetry boundary for the reason of assuming that the height of the cylinder is infinite.

Discretization format

In this study, the governing equations are the three dimensional incompressible Navier-Stokes (N-S) equations. Since the SST-IDDES approaches are employed here. The governing equations are discretized using a finite volume method for solving the incompressible Navier-Stokes equations using solver naoe-FOAM-SJTU, with a newly implemented SST-IDDES turbulence model. The time discretization is done using second order implicit Euler scheme. A second order Gauss integration is used for spatial gradient calculations. The convection operator is discretized using a total variation diminishing (TVD) scheme.

The physics of the flow is considered as transient, incompressible and turbulent flow. Hence, we chose pimpleFoam as the solver for this case. pimpleFoam algorithm uses an inner PISO loop to get an initial solution, which is then under relaxed and corrected using an outer SIMPLE loop. This method enables unsteady simulations at Courant-Friedrichs-Lewy number (CFL) numbers larger than 1. In theory, very large CFL numbers could be maintained if a large number of SIMPLE correction loops were applied along with large under-relaxation factors. The equation of Courant-Friedrichs-Lewy (CFL) number is as follows:

$$\text{CFL} = \frac{\Delta t U_{\infty}}{\Delta x} \quad (23)$$

A too large time step generally smears the solution, regardless of the grid resolution. Hence time step is chosen in such a way that CFL is less than 1.

Results and Discussions

Figure3 demonstrates the time histories of the present lift (C_l) and drag (C_d) coefficients for a period of 80 vortex shedding cycles.

The drag non-dimensional forces coefficient is given by:

$$C_d(t) = \frac{2F_x(t)}{\rho H D U^2} \quad (24)$$

Where $F_x(t)$ is the force in the in-line direction; U is the velocity of flow; ρ is the density of water; D , in this case, is the projected length (or characteristic diameter).

The lift non-dimensional forces coefficient is given by:

$$C_l(t) = \frac{2F_y(t)}{\rho H D U^2} \quad (25)$$

Where $F_y(t)$ is the force in the transverse direction; U is the velocity of flow; ρ is the density of water; D , in this case, is the projected length (or characteristic diameter).

The Reynolds number is defined as:

$$Re = \frac{UD}{\nu} \quad (26)$$

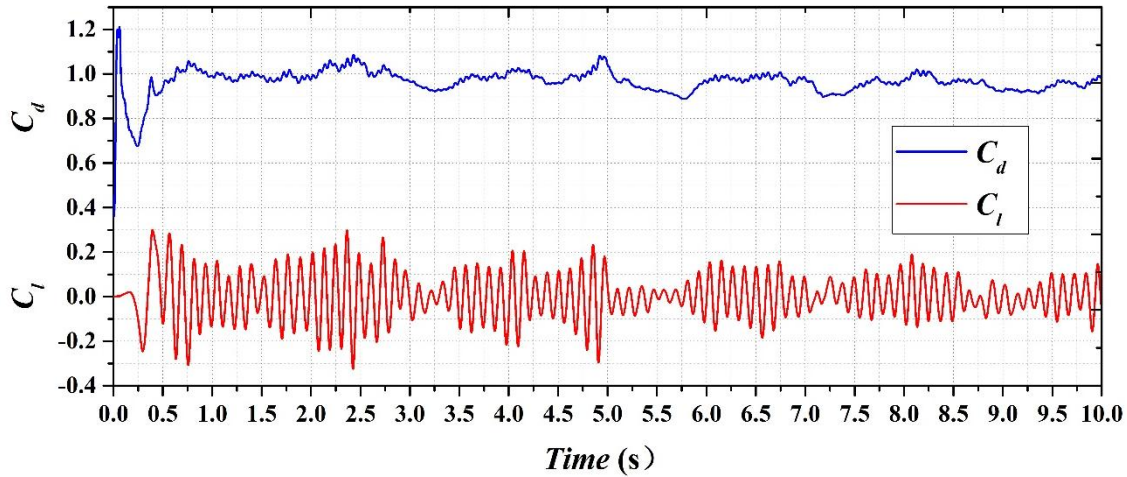


Fig. 3. Fluctuation of Coefficient of Drag(C_d) and Lift (C_l) Vs Time in second . Upper curves, C_d ; lower, C_l

Initially the fluctuation goes on increases but after some time period it stabilizes and fluctuates between the constant values. Further, the time-averaged drag coefficient is C_d calculated and listed in Table 1. The values obtained by the SST-DDES and SST-IDDES are in good agreement with the experimental data. The analysis of the vortex shedding frequency, or the Strouhal number (St), is obtained and given in Table 1. The values by the SST-DDES and SST-IDDES agree well with the experimental data. The present SST-IDDES and SST-DDES suggest a much longer recirculation bubble ($L_{rec}/D = 1.5$) than that indicated by experimental data of Lourenco and Shih^[3](1993).

Table 1. Overall flow parameters of the flow past a circular cylinder, $Re=3900$

Data Source	C_d	$-C_{pb}$	St	L_{rec}/D	θ
PIV Exp. of Lourenco and Shih ^[3] (1993)	0.99	0.88	0.215	1.33	89
DNS of Ma et al. ^[20] (2000)	0.84	-	0.220	1.59	-
SST-DES of Xu et al. ^[9] (2010)	1.08	-	0.220	0.98	-
DNS of Frederic, Tremblay ^[10] (2002)	1.03	0.93	0.220	1.30	85.7
LES of Frederic, Tremblay ^[10] (2002)	1.14	0.99	0.210	1.04	87.3
LES of Kravchenko ^[5] (2000)	1.04	0.94	0.210	1.35	88.0
Present SST-DDES	0.99	0.83	0.207	1.27	87.1
Present SST-IDDES	0.97	0.87	0.215	1.20	86.5

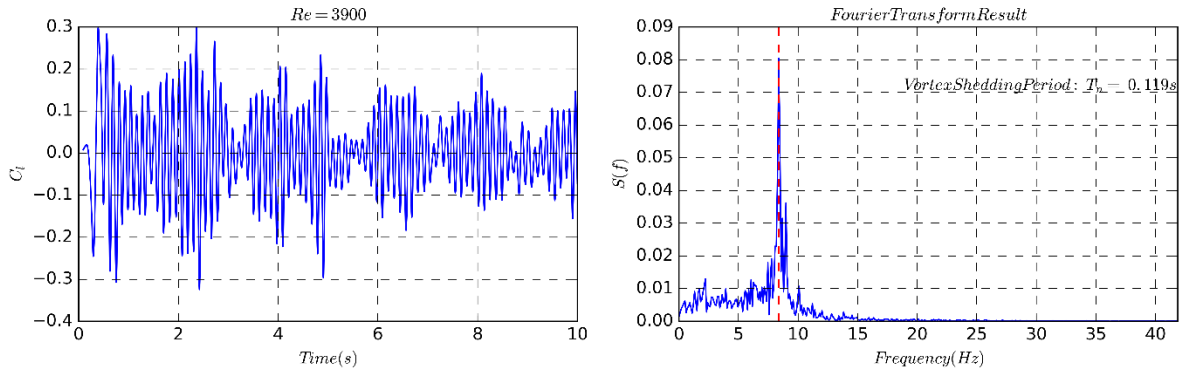


Fig. 4 Time history and Fourier transform result of Coefficient of Lift (C_l)

Figure 4 shows the time history and Fourier transform result of coefficient of lift (C_l). After the Fourier transform processing of data, the vortex shedding periods can be got with $T_n=0.119s$ (i.e. the vortex shedding frequency f_n is equal to $f_n=1/T_n=8.4$ Hz).

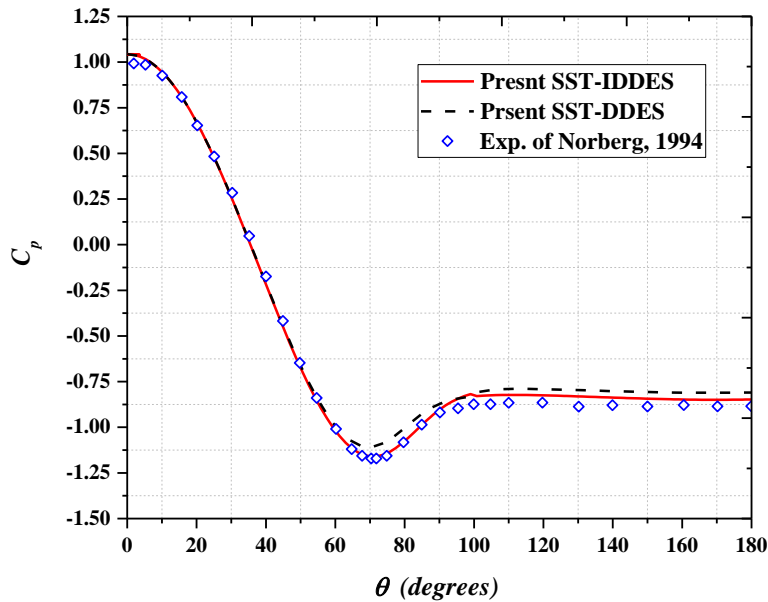


Fig.5 Time-averaged pressure coefficient on the surface of the cylinder with the experimental data and the SST-DDES results

Figure 5 Shows the time-averaged pressure distribution of SST-IDDES is reasonably below the SST-DDES result, and showing a good agreement with the experimental data^[2].

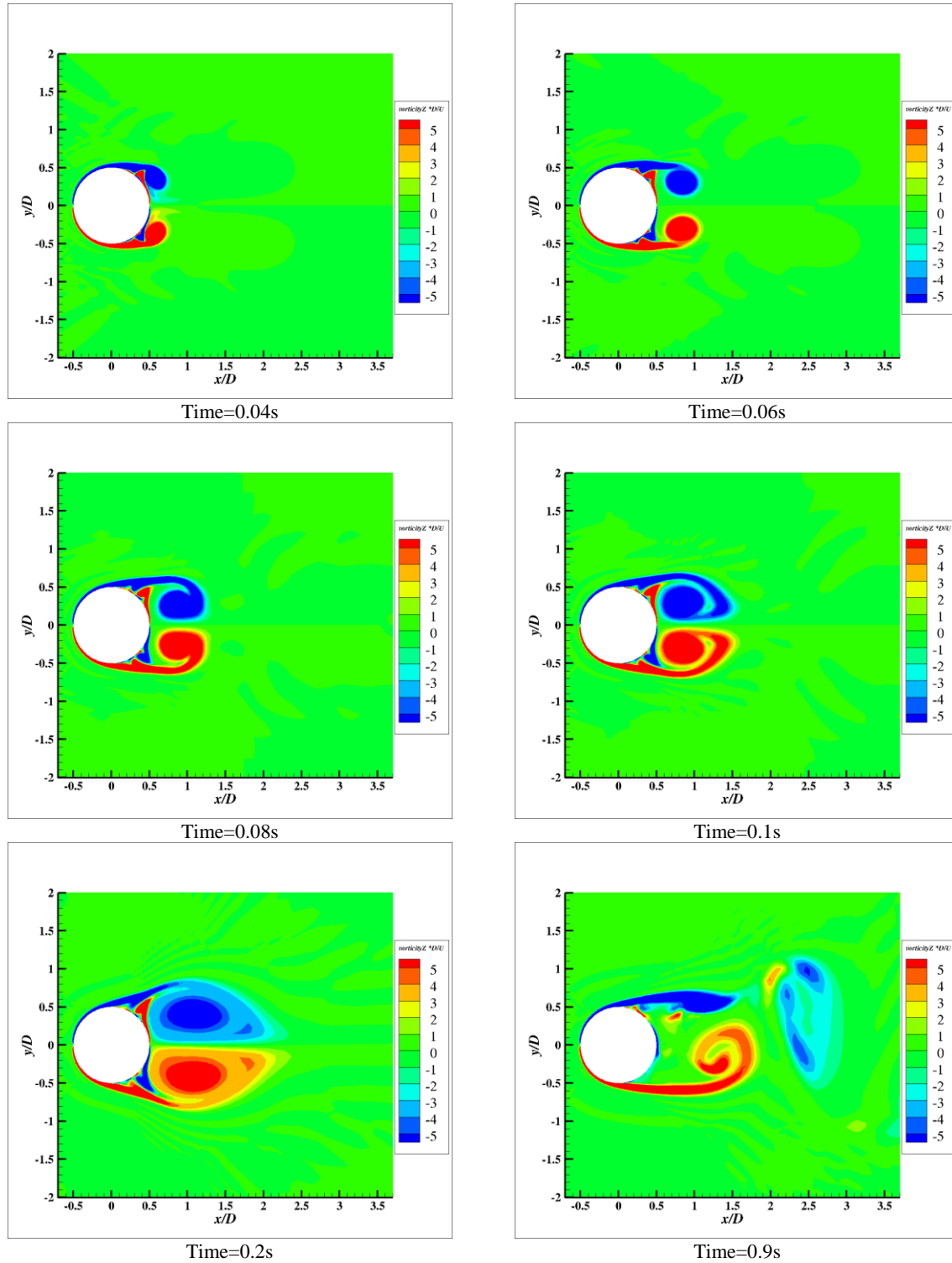


Fig. 6 The time evolution of instantaneous vorticity Z contour onto a horizontal plane ($z=0$)

Figure 6 shows the evolution of instantaneous vorticity Z at different time. The formation of the free shear layers from the cylinder surface took place identically for all runs. As the boundary layers from the upper and lower parts of the cylinder surface detach, free-shear layers are formed. This process is also displayed in Fig. 7. The separating shear layers behind the cylinder become unstable, and small-scale vortices can be clearly observed in the shear layers.

Figure 7 shows the iso-surface of the Q-criterion, flow past a circular cylinder with present SST-IDDES simulation. The slight difference vortex street pattern among each plane indicating the three dimensional turbulent structures.

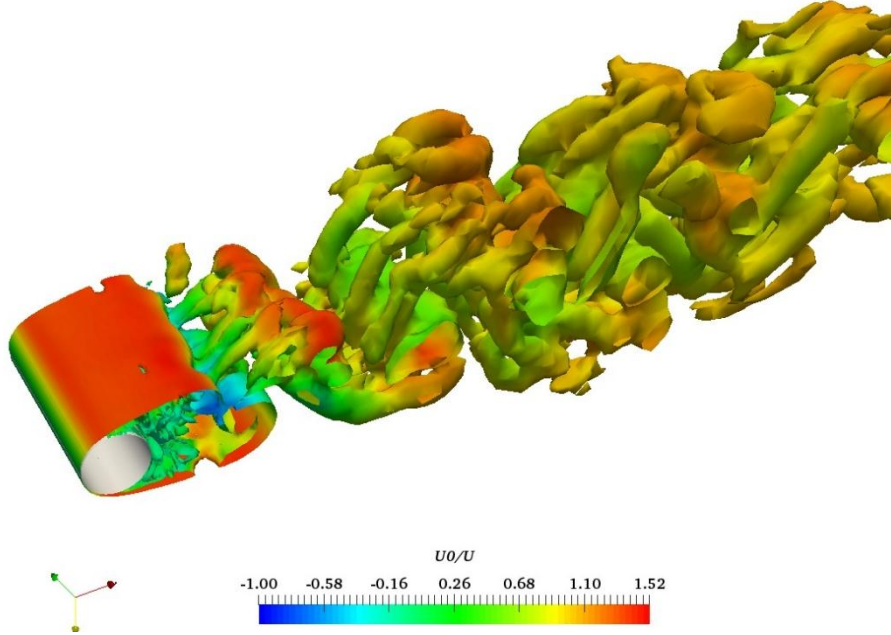


Fig. 7 Iso-surface of the Q-criterion, flow past a circular cylinder with present SST-IDDES simulation.

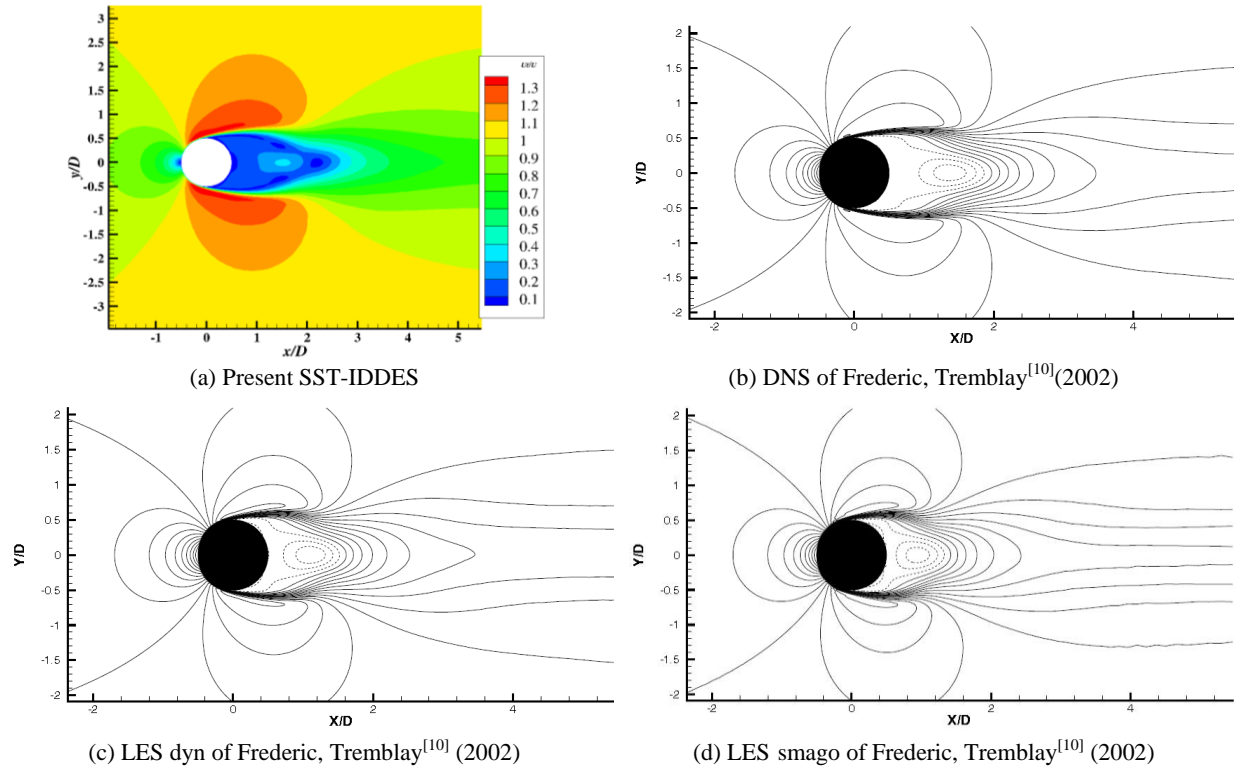


Fig. 8 Contours of mean velocity onto a horizontal plane ($z=0$)

Figure 8 shows contour and streamlines of mean velocity onto a horizontal plane. We again notice the shortening of the recirculation length obtained by our SST-IDDES computations as compared to the DNS result. We note once again the shorter recirculation bubble of the two LES computations by Frederic, Tremblay^[10].

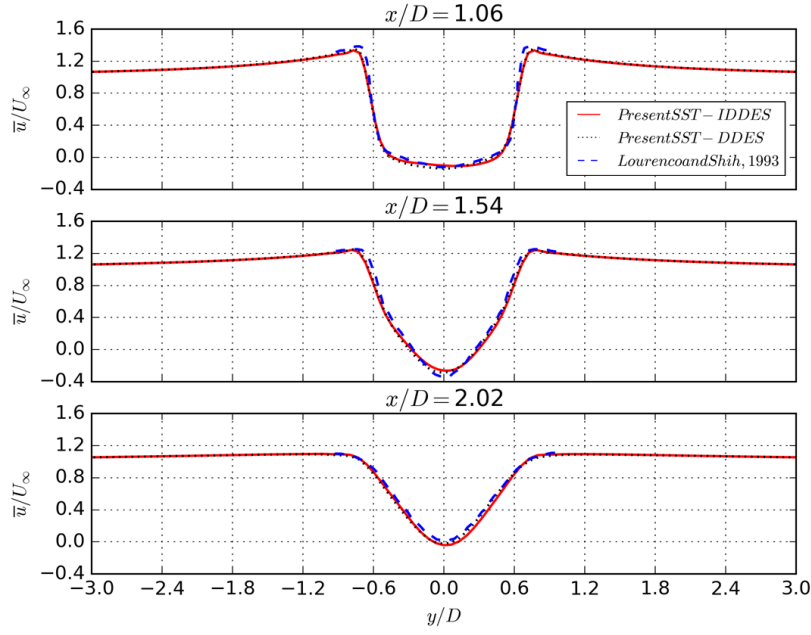


Fig. 9. Mean crossline velocities \bar{u}/U_∞ in different cross-stream profiles of near wake behind the cylinder

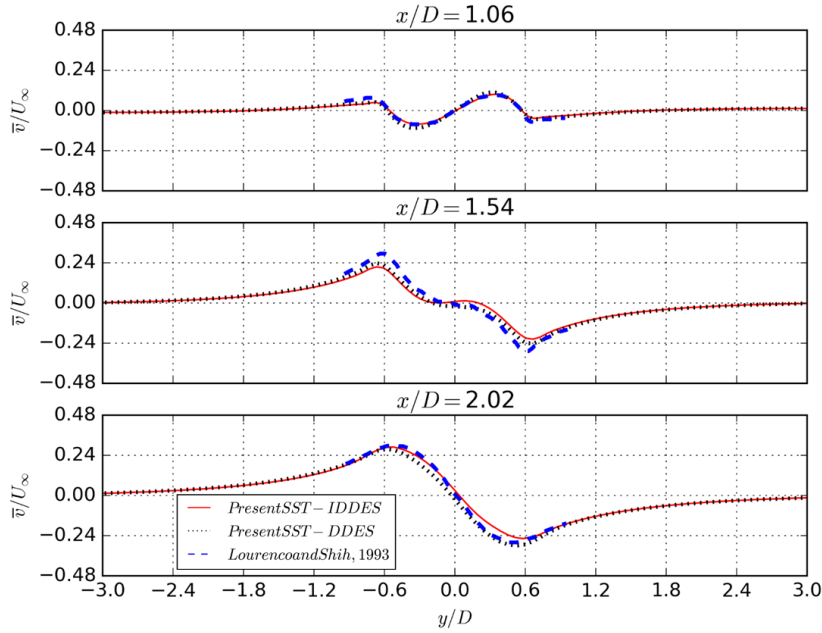


Fig. 10. Mean inline velocities \bar{v}/U_∞ in different cross-stream profiles of near wake behind the cylinder

Figure 9 shows mean velocities in different cross-stream profiles of downstream. At $x/D = 1.06$, the axial velocity profile exhibits a typical U shape that many others have also found in their LES and DNS studies. At $x/D = 1.54$ and $x/D = 2.02$, both predictions yield V shaped profiles. These cross-stream profiles of the mean axial velocity for both meshes are consistent with the observations made earlier with regard to the length of the recirculation bubble. It has been discussed by Kravchenko^[5] that the difference between a U shape and a V shape of the mean streamwise velocity profile in the near wake is mainly due to the shear layer dynamics

which affect this region. Figure 10 contains averaged over sufficient time inline velocities plots in series of cross-stream profiles along the downstream past the cylinder.

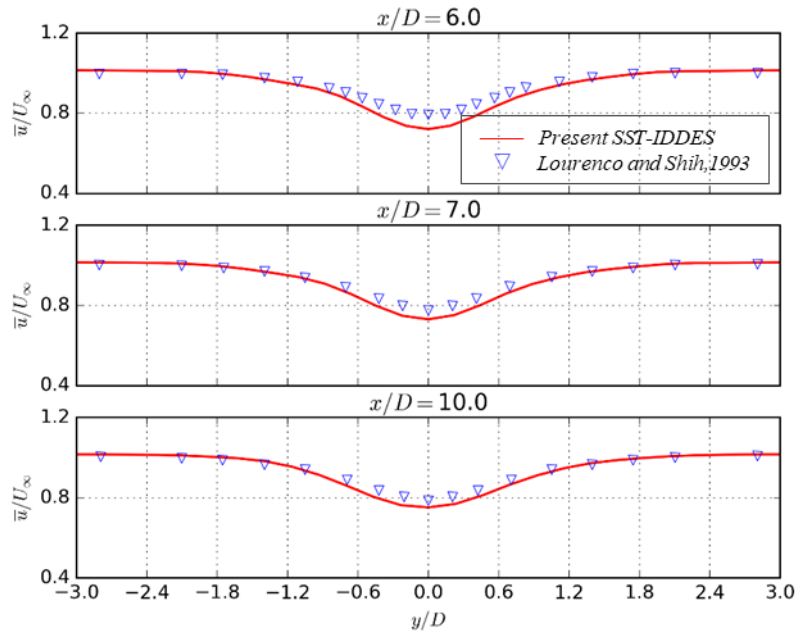


Fig. 11 Mean crossline velocities in different cross-stream profiles of near wake behind the cylinder

The normalized mean stream-wise velocity U/U_∞ in the far wake of the cylinder ($x/D \leq 10$) is plotted in Figure 11. Obviously, SST-IDDES didn't predicts the minimum value along the centerline of the cylinder. It could be due to the rather large grid spacing in the far wake of the cylinder.

Conclusion and Outlook

The separated turbulent flow around a circular cylinder is investigated using SST-IDDES. Some typical results, e.g., the mean pressure and drag coefficients, velocity Profiles and Strouhal number are calculated. A detailed validation of the SST-IDDES implementation for the turbulent flow around a circular cylinder has been conducted, with impressive levels of agreement achieved with experimental PIV data for both time-averaged and instantaneous quantities. In general, the prediction of SST-IDDES is satisfactory compared with the experiment data. Moreover, the hybrid RANS/LES models are able to simulate small turbulence structures and three dimensional effect even with limited computational meshes. The good agreement validated the CFD solver naoe-FOAM-SJTU is applicable and reliable to study such problems. Nevertheless, high Reynolds number (supercritical Re) flows are still challenging for the author to use present SST-IDDES models. Maybe more numerical try will need to be carried out in the near future.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Shanghai Key Laboratory of Marine Engineering (K2015-11), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical

Tank of Ministry of Industry and Information Technology of China(2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

- [1] Ong, L., and Wallace, J. (1996) The velocity field of the turbulent very near wake of a circular cylinder, *Experiments in Fluids* **20**(6), 441–453.
- [2] Norberg, C. (1994) An experimental investigation of the flow around a circular cylinder: influence of aspect ratio, *Journal of Fluid Mechanics* **258**(1), 287–316.
- [3] Lourenco L, S.C. (1994) Characteristics of the plane turbulent near wake of a circular cylinder, Technical report TF-62, CTR Annual Research Briefs, NASA Ames / Stanford University.
- [4] Parnaudeau, P., Carlier, J., Heitz, D., and Lamballais, E. (2008) Experimental and numerical studies of the flow over a circular cylinder at Reynolds number 3900, *Citation: Physics of Fluids* **20**(085101), 1-14.
- [5] Kravchenko, A.G., and Moin, P. (2000) Numerical studies of flow over a circular cylinder at $Re_D=3900$, *Physics of Fluids* **12**(2), 403–417.
- [6] Zhao, R., Liu, J., and Yan, C. (2012) Detailed Investigation of Detached-Eddy Simulation for the Flow Past a Circular Cylinder at $Re=3900$, Springer, Berlin, Heidelberg, 401–412.
- [7] DONG, S., KARNIADAKIS, G.E., EKMEKCI, A., and ROCKWELL, D. (2006) A combined direct numerical simulation–particle image velocimetry study of the turbulent near wake, *Journal of Fluid Mechanics* **569**, 185–207.
- [8] Mockett, C., Perrin, R., Reimann, T., Braza, M., and Thiele, F. (2010) Analysis of Detached-Eddy Simulation for the Flow Around a Circular Cylinder with Reference to PIV Data, *Flow, Turbulence and Combustion* **85**(2), 167–180.
- [9] XU, C., CHEN, L., and LU, X. (2007) Large-Eddy and Detached-Eddy Simulations of the separated flow around a circular cylinder, *Journal of Hydrodynamics, Ser B* **19**(5), 559–563.
- [10] Frederic, T. (2002) Direct and large-eddy simulation of flow around a circular cylinder at subcritical Reynolds numbers, Fakultät für Maschinenwesen, Thesis, 1-146.
- [11] Pereira, F.S., Vaz, G., Eça, L., and Girimaji, S.S. (2018) Simulation of the flow around a circular cylinder at $Re=3900$ with Partially-Averaged Navier-Stokes equations, *International Journal of Heat and Fluid Flow* **69**, 234–246.
- [12] WANG, J. song (2010) Flow around a circular cylinder using a finite-volume TVD scheme based on a vector transformation approach, *Journal of Hydrodynamics* **22**(2), 221–228.
- [13] Krishnan, V., Squires, K.D., and Forsythe, J.R. (2006) Prediction of the Flow Around a Circular Cylinder at High Reynolds Number, *44th AIAA Aerospace Sciences Meeting and Exhibit, 9-12 January 2006*, Reno, Nevada, United States, 10765–10774.
- [14] Spalart, P.. (2000) Strategies for turbulence modelling and simulations, *International Journal of Heat and Fluid Flow* **21**(3), 252–263.
- [15] Menter, F.R., Kuntz, M., and Langtry, R. (2003) Ten years of industrial experience with the SST turbulence model, *Turbulence, Heat and Mass Transfer* **4**(1), 625–632.
- [16] Menter, F., and Egorov, Y. (2005) A Scale Adaptive Simulation Model using Two-Equation Models, *43rd AIAA Aerospace Sciences Meeting and Exhibit*, Reston, Virginia, American Institute of Aeronautics and Astronautics, 1095–1107.
- [17] Gritskevich, M.S., Garbaruk, A. V., Schütze, J., and Menter, F.R. (2012) Development of DDES and IDDES formulations for the $k-\omega$ shear stress transport model, *Flow, Turbulence and Combustion* **88**(3), 431–449.
- [18] Zhao, WW., Wan, DC, and Sun, R. (2016) Detached-Eddy Simulation of Flows over a Circular Cylinder at High Reynolds Number, *Proceedings of the Twenty-Sixth (2016) International Ocean and Polar Engineering Conference* **96**(10–11), 1528–1536.
- [19] Travin, A., Shur, M., Strelets, M., and Spalart, P. (2000) Detached-Eddy Simulations Past a Circular Cylinder, *Flow, Turbulence and Combustion* **63**(1/4), 293–313.
- [20] MA, X., KARAMANOS, G.-S., and KARNIADAKIS, G.E. (2000) Dynamics and low-dimensionality of a turbulent near wake, *Journal of Fluid Mechanics* **410**, S0022112099007934.

Numerical Simulation of Ship Bow Wave Breaking using DES and RANS

*Zhen Ren, Jianhua Wang and †Decheng Wan

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering,
Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration,
Shanghai 200240, China

*Presenting author: renzhen90@163.com

†Corresponding author: dcwan@sjtu.edu.cn

Abstract

The phenomena of wave breaking, known as white water, has attracted many researchers since its complex mechanism and effects on the performance of ship. The CFD solver naoe-FOAM-SJTU, is used to investigate the wave breaking phenomena of the bow wave of KCS model without sinkage and trim. In the present work, the DES and RANS turbulence model are adopted to simulate the bow wave breaking of KCS in different advance speeds, i.e. $Fr=0.26$, 0.30 , 0.35 . In the simulations, volume of fluid (VOF) is employed to capture the free surface. For the $Fr=0.26$ case, the predicted resistance and wave patterns via both turbulence models are in good agreement with the available experiment data. For the $Fr = 0.35$ case, the process of overturning and breaking of bow wave were obtained by the both turbulence models, and the scars are more visible by the DES approach. The detailed flow visualizations obtained through both turbulence models, such as wake profiles, vorticity and wave patterns, are presented to illustrate the hydrodynamic performance of high speed surface ship. The present simulations give a better understanding of the ability of both turbulence models to predict the ship bow waves breaking.

Keywords: wave breaking; KCS; free surface; OpenFOAM, RANS, DES

Introduction

Despite of the high accuracy of the resistance prediction, it is still challenging to accurately resolve the breaking wave phenomenon, which has long been recognized. Extensive experiments have been performed to try to give the physical understanding of the breaking wave mechanism and provide experimental data for CFD validation. Dong et al.^[1] conducted experimental study using particle-image-velocimetry (PIV) measurements to analyze the velocity component and vorticity distribution at $Fr=0.28$ and $Fr=0.45$. The vorticity and its associated lateral flow field are analyzed. Duncan^[2] studied the breaking and non-breaking wave resistance of a 2-D hydrofoil via the experiments in which they found the drag associated with breaking was more than 3 times the drag theoretically obtained with non-breaking waves. Kayo and Takekuma^[3] investigated bow wave breaking phenomenon around full ship models by velocity field measurements and by a flow visualization technique. They found that there was a shear flow on the free surface. Roth et al.^[4] utilized PIV to measure the flow structure and turbulence within the bow wave of DDG-51 model 5422. In the measurements, they also found that the negative vorticity originated at the toe of the wave while the positive vorticity was generated on the crest of the wave. In addition, they

discovered that the great energy losses were experienced at the toe. Longo and Stern^[5] performed mean velocity measurements using a five-hole Pitot and wave elevation measurements using capacitance wires and point gauges for the static drift condition showing the presence of a bow wave breaking induced vortex on the windward side of the model. Olivieri et al.^[6] studied the bow wave breaking of model DTMB 5415 and they analyzes the scars and vortices induced by bow and shoulder wave breaking. In that study, the large extents of plunging bow was observed at $Fr=0.35$ and spilling shoulder wave breaking was investigated.

Despite the extensive study through experiment, numerical simulation has also been adopted to predict and analyze the ship wave breaking phenomena. In the numerical simulation of Wilson et al.^[7], the unsteady single-phase level set method was adopted to resolve and investigate bow wave breaking around a surface combatant advancing in calm water, including induced vortices and free surface scars. The velocity component and vorticity distribution were in good agreement with the experiment data. Moraga et al.^[8] proposed a sub-grid model for air entrainment in the bow waves breaking and applied for the simulation of naval surface ship DTMB 5415 and Athena. The model compared favorably with data at laboratory scale and also presented the right trends at full-scale. Marrone et al.^[9] used a 2D+t SPH model to analyze the breaking wave pattern of the vessel DTMB 5365, and the overturning and breaking of bow wave were captured clearly. Marrone et al.^[10] developed a 3D SPH solver to simulate the bow wave breaking of Alliance vessels. The results achieved by the solver were compared with the experimental measurements and numerical results from RANS simulation in which the level set method was applied to resolve the free surface. Noblesse et al.^[11] reviewed the recent results about the overturning and breaking bow wave regimes, and the boundary that divides these two basic flow regimes. Questions and conjectures about the energy of breaking ship bow waves, and free-surface effects on flow circulation, are also noted.

In present work, high resolution Volume of Fluid (VOF) method is used to accurately resolve the large deformation of free surface. The main framework of this paper goes as following. The first part is the numerical methods, where DES and RANS turbulence models are presented. The second part is the geometry model and grid generation. Then comes the simulation part, where wave breaking simulations are present at different Froude numbers. In this part, extensively comparisons are performed between the results obtained via different turbulence models including ship resistance, wave patterns and wake fields at longitudinal slices. Finally, a conclusion of this paper is drawn.

Numerical methods

Governing equations

The in-house CFD solver naoe-FOAM-SJTU^[12-14], developed on open source platform OpenFOAM, is applied in this study and VOF method is used to capture free surface around the complex geometry models. The governing equations are written as a mass conservation equation and a momentum conservation equation^[15]:

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0 \quad (1)$$

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial \bar{u}_j \bar{u}_i}{\partial x_j} = -\frac{\partial \bar{P}}{\partial x_i} + \frac{\partial}{\partial t} \left[\nu \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) \right] - \frac{\partial \tau_{ij}}{\partial x_j} \quad (2)$$

Where, ν is the molecular viscosity, τ_{ij} is the Reynolds stress tensor or subgrid-scale stress tensor. According Boussinesq hypothesis, τ_{ij} can be expressed as

$$\tau_{ij} = \frac{2}{3} \delta_{ij} k - \nu_t \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) \quad (3)$$

Turbulence Model

The SST model was proposed by Menter^[16] and is a two-equation turbulence model. SST combines the k - ϵ and k - ω models to treat the flow in the boundary layer region of the near wall with k - ω and the flow in the free shear flow region with k - ϵ . This not only preserves the ability of the k - ω model to handle the boundary conditions near the wall surface, but also avoids the insensitivity of the k - ω model to the inlet parameters in the free shear flow region. The DES model assumes the eddy viscosity ν_t is a function of turbulent kinetic energy k and specific turbulence dissipation rate ω , and strain invariant S .

$$\nu_t = \frac{a_1 k}{\max(a_1 \omega, SF_2)} \quad (4)$$

In which, the transport equations^[17] for the turbulent kinetic energy k and the specific turbulent dissipation rate ω are denoted as:

$$\frac{\partial k}{\partial t} + \frac{\partial(u_j k)}{\partial x_j} = \tilde{G} - \frac{k^{3/2}}{l_{DDES}} + \frac{\partial}{\partial x_j} \left[(\nu + \alpha_k \nu_t) \frac{\partial k}{\partial x_j} \right] \quad (5)$$

$$\frac{\partial \omega}{\partial t} + \frac{\partial(u_j \omega)}{\partial x_j} = \gamma S^2 - \beta \omega^2 + \frac{\partial}{\partial x_j} \left[(\nu + \alpha_\omega \nu_t) \frac{\partial \omega}{\partial x_j} \right] + (1 - F_1) CD_{k\omega} \quad (6)$$

The l_{DDES} in k -equation is mixing length scale which control the switch of solution between RANS and LES.

VOF method and surface tension

For the wave breaking simulations, the free surface capture method plays an important role in the accuracy of predicted results. In the present work, VOF method with bounded compression technique^[18] is applied to capture free surface and the transport equation is expressed as:

$$\frac{\partial \alpha}{\partial t} + \nabla \cdot [(\mathbf{U} - \mathbf{U}_s) \alpha] + \nabla \cdot [\mathbf{U}_r (1 - \alpha) \alpha] = 0 \quad (7)$$

Where α volume of fraction, 0 and 1 represent that the cell is filled with air and water respectively and $0 < \alpha < 1$ stands for the interface between two-phase fluids. \mathbf{U}_r in Eqn. (7) is the velocity field used to compress the interface and it only takes effect on the free surface due to the term $(1 - \alpha) \alpha$.

According to the literature concerning wave breaking, small scale wave breaking is strongly influenced by surface tension. The role played by the surface tension is quite different for breaking and non-breaking waves since the surface tension pressure jump depends on the magnitude of the radius of curvature of the free surface. In order to reappear the wave patterns of the experiment, the surface tension is taken account in the present simulation and the surface tension is expressed as:

$$f_\sigma = \sigma \kappa \nabla \alpha \quad (8)$$

Where σ stands for the surface tension, κ is the curvature of free surface and it is defined as:

$$\kappa = -\nabla \cdot \mathbf{n} = -\frac{\sum_f \mathbf{S}_f \cdot \mathbf{n}_f}{V_i} \quad (9)$$

V_i represents the volume of cell i , $\sum_f S_f$ stands for the sum of value on each face of cell.

Geometry model and grid generation

Geometry model

As a full-formed hull, the wave breaking of KRISO Container Ship (KCS) has aroused significant interest of researchers. The geometry model that is the 6.0702 m replica with rudder in numerical simulations is shown in Fig. 1, and its principle parameters are listed in Table 1. Extensive experiments have been conducted for the ship model under various Froude numbers except for Fr above 0.30. But the phenomena of wave breaking of KCS will be observed clearly at Fr=0.35 according to the previous work in which the study of the effects of different speeds on the wave breaking are carried out. In the present work, the wave breaking simulations of KCS at Fr= 0.35 are conducted. In the numerical simulation, the model is fixed without the sinkage and trim.



Fig. 1 Geometry model of KCS (Case 2.10)

Table 1 Principle dimensions of KCS

Main particulars		Full scale	Model scale
Length between perpendiculars	L_{pp} (m)	230	6.0702
Maximum beam of waterline	B_{WL} (m)	32.2	0.8498
Draft	T (m)	10.8	0.2850
Displacement volume	Δ (m ³)	52030	0.9565
Wetted surface area (with rudder)	S_0 (m ²)	9645	6.7182

Grid Generation

Due to the high computational costs and fixed ship model condition, only half of the computational domain is adopted for the numerical simulations. Fig. 2 shows the computational domain and the boundary conditions.

All-hexahedral unstructured grid adopted in the present simulations is generated by the software, *Hexpress*. The grid number in x , y , z direction is 100×30×45, respectively. To better resolve the bow wave breaking and free surface wave pattern, several blocks are adopted to refine the regions around the hull, bow and local free surface, as shown in Fig. 3. Block 1 is the region that wraps the hull surface. The 18.3 million grid scheme is obtained via the refinement in several blocks, as listed in Table 2. The scale of the size of the highest-level refinement region to L_{pp} is about 1.56e-3. Fig. 4 presents the global and local profile of grid distribution.

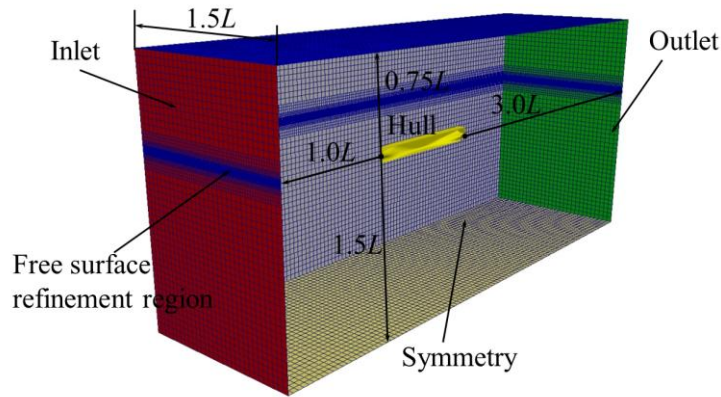


Fig. 2 Domain and boundary conditions.

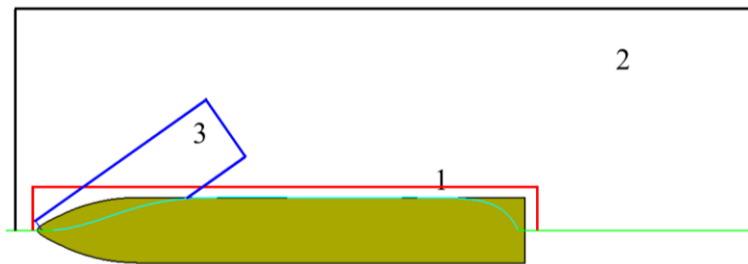


Fig. 3 Refinement regions.

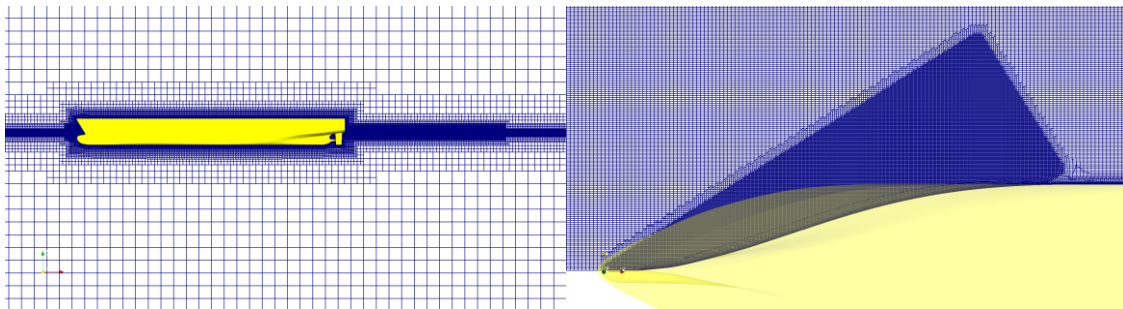


Fig. 4 Profile of grid distribution, global (left), local grid of bow (right)

Table 2 Grid refinement strategy

Total	No.	Refinement level (x, y, z)	Ratio($/L_{pp}$) ($\times 10^{-3}$)
	1	$3 \times 3 \times 3$	$6.25 \times 6.25 \times 6.25$
18.3Million	2	$4 \times 4 \times 5$	$3.125 \times 3.125 \times 1.5625$
	3	$5 \times 5 \times 5$	$1.5625 \times 1.5625 \times 1.5625$

Analysis of simulation results

Validation of Numerical Scheme

The focus of the present study is on the resolution of bow wave breaking of KCS via DES and RANS at different speeds, i.e. $Fr = 0.26, 0.30, 0.35$. The sinkage and trim of the hull is not taken into consideration in order to simplify the numerical simulation. The condition at $Fr = 0.26$ is selected to validate the prediction accuracy of the current numerical scheme.

Table 3 Comparison of experimental and predicted resistance

Parameters	EFD	CFD	
		RANS	DES
$U(\text{m/s})$	2.017	2.017	2.017
Sinkage ($/L_{pp}$)	-2.074e-3	None	None
Trim (deg)	-0.1646	None	None
Wetted surface area with rudder	6.6978	6.7449	6.7449
$C_t(\times 10^3)$	3.835	3.662	3.644
Error		-4.51%	-4.98%

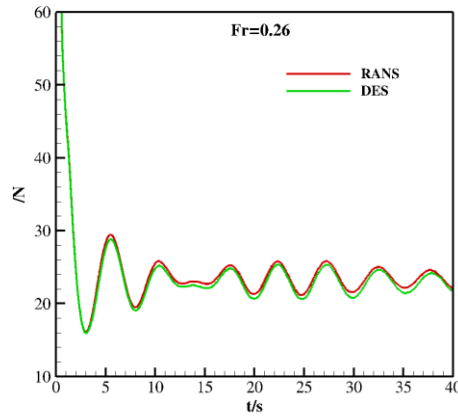


Fig. 5 Comparison of resistance obtain at $Fr=0.26$ via DES and RANS

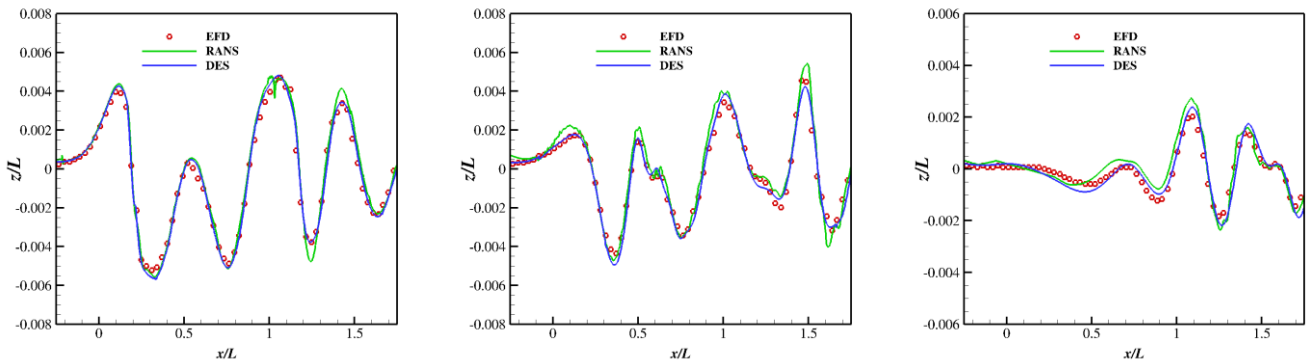


Fig. 6 Comparison of free-surface between experiments (circles) and computational results (green: RANS; blue: DES) at different cutting planes.

Fig. 5 shows the time history of total resistance via both turbulence models. As we can see, the convergence trends of total resistance via both numerical schemes are consistent. Table 3 shows the comparison between the predicted resistance and the experiment data. The errors via RANS and DES turbulence model are about -4.51% and -4.98%, respectively. Since the ignorance of the hull posture, the errors of resistance are acceptable. **Fig. 6** shows the wave height of the three profiles ($y/L = 0.0741, 0.1509, 0.4224$), obtained by experiment and both numerical simulations schemes. From the near field to the far field, the calculated free surface via both numerical schemes are consistent with the experimental measurements. Compared with the results via RANS schemes, the results via DES scheme is slightly better agreement with the experiment data. The results show that the numerical scheme that DES turbulence

model may be more suitable to simulate for capturing the free surface. On the other hand, the above results prove that the numerical scheme in the present work are reliable and robust.

Simulation results at $Fr=0.35$

According to the simulation results, the bow wave breaking is observed at $Fr=0.35$, so the numerical results at $Fr=0.30$ is not shown here. In the simulation of bow wave breaking of KCS at high speed, the hull also is fixed without trim and sinkage.

The time history of resistance at $Fr=0.35$ via both numerical schemes are shown in Fig. 7. The convergence trends of both lines are consistent and almost same basically. The total resistance coefficients achieved via both numerical schemes are compared in Table 4. The relative error of both resistances is 0.354% demonstrates the results via the present numerical schemes are reliable and robust.

Table 4 Comparison of resistance via DES and RANS

Parameter	$U(m/s)$	Sinkage (/Lpp)	Trim (deg)	$C_t(\times 10^3)$	
				RANS	DES
Value	2.701	None	None	5.084	5.066

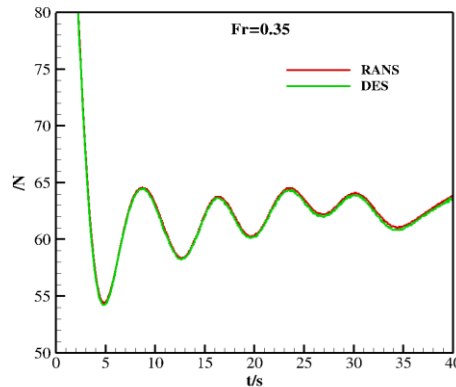


Fig. 7 Time history of total resistance via DES and RANS

The wave patterns of free surface via both turbulence models are presented in Fig. 8. As depicted in Fig. 8(a), the global profiles are same basically except bow wave. The result obtained via RANS approach is smoother than that obtained via DES approach. There are more stripes near the hull in the results achieved by the DES turbulence model. Near the bow, four scars are observed by DES method while there is only scar achieved by RANS method, as shown in Fig. 8 (b).

In the both numerical simulation, the marked difference is observed in the evolution of bow wave. Fig. 9 shows the difference between both simulation results more clearly. In the results of DES scheme, the nearest scar that is formed by the first overturning of bow wave starts at $x/L=0.09$ and disappears at $x/L=0.22$. When the bow wave plunges into the below free surface secondly, the second scar is formed at $x/L=0.16$ and it disappears at $x/L=0.28$. The third scar is also observed clearly between $x/L=0.21$ and $x/L=0.30$. The farthest scar is not visible apparently and it starts at $x/L=0.25$ and vanishes at $x/L=0.34$. Through the RANS scheme, the only one scar that starts at $x/L=0.09$ and vanishes at $x/L=0.19$ is visible clearly.

While the abrupt change and breaking of bow wave is very obvious near $x/L=0.21$ and the bow wave captured in the DES scheme seems to have better continuity and smoothness.

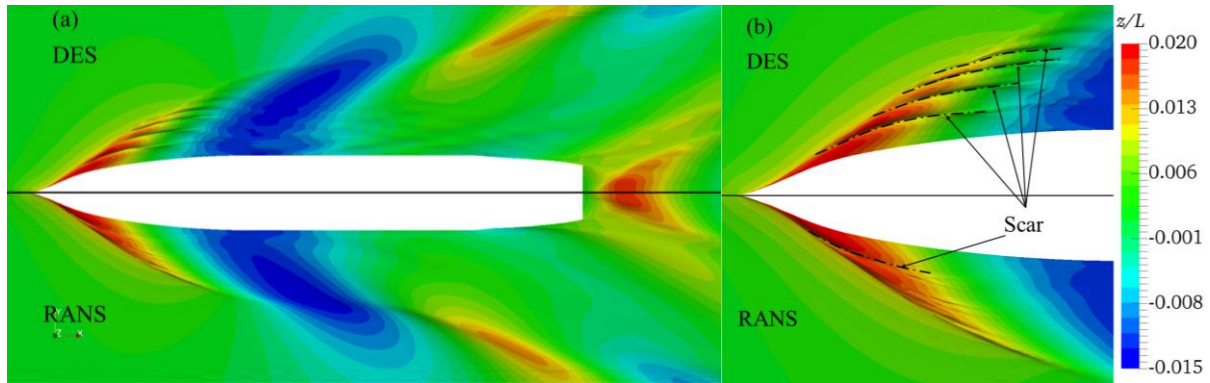


Fig. 8 Comparison of free surface via RANS (below) and DES (top), a: global profile, b: local profile

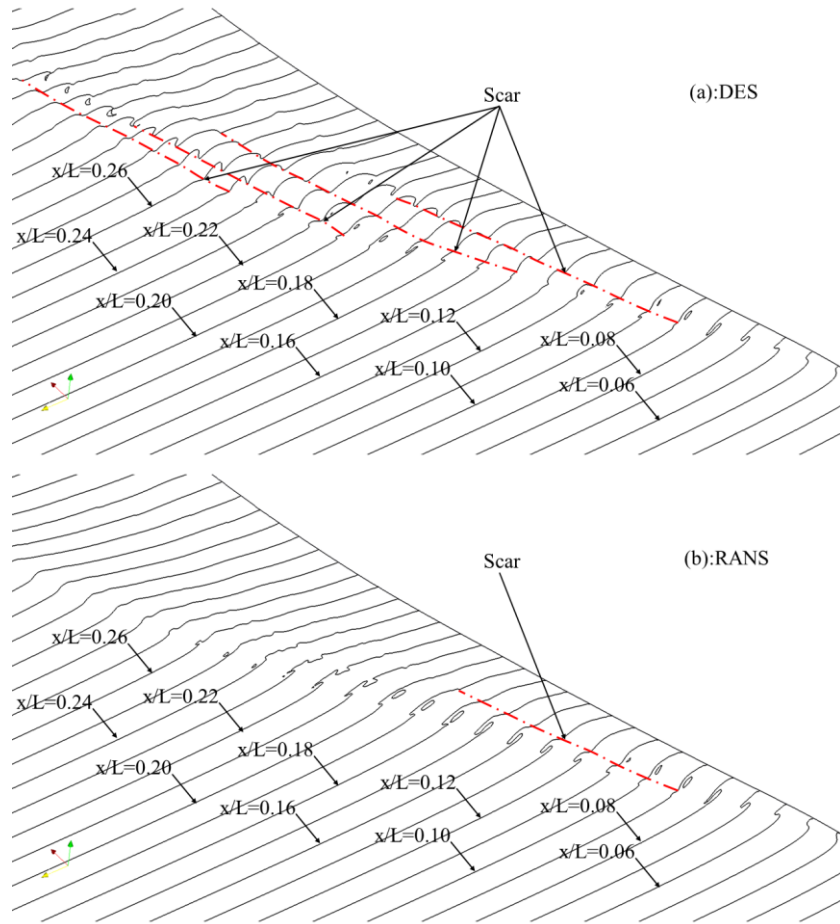


Fig. 9 Scar in bow wave breaking (a: DES, b: RANS)

The general view of the vertical structure at $Fr=0.35$, is given in Fig. 10 which presents an isosurface of the dimensionless value $Q=10$ ($Q = (S^2 + \Omega^2)/2$, S and Ω being the symmetric and antisymmetric component of velocity) colored by the velocity. With the DES approach, three vortices near the bow are visible clearly and maintained until the middle of the hull. Except the three main vortices, some fragmented vortices occur near the bow wave. The main vortices vanish near the middle of the hull. On the other hand, some larger vortices also

appear near the middle of the hull. Through the RANS approach, some small intermittent vortices are yielded and these vortices vanish rapidly in the wake.

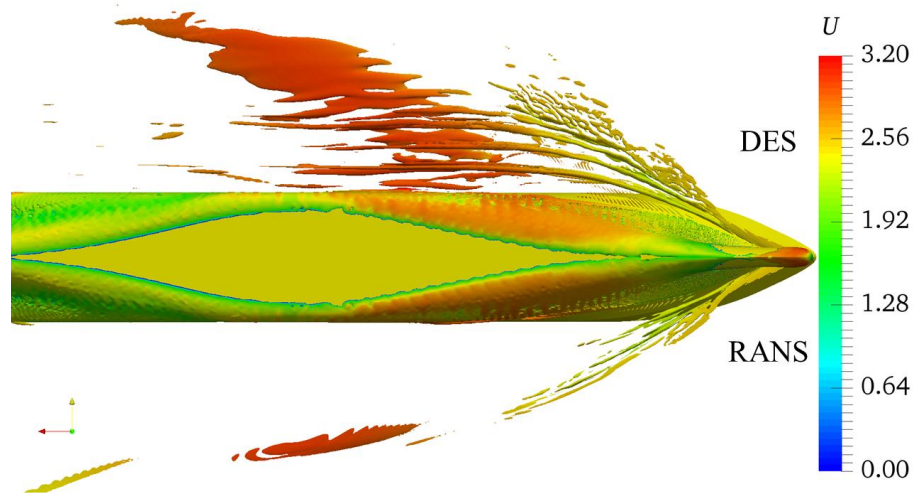


Fig. 10 Vortical structure (top: DES, below: RANS)

In order to compare the evolution of vortices more clearly, the axial vorticity distribution at different cutting planes are presented in Fig. 11. At $x/L=0.05$, the axial vorticity distribution obtained by both numerical schemes are similar. At the tip of the initial plunger, the axial vorticity is positive while the negative vorticity occurs at the high curvature region of the overturning bow wave. At $x/L=0.07$, the initial plunger generated due to the gravitational and inertial forces develops outboard and is going to reconnect with the free surface below. Although the axial vorticity distribution of both models are similar, the initial plunger yielded by DES model is thicker than that yielded by RANS approach. In addition, the initial plunger obtained by RANS scheme has shown the indication of breaking. Axial vorticity distribution is significantly different at $x/L=0.14$. The scar yielded by DES approach is clearer than that captured by RANS scheme. In the simulation of DES model, the shape of plungers is not observed while the second plunger is visible clearly in the prediction of RANS approach. In addition, the negative vorticity obtained by DES approach is much larger than that achieved by RANS scheme. In the simulation of DES model, the positive vorticity is much larger than the other scheme and concentrated on the free surface. A counter-rotating vortex pair is generated near the scar and the vortex pair pumps fluid outboard.

Though the comparison of vorticity, the vorticity yielded by both numerical schemes at $x/L=0.14$ differs greatly. In order to analyze the mechanism of bow wave, the wake field at $x/L=0.14$ is presented in Fig. 12. In the axial direction, the velocity distributions from both numerical schemes are similar. The variation of axial velocity is concentrated near the bow wave. The lowest axial velocity occurs near the wave crest and it increases with increasing distance from the hull. The obvious difference mainly due to the wave pattern is concentrated near the second plunger. Similar to the axial velocity distribution, the variation of lateral velocity is also concentrated near the bow wave. The highest lateral velocity occurs near the free surface. And the lateral velocity decreases with increasing distance from the model. The vertical velocity shows the obvious difference. The significant variation occurs near the free surface. Positive and negative vertical velocity appear alternately in the simulation of DES model leads to the counter-rotating flow so that the counter-rotating vortex pair is generated near the scar. The result obtained by the RANS approach also has this trend but it is not very clear.

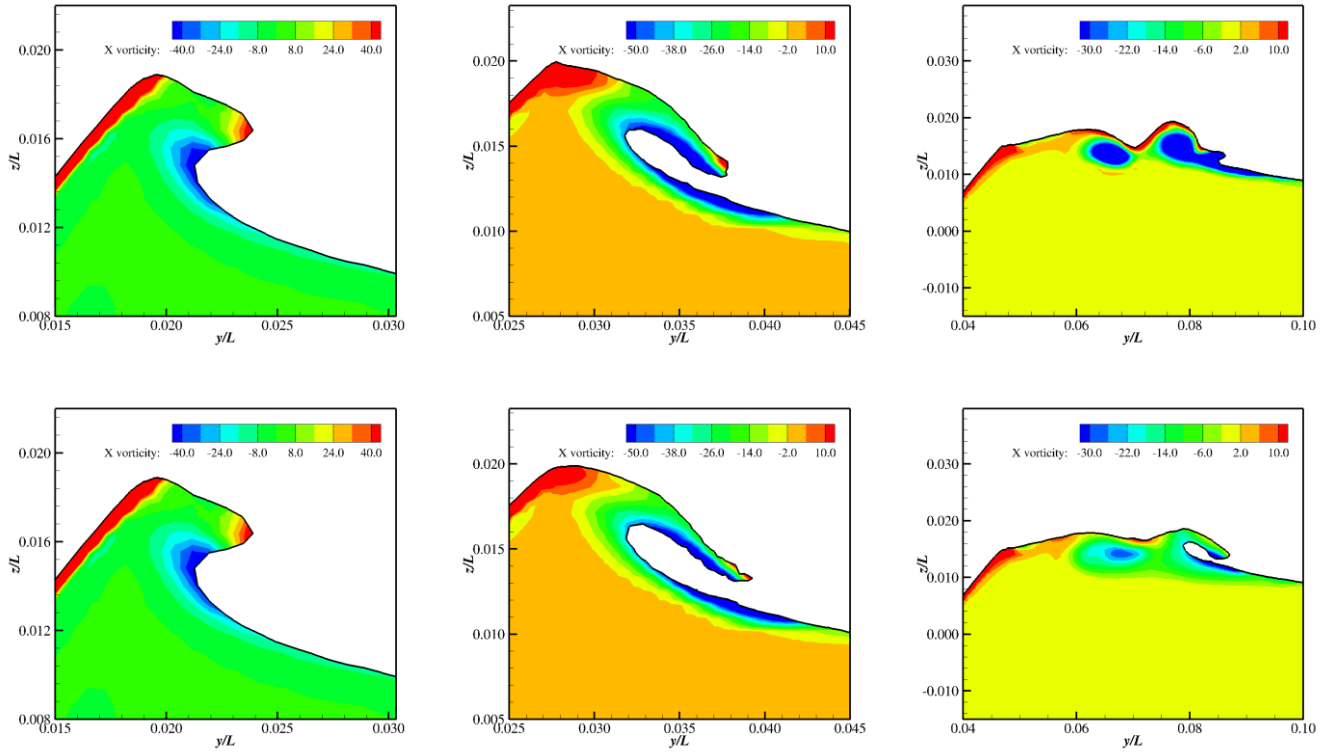


Fig. 11 Axial vorticity distribution at $x/L=0.05$ (left), 0.07 (middle) and 0.14 (right) (top: DES, below: RANS)

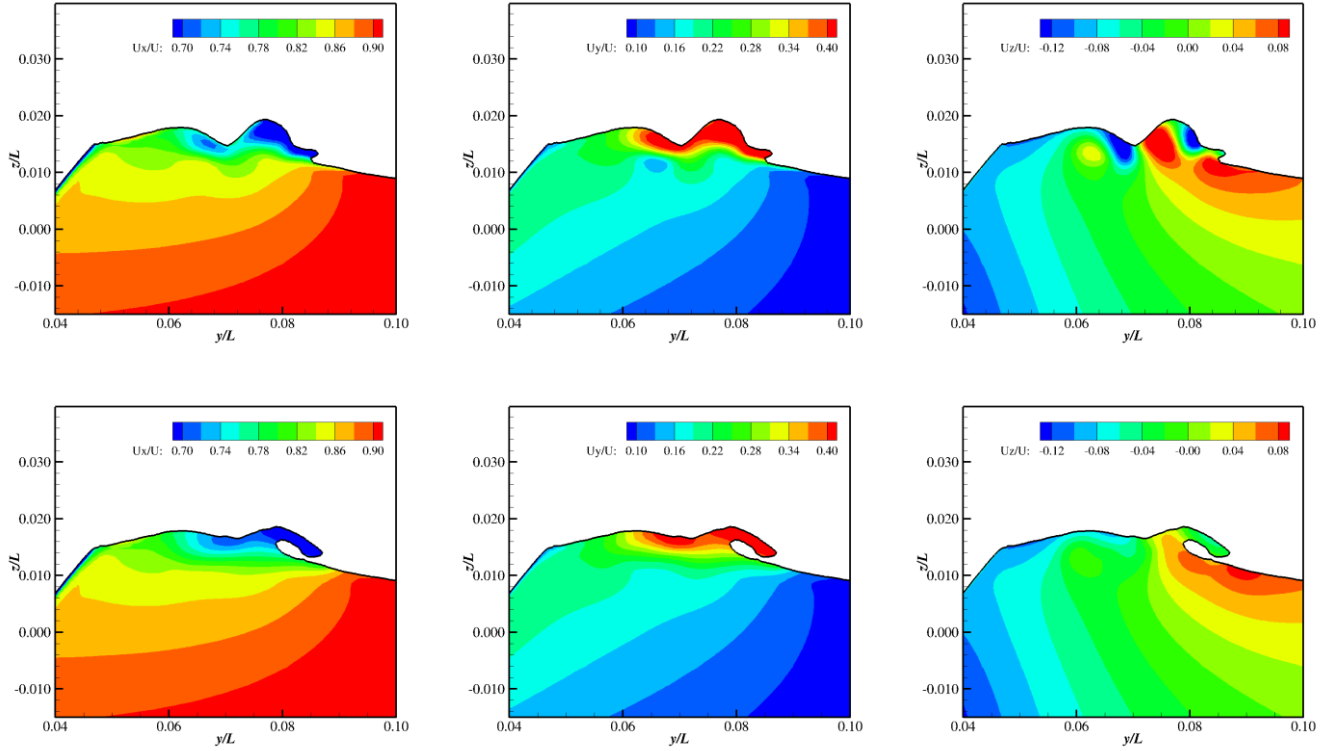


Fig. 12 Wake field distribution at $x/L=0.14$ (top: DES, below: RANS)

Conclusion

In the present study, both turbulence models, DES and RANS, are adopted to simulate the bow wave breaking of KCS. At low speed, $Fr=0.26$, in terms of resistance, the predicted result of RANS approach is slightly more accurate than DES model. While the prediction of wave pattern obtained by the latter is slightly better agreement with experiment data. At high speed, $Fr=0.35$, the resistance achieved by both turbulence models are almost same basically. Though four scars are captured by the DES model, while only one scar is observed via RANS approach. In terms of the 3D vortical structure, DES model yields three main vortices and more fragmented vortices near the bow and the main vortices vanishes until the mid-ship, while RANS approach captures some intermittent vortices that vanish rapidly in the wake. At the cutting planes, $x/L=0.05, 0.07$, the axial vorticity distribution obtained by both numerical schemes are similar. At $x/L=0.14$, the negative vorticity obtained by DES approach is much larger than that achieved by RANS scheme. A counter-rotating vortex pair, which induces the scar, is generated near the free surface. In the wake field, the axial and lateral velocity do not present significant difference. In the results of DES method, positive and negative velocity appear alternately in the vertical direction, while RANS approach only provides a fuzzy distribution for this trend. In the future work, some small scale features, such as air entrainment, capillary wave, should be paid more attention.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145, 11272120), Chang Jiang Scholars Program (T2014099), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23) and Lloyd's Register Foundation for doctoral students, to which the authors are most grateful.

References

- [1] Dong, R.R., Katz, J., and Huang, T.T. (1997) On the structure of bow waves on a ship model, *Journal of Fluid Mechanics* 346, 77–115.
- [2] Duncan, J.H. (1983) The breaking and non-breaking wave resistance of a two-dimensional hydrofoil, *Journal of Fluid Mechanics* 126, 507–520.
- [3] Kayo, Y., and Takekuma, K. (1981) On the Free-Surface Shear Flow related to Bow Wave-Breaking of Full Ship Models, *Journal of the Society of Naval Architects of Japan* 149, 11–20.
- [4] Roth, G.I., Mascenik, D.T., and Katz, J. (1999) Measurements of the flow structure and turbulence within a ship bow wave, *Physics of Fluids* 11(11), 3512–3523.
- [5] Longo, J., and Stern, F. (2002) Effects of drift angle on model ship flow, *Experiments in Fluids* 32(5), 558–569.
- [6] Olivieri, A., Pistani, F., Wilson, R., Campana, E.F., and Stern, F. (2007) Scars and Vortices Induced by Ship Bow and Shoulder Wave Breaking, *Journal of Fluids Engineering* 129(11), 1445–1459.
- [7] Wilson, R.V., Carrica, P.M., and Stern, F. (2006) URANS simulations for a high-speed transom stern ship with breaking waves, *International Journal of Computational Fluid Dynamics* 20(2), 105–125.
- [8] Moraga, F.J., Carrica, P.M., Drew, D.A., and Lahey, R.T. (2008) A sub-grid air entrainment model for breaking bow waves and naval surface ships, *Computers & Fluids* 37(3), 281–298.

- [9] Marrone, S., Colagrossi, A., Antuono, M., Lugni, C., and Tulin, M.P. (2011) A 2D+t SPH model to study the breaking wave pattern generated by fast ships, *Journal of Fluids and Structures* 27(8), 1199–1215.
- [10] Marrone, S., Bouscasse, B., Colagrossi, A., and Antuono, M. (2012) Study of ship wave breaking patterns using 3D parallel SPH simulations, *Computers & Fluids* 69, 54–66.
- [11] Noblesse, F., Delhommeau, G., Liu, H., Wan, D., and Yang, C. (2013) Ship bow waves, *Journal of Hydrodynamics, Ser B* 25(4), 491–501.
- [12] Shen, Z., Cao, H., Ye, H., and Wan, D. (2012) The manual of CFD solver for ship and ocean engineering flows: naoe-FOAM-SJTU, Shanghai Jiao Tong University.
- [13] Cao, H., and Wan, D. (2014) Development of Multidirectional Nonlinear Numerical Wave Tank by naoe-FOAM-SJTU Solver, *International Journal of Ocean System Engineering* 4(1), 52–59.
- [14] Shen, Z., Wan, D., and Carrica, P.M. (2015) Dynamic overset grids in OpenFOAM with application to KCS self-propulsion and maneuvering, *Ocean Engineering* 108, 287–306.
- [15] Zhao, W., Wan, D. (2017) CFD study of VIM of a paired- column semi-submersible platform, *Proceeding of 27th International Ocean and Polar Engineering Conference*. California, USA, June 25-30, 2017, 693-700.
- [16] Menter, F. R., Kuntz, M., Langtry, R. (2003) “Ten years of industrial experience with the SST turbulence model, Turbulence,” *Heat and Mass Transfer* 4, 625-632.
- [17] Zhao, W., Wan, D. (2016) Numerical study of 3D flow past a circular cylinder at subcritical Reynolds number using SST-DES and SST-URANS, *Chinese Journal of Hydrodynamics*, 31(1), 1-8.
- [18] Weller, H.G. (2008) A new approach to VOF-based interface capturing methods for incompressible and compressible flow, OpenCFD Ltd, Report TR/HGW/04.

Numerical Study of Bubble Coalescence by Multiphase MPS Method

Xiao Wen*, Decheng Wan[†]

Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration,
State Key Laboratory of Ocean Engineering, School of Naval Architecture,
Ocean and Civil Engineering, Shanghai Jiao Tong University,
Shanghai, China

*Presenting author: wenxiao2016@sjtu.edu.cn

[†]Corresponding author: dcwan@sjtu.edu.cn

Abstract

In present study, a new mesh-less multiphase method is developed based on the IMPS (Improved Moving Particle Semi-implicit) method and applied to simulate bubble coalescence. In this method, the multiphase system is treated as a multi-density and multi-viscosity system and special interface treatments are introduced for interface region. The new multiphase MPS method is validated through comparisons with data in open literature. In particular, the multiphase MPS method is verified against Hysing et al.'s quantitative benchmark computations of two-dimensional bubble dynamics. Good agreements achieved for the benchmark quantities, including center of mass, and mean rise velocity, which demonstrate the accuracy and stability of the present multiphase MPS method. Then, a numerical study of the coalescence of bubble pairs is conducted and comparison between MPS and experimental results is made. In general, good agreement can be achieved.

Keywords: **Multiphase; MPS; bubble coalescence; CFD**

Introduction

In fluid dynamics, bubbly flow is one of the most important issues and can be found in various natural and industrial processes such as chemical reactors, petroleum refining and boiling. The rising of bubbles in a viscous liquid due to buoyancy is one of fundamental bubbly flow. And when the distances between bubbles become close enough during rising process, interactions of bubbles occur and coalescence may be observed, which can greatly affect the bubble dynamics. Therefore, in order to obtain sufficient acknowledge of the bubble dynamics, the coalescence of bubbles should be primarily investigated.

In recent decades, there have been many experimental and theoretical researches on the coalescence of bubbles [1]-[6]. However, both experimental and theoretical approaches have their limitations. On the one hand, only a small part of flow information can be obtained due to the difficulties in experiments, which results in a limited understanding on coalescence of bubbles. On the other hand, theoretical analysis is also not competitive, because approximate theoretical solutions can be derived only in the cases of small bubble deformations. Therefore, numerical simulations seem to be an effective alternative approach for the study of coalescence of bubbles [7]-[10], with the development of computer technology and computational fluid dynamics (CFD) in recent years. However, the numerical study of bubble coalescence is challenging due to the tracing of largely deformation of multiphase interface. For traditional numerical methods which are based on meshes or grids, there is difficulty in the maintenance of sharp or fragmented interfaces and the meshes containing the interface

need extra and complicated treatments. Besides, the numerical diffusion induced by the advection term is also an unavoidable problem when using these methods.

Through the past few years, the mesh-less particle methods, as a new generation of numerical methods, have been developed energetically and provide a good solution to the problems mentioned above. In the mesh-less particle method, the calculation domain is dispersed by a set of particles with Lagrangian description, which means the profile of interface is automatically presented by the particles distribution, without need of additional capture or tracking procedure. Therefore, particle methods are particularly suitable for problems with large deformation of interface or free surface, like the coalescence of bubbles. Furthermore, numerical diffusion in the discretization of advection term is eliminated because of the use of substantial derivative in governing equations for Lagrangian system. In this paper, the moving particles semi-implicit (MPS) method, which is one of the most commonly known mesh-less particle methods, is employed to simulate the bubble coalescence. Since its being put forward, the MPS method has been successfully applied in a wide range of hydraulic engineering problems, including liquid sloshing [11][12], dam breaks [13][14], water entry [15], and fluid-structure interaction [16]. However, the original MPS method considered only the problems with a single fluid and is not applicable to multiphase flows [17].

A number of significant numerical works have been devoted to extend the MPS method to multiphase flows. The first MPS multiphase method is proposed for solid-liquid two-phase flows [18]. After that, a hybrid MPS-FVM method is developed for the viscous, incompressible, multiphase flows, in which the heavier fluid is represented by moving particles while the lighter fluid is defined on the mesh [19]. Shakibaeinia and Jin [20] studied the multiphase flows by treating the multiphase system as a multi-viscosity and multi-density system, but the unphysical penetration is observed due to the less consideration of interface tension force. Khayyer and Gotoh [21] derived a first-order density smoothing scheme for multiphase flows characterized by high density ratios, and successfully applied this scheme to various multiphase problems including bubble rising. Two multiphase MPS methods, the MMPS-HD and MMPS-CA method are developed by Duan et al. [22]. In MMPS-HD method, harmonic mean density is utilized to avoid extremely large acceleration at interface. In MMPS-CA method, new stable formulations are developed to keep acceleration continuous at interface. Chen et al. [23] performed a two-dimensional numerical simulation of the motion and coalescence of bubble pairs rising in the stationary liquid pool, in which the single-phase MPS method is utilized and the bubble region is simply regarded as a vacuum.

In present study, a multiphase method is developed based on the IMPS method and special treatments on the two-phase interface. In our method, the multiphase system is treated as the multi-density and multi-viscosity fluid. The interparticle viscosity defined by the harmonic mean viscosity of two particles is firstly adopted to consider the interaction between different phases. Then the density smoothing technique is employed to reduce pressure discontinuity crossing the interface and obtain the continuous acceleration and velocity fields. The influence of surface tension force on the interface is considered through a contoured continuum surface force (CCSF) model. The present multiphase MPS is then verified against Hysing et al.'s quantitative benchmark computations of two-dimensional bubble dynamics [24]. Good agreements are achieved for the benchmark quantities, including center of mass, and mean rise velocity, which demonstrates the accuracy and stability of the present multiphase MPS method. Finally, a numerical study of the coalescence of bubble pairs is conducted and comparison between MPS and experimental results is made. In general, good agreement can be achieved.

Numerical Methods

Improved MPS Method

Due to less accurate formulations, the original MPS method suffered from some non-physical fluid motions and pressure oscillations. In this paper, we employed an improved MPS method, which includes improved schemes. In IMPS method, the governing equations for incompressible viscous fluid in Lagrangian system are employed, which can be expressed by continuity and momentum equations:

$$\frac{D\rho}{Dt} = -\rho(\nabla \cdot \mathbf{V}) = 0 \quad (1)$$

$$\rho \frac{D\mathbf{V}}{Dt} = -\nabla P + \mu \nabla^2 \mathbf{V} + \mathbf{f} \quad (2)$$

where \mathbf{V} , t , ρ , P , μ and \mathbf{f} represent the velocity vector, time, fluid density, pressure, dynamic viscosity and the body force, respectively. Similar with the solving strategy adopted by Shakibaeinia and Jin [20] and Duan et al. [22], in present method, the governing equations for all fluids belonging to liquid or bubble are uniform and solved together, thus the multiphase system is treated as a single fluid with multi-density and multi-viscosity.

In particle method, a particle interacts with its neighboring particles covered with kernel function $W(r)$. In present study, the kernel function without singularity is employed [25].

$$W(r) = \begin{cases} \frac{r_e}{0.85r + 0.15r_e} - 1 & 0 \leq r < r_e \\ 0 & r_e \leq r \end{cases} \quad (3)$$

where r is distance between particles and r_e is the effect radius.

In MPS method, all terms of differential operators on the right hand of governing equations are replaced by particle interaction models, including the gradient, divergence and Laplacian model, defined as

$$\langle \nabla \phi \rangle_i = \frac{D}{n^0} \sum_{j \neq i} \frac{\phi_j - \phi_i}{|\mathbf{r}_j - \mathbf{r}_i|^2} (\mathbf{r}_j - \mathbf{r}_i) \cdot W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (4)$$

$$\langle \nabla \cdot \boldsymbol{\Phi} \rangle_i = \frac{D}{n^0} \sum_{j \neq i} \frac{(\boldsymbol{\Phi}_j - \boldsymbol{\Phi}_i) \cdot (\mathbf{r}_j - \mathbf{r}_i)}{|\mathbf{r}_j - \mathbf{r}_i|^2} W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (5)$$

$$\langle \nabla^2 \phi \rangle_i = \frac{2D}{n^0 \lambda} \sum_{j \neq i} (\phi_j - \phi_i) \cdot W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (6)$$

where ϕ is an arbitrary scalar function, $\boldsymbol{\Phi}$ is an arbitrary vector, D is the number of space dimensions, n^0 is the particle number density at initial arrangement, λ is a parameter defined as

$$\lambda = \frac{\sum_{j \neq i} W(|\mathbf{r}_j - \mathbf{r}_i|) \cdot |\mathbf{r}_j - \mathbf{r}_i|^2}{\sum_{j \neq i} W(|\mathbf{r}_j - \mathbf{r}_i|)} \quad (7)$$

which is applied to keep the variance increase equal to that of the analytical solution.

In case of pressure gradient model, direct application of Eq. (4) may bring about the problem of tensile instability. In this paper, we employed the modified pressure gradient model [26], written as:

$$\langle \nabla P \rangle_i = \frac{D}{n^0} \sum_{j \neq i} \frac{P_j - P_{i_{\min}}}{|\mathbf{r}_j - \mathbf{r}_i|^2} (\mathbf{r}_j - \mathbf{r}_i) \cdot W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (8)$$

where $P_{i_{\min}}$ represents the minimal pressure among neighboring particles of particle i .

Fluid density in MPS method is represented by particle number density. Therefore, the incompressible condition can be satisfied by keeping the particle number density constant. For this purpose, each time step in MPS method is divided into two stages: In the first stage, temporal velocity of particles is obtained using viscous and body forces terms, which can be explicitly calculated. In the second stage, pressure term is implicitly calculated by solving the Pressure Poisson Equation (PPE), and the velocity and position of particles are updated to make the fluid incompressible. The PPE in present MPS solver is defined as

$$\langle \nabla^2 P^{n+1} \rangle_i = (1 - \gamma) \frac{\rho}{\Delta t} \nabla \cdot \mathbf{V}_i^* - \gamma \frac{\rho}{\Delta t^2} \frac{\langle n^* \rangle_i - n^0}{n^0} \quad (9)$$

where γ is a blending parameter with a value less than 1. The range of $0.01 \leq \gamma \leq 0.05$ is recommended according to the numerical tests conducted by Lee et al. [27]. In this paper, $\gamma = 0.01$ is adopted for all simulations.

The kinematic and dynamic free surface boundary conditions should be imposed on the free surface particles. The kinematic boundary condition is automatically satisfied in Lagrangian method, while the dynamic free surface boundary condition is implemented by setting zero pressure on the free surface particles. To impose the dynamic free surface boundary condition, free surface particles should be detected at first. In present method, a function [25] based on the asymmetric arrangement of neighboring particles of the center particle is defined as:

$$\langle \mathbf{F} \rangle_i = \frac{D}{n^0} \sum_{j \neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} (\mathbf{r}_i - \mathbf{r}_j) W(r_{ij}) \quad (10)$$

where the vector function \mathbf{F} represents the asymmetry of arrangements of neighbor particles. Particles satisfying

$$\langle |\mathbf{F}| \rangle_i > 0.9 |\mathbf{F}|^0 \quad (11)$$

are judged as free surface particles, where $|\mathbf{F}|^0$ is the value of $|\mathbf{F}|$ for surface particles at initial arrangement.

Special Treatments for Multi-phase Flows

To extend the IMPS method into multiphase flows, special treatments are introduced in present paper. For multiphase flows, the mathematical discontinuity of density at two-phase interface causes a discontinuous acceleration field and accordingly numerical instabilities. In present study, the density smoothing scheme [20][21] is adopted for interface particles, which is based on a simple spatial averaging as follow:

$$\langle \rho \rangle_i = \frac{\sum_{j \in I} \rho_j W(r_{ij}, r_e)}{\sum_{j \in I} W(r_{ij}, r_e)} \quad (14)$$

where I include particles i and all its neighboring particles.

In present paper, the multi-viscosity model [20] is employed to deal with viscosity discontinuity cross interface. With the multi-viscosity model, the viscous term in Eq. (2) can be presented as:

$$\langle \mu \nabla^2 \mathbf{V} \rangle_i = \frac{2D}{\lambda n^0} \sum_{j \neq i} \mu_{ij} (\mathbf{V}_j - \mathbf{V}_i) W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (15)$$

$$\mu_{ij} = \frac{2\mu_j\mu_i}{\mu_j + \mu_i} \quad (16)$$

where μ_{ij} represents the interparticle viscosity between particles i and j . It is obvious that μ_{ij} is equal to the real viscosity of one fluid ($\mu_{ij}=\mu_i=\mu_j$) when particles i and j belong to a same phase.

For multiphase flows, interface tension effects are important when deformations of fluid interfaces are involved. One of the most widely used approach for incorporation of interface tension effects is the continuum surface force (CSF) method [28]. In this approach, surface/interface tension is converted into a body force in a transition region across the interface. Therefore, for particles in this transition region, the momentum equation Eq. (2) of governing equations transforms into

$$\rho \frac{D\mathbf{V}}{Dt} = -\nabla P + \mu \nabla^2 \mathbf{V} + \mathbf{f} + \mathbf{f}_s \quad (17)$$

$$\mathbf{f}_s = \sigma \kappa \nabla C \quad (18)$$

where \mathbf{f}_s is the interface tension, σ is the surface tension coefficient, κ is the interface curvature, C is a color function defined as:

$$C_i = \begin{cases} 0 & \text{particle } i \text{ belongs to the specified phase} \\ 1 & \text{particle } i \text{ belongs to the other phase} \end{cases} \quad (19)$$

The surface tension coefficient σ is known for a certain fluid. Therefore, the most important issue in CSF model is the calculation of interface curvature κ , which is also the most difficult issue. In present study, the curvature is analytically calculated based on the contour of color function [29]. The contour of color function C is approximately regarded as the profile of interface. With the contour of color function obtained, the interface curvature can be calculated. The calculation process of curvature can be simply divided into four steps. First, the value of the color function f at an arbitrary location (x, y) in the vicinity of the target particle i can be obtained by performing a spatial weighted averaging of all the neighboring particles of particle i , through the implementation of a Gaussian kernel function G :

$$f(x, y) = \frac{\sum_j C_j G(|\mathbf{r} - \mathbf{r}_j|, r_s)}{\sum_j G(|\mathbf{r} - \mathbf{r}_j|, r_s)} \quad (20)$$

$$G(r_{ij}, r_s) = \frac{9}{\pi r_s^2} \exp\left(-\frac{9r_{ij}^2}{r_s^2}\right) \quad (21)$$

where r_s represents the effect radius of the Gaussian kernel function.

Second, $f(x, y)$ is expanded at point i through the Taylor series expansion:

$$\begin{aligned} f(x, y) = & f(x_i, y_i) + f_{x,i}(x - x_i) + f_{y,i}(y - y_i) + \frac{1}{2} f_{xx,i}(x - x_i)^2 \\ & + f_{xy,i}(x - x_i)(y - y_i) + \frac{1}{2} f_{yy,i}(y - y_i)^2 + O(r_s^3) \end{aligned} \quad (22)$$

where the subscripts x and y represent the partial derivatives with respect to x and y at particle i , respectively. The calculation of the partial derivatives can be found in [29]. $O(r_s^3)$ is the high order error term which has no influence on the accuracy of curvature calculation and can be omitted [29].

Third, considering that the local contour of color function passing through particle i must satisfy $f(x, y) = f(x_i, y_i)$, we can obtain the equation of the local contour at particle i from Eq. (22):

$$f_{x,i}(x - x_i) + f_{y,i}(y - y_i) + \frac{1}{2} f_{xx,i}(x - x_i)^2 + f_{xy,i}(x - x_i)(y - y_i) + \frac{1}{2} f_{yy,i}(y - y_i)^2 = 0 \quad (23)$$

Finally, the curvature at particle i can be analytically calculated as follow:

$$\kappa_i = \frac{y''}{(1 + y_i')^{3/2}} = \frac{2f_{x,i}f_{y,i}f_{xy,i} - f_{x,i}^2f_{yy,i} - f_{y,i}^2f_{xx,i}}{(f_{x,i}^2 + f_{y,i}^2)^{3/2}} \quad (24)$$

Numerical Validation

The new multiphase method is verified against Hysing et al.'s quantitative benchmark computations of two-dimensional bubble rising [24]. The initial condition and parameters for the test case are shown in Fig. 1 and Table 1. On the one hand, the qualitative results of the bubble shapes during rising process are compared. On the other hand, some quantitative comparisons are also provided, including the center of mass (y_{bubble}) and rise velocity of the bubble (u_{bubble}), which are defined as below:

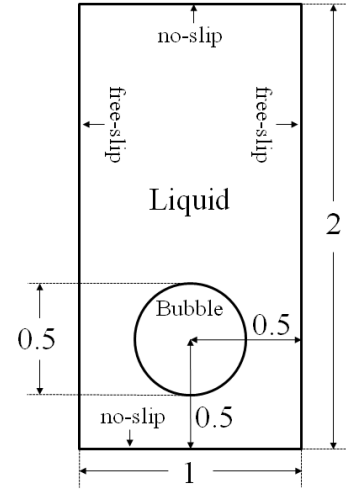
$$y_{\text{bubble}} = \left(\sum_i^{N_{\text{bubble}}} y_i \right) / N_{\text{bubble}} \quad (25)$$

$$u_{\text{bubble}} = \left(\sum_i^{N_{\text{bubble}}} u_{yi} \right) / N_{\text{bubble}} \quad (26)$$

Where N_{bubble} is the total number of bubble particles, y_i is the vertical coordinate of the bubble particle i , u_{yi} is the vertical velocity of the bubble particle i .

Table 1. Parameters for the simulation of bubble rising

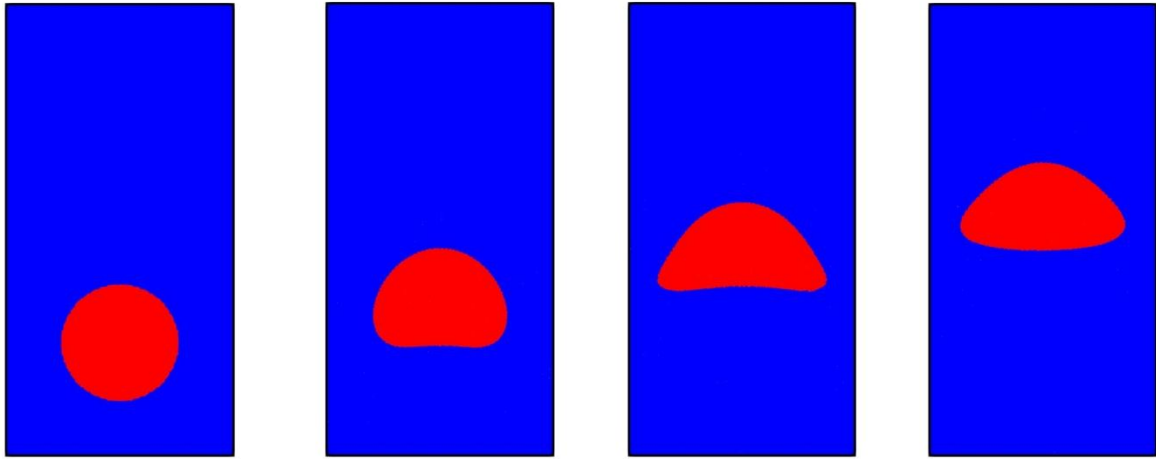
Parameters	Value	Unit
Computational domain	1×2	m
Particle spacing	0.01	m
Total particle number	25389	/
Bubble diameter (r_0)	0.5	m
Liquid density	1000	kg/m ³
Bubble density	100	kg/m ³
Liquid viscosity	10	kg/(m. s)
Bubble viscosity	1	kg/(m. s)
Time step size	0.0001	s
Gravitational acceleration	-0.98	s/m ²


Figure 1. Initial condition for the validation case

In present case, the non-dimensional Bo number can be defined as

$$Bo = \frac{\rho_1 U_g^2 L}{\sigma} \quad (27)$$

where ρ_1 refers to the density surrounding heavier fluid, $L=2r_0$ is the characteristic length scale, $U_g^2=g2r_0$ is the gravitational velocity. The Bo number is relatively small in this case, equal to 10, which means the surface tension effects are strong enough compared with gravitational forces and can hold the bubble together. As shown in Fig. 2, the rising bubble ends up in the ellipsoidal regime without break up. The bubble shapes simulated by present method at $t = 3$ s are compared with the benchmarks results, where good agreement are obtained. Furthermore, the results of the evolution of center of mass and rise velocity are presented in Fig. 4. The bubble rises at a growing speed at the early stage, until a largest velocity appears. After that, the rising of the bubble gradually slows down and a steady velocity can be finally reached, which can be called terminal velocity. In general, good agreements with benchmark results can be observed. Besides, the fluctuations of the rise velocity in some other particles methods are not found in present results, due to the improved accuracy and stability of the IMPS.


Figure 2: Time evolution of the bubble shapes during rising process. ($t = 0$ s, 1 s, 2 s, 3 s)

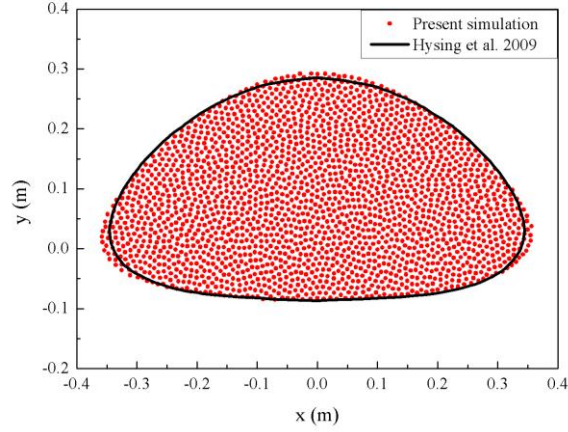


Figure 3: Comparison of bubble shapes at $t = 3$ s for bubble rising

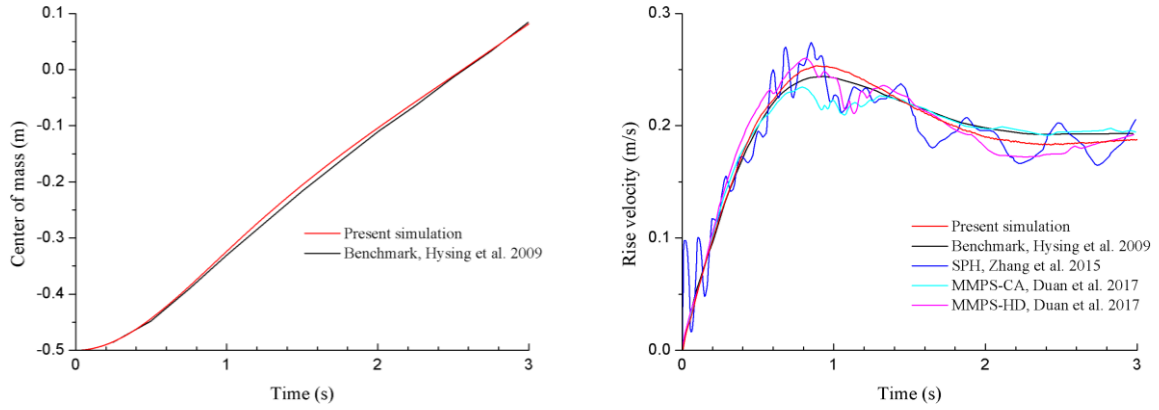


Figure 4: Comparison of center of mass and rise velocity for bubble rising

Numerical Simulation of Bubble Coalescence

In this section, the bubble coalescence is numerically simulated by the newly developed method and comparison between numerical and experimental results is carried out. As shown in Fig. 5, two bubbles are initially arranged side by side at horizontal direction. Then the two bubbles will rise under the effects of buoyancy. According to the Bernoulli's law, the two bubbles will move closer to each other during the rising process, due to the faster fluid velocity and lower pressure between the two bubbles, and bubble coalescence can occur. The initial condition and parameters for the test case are shown in Fig. 5 and Table 2.

Fig. 6 shows the snapshots of the coalescence of two bubbles obtained by present multiphase MPS method and observed in experiment [2] at some characteristic time instants. It can be seen from Fig. 6 that as the two bubbles rose, they would at first approach and finally get in touch with each other. Then a gas-bridge forms between and connected the two bubbles. Under the effects of interface tension force, the liquid between two bubbles are rapidly excluded and the gas bridge expands in the vertical direction, which results in the coalescence of the two bubbles. It can be clearly seen from comparison between Fig. 6(a) and Fig. 6(b) that the coalescence processes obtained the MPS simulation and experiment shows good accordance in general.

Table 2. Parameters for the bubble coalescence

Parameters	Value	Unit
Computational domain	0.05×0.1	m
Particle spacing	0.00025	m
Total particle number	90489	/
Bubble diameter ($2r_0$)	0.01	m
Liquid density	1000	kg/m ³
Bubble density	10	kg/m ³
Liquid viscosity	1.31×10^{-2}	kg/(m. s)
Bubble viscosity	1.31×10^{-4}	kg/(m. s)
Time step size	0.00000125	s
Gravitational acceleration	9.8	s/m ²

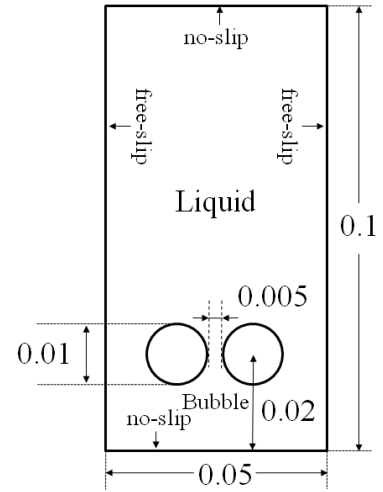


Figure 5. Initial condition for the bubble coalescence

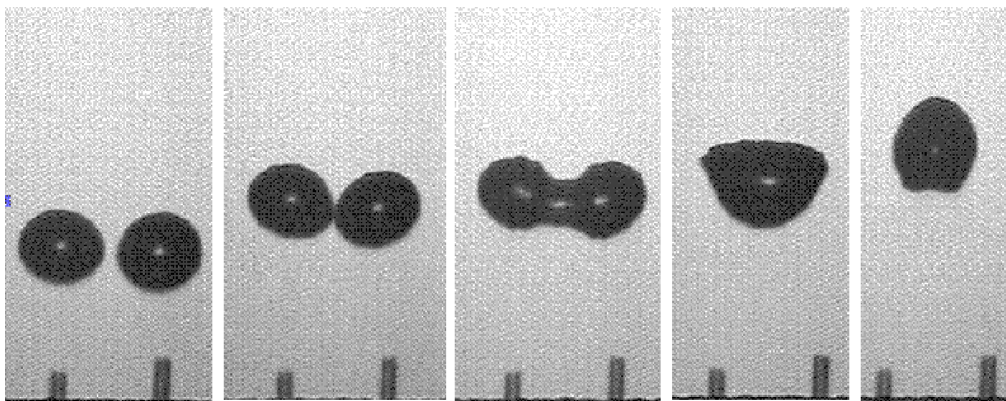
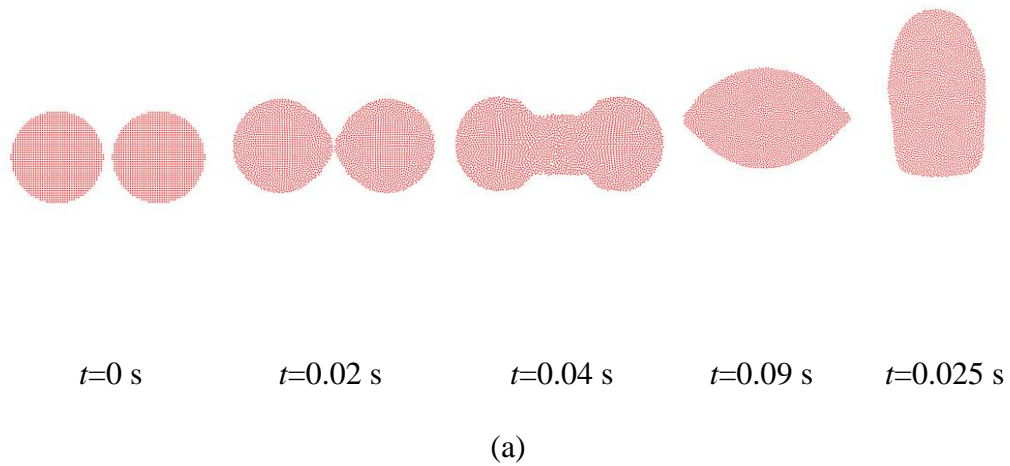


Figure 6. Comparison of coalescence of two bubbles between present MPS simulation and experiment in the literature: (a) simulative result; (b) experimental result from literature [2].

Conclusions

In this study, the numerical study of bubble coalescence is carried out by using a newly developed multiphase MPS method based on the IMPS method. The present method treats the multiphase system as the multi-density and multi-viscosity fluid, thus only a single set of equations needs to be solved for all phases. Besides, extra density smoothing technique, interparticle viscosity model and surface tension model are included in the present method for interface particles. The new method is verified against Hysing et al.'s quantitative benchmark computations of two-dimensional bubble dynamics [24]. In general, both the qualitative comparison of bubble shapes and quantitative comparisons of center of mass and mean rise velocity show good agreement, validating the accuracy and stability of the present multiphase MPS method. Then, the newly developed method is applied to simulate bubble coalescence. The coalescence process obtained by present method is in good agreement with the experimental results observed by Duineveld [2].

Acknowledgement

This work is supported by the National Natural Science Foundation of China (51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

- [1] Prince, M. J. and Blanch, H. W. (1990) Bubble coalescence and break-up in air sparged bubble columns, *Aiche Journal* 36(10), 1485-1499.
- [2] Duineveld, P. C. (1998) Bouncing and coalescence of bubble pairs rising at high Reynolds number in pure water or aqueous surfactant solutions, *Applied Scientific Research* 58(1-4), 409-439.
- [3] Sanada, T., Watanabe, M. and Fukano, T. (2005) Effects of viscosity on coalescence of a bubble upon impact with a free surface, *Chem. Eng. Sci.* 60, 5372-5384.
- [4] Sanada, T., Sato, A., Shiota, M. and Watanabe, M. (2009) Motion and coalescence of a pair of bubbles rising side by side, *Chemical Engineering Science* 64(11), 2659-2671.
- [5] Suñol, F. and González-Cinca, R. (2010) Rise, bouncing and coalescence of bubbles impacting at a free surface, *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 365(1-3), 36-42.
- [6] Kemiha, M., Dietrich, N., Poncin, S. and Li, H. Z. Coalescence between bubbles in non-newtonian media, *Proceedings of 8th World Congress of Chemical Engineering*, Montreal, Canada, 2009, August 23-27.
- [7] Legendre, D., Magnaudet, J. and Mougin, G. (2003), Hydrodynamic interactions between two spherical bubbles rising side by side in a viscous liquid, *Journal of Fluid Mechanics* 497(497), 133-166.
- [8] Farhangi, M. M., Passandideh-Fard, M. and Moin, H. (2010), Numerical study of bubble rise and interaction in a viscous liquid, *International Journal of Computational Fluid Dynamics* 24(1), 13-28.
- [9] Liu, J., Zhu, C., Fu, T., Ma, Y. and Li, H. (2013) Numerical simulation of the interactions between three equal-interval parallel bubbles rising in non-Newtonian fluids, *Chemical Engineering Science* 93, 55-66.
- [10] Liu, J., Zhu, C., Wang, X., Fu, T., Ma, Y. and Li, H. (2015) Three-dimensional numerical simulation of coalescence and interactions of multiple horizontal bubbles rising in shear-thinning fluids, *AIChE Journal* 61(10), 3528-3546.
- [11] Zhang, Y. X. and Wan, D. C. (2014) Comparative study of MPS method and Level-Set method for sloshing flows, *Journal of Hydrodynamics* 26(4), 577-585.
- [12] Lee, B. H., Park, J. C., Kim, M. H. and Hwang, S. C. (2011) Moving particle simulation for mitigation of sloshing impact loads using surface floaters, *Computer Modeling in Engineering and Sciences* 75(2), 89-112.

- [13]Zhang, Y. X. and Wan, D. C. (2011) Application of MPS in 3D dam breaking flows, *Scientia Sinica* 41, 140-151. (in Chinese)
- [14]Shakibaeinia, A. and Jin, Y. C. (2010) A weakly compressible MPS method for simulation open-boundary free-surface flow, *International Journal for Numerical Methods in Fluids* 63(10), 1208–1232.
- [15]Tang, Z. Y., Wang, J. H., Zhang, Y. L. and Wan, D. C. (2015) Numerical investigation of 3D water entry based on IMPS method and overset method, *Proceedings of 3rd International Conference on Violent Flows*, Osaka, Japan, 2015.
- [16]Zhang, Y. L. and Wan, D. C. (2017) MPS-FEM coupled method for sloshing flows in an elastic tank, *Ocean Engineering* 152, 416-427.
- [17]Koshizuka, S. and Oka, Y. (1996) Moving-particle semi-implicit method for fragmentation of incompressible fluid, *Nuclear Science & Engineering* 123(3), 421-434.
- [18]Gotoh, H. and Fredsøe, J. Lagrangian two-phase flow model of the settling behavior of fine sediment dumped into water, *Proc. ICCE*, Sydney, 2000, 3906–3919.
- [19]Liu, J., Koshizuka, S. and Oka, Y. (2005) A hybrid particle-mesh method for viscous, incompressible, multiphase flows, *Journal of Computational Physics* 202(1), 65-93.
- [20]Shakibaeinia, A. and Jin, Y. (2012) MPS mesh-free particle method for multiphase flow, *Comput Methods Appl Mech Eng* 229-232, 13-26.
- [21]Khayyer, A. and Gotoh, H. (2013) Enhancement of performance and stability of MPS mesh-free particle method for multiphase flows characterized by high density ratios, *Journal of Computational Physics* 242, 211-233.
- [22]Duan, G., Chen, B., Koshizuka, S. and Xiang, H. (2017) Stable multiphase moving particle semi-implicit method for incompressible interfacial flow, *Comput Methods Appl Mech Engrg* 318, 636–666.
- [23]Chen, R. H., Tian, W. X., Su, G. H., Qiu, S. Z., Ishiwatari, Y. and Oka, Y. (2011) Numerical investigation on coalescence of bubble pairs rising in a stagnant liquid, *Chemical Engineering Science* 66(21), 5055–5063.
- [24]Hysing, S., Turek, S., Kuzmin, D., Parolini, N., Burman, E., Ganesan, S. and TobiskaL, T. (2009) Quantitative benchmark computations of two-dimensional bubble dynamics, *International Journal for Numerical Methods in Fluids* 60(11), 1259-1288.
- [25]Zhang, Y. X. and Wan, D. C. Apply MPS method to simulate liquid sloshing in LNG tank, *Proceedings of the 22nd international offshore and polar engineering conference*, 2012, 381-391.
- [26]Koshizuka, S., Nobe, A. and Oka, Y. (1998) Numerical analysis of breaking waves using the moving particle semi-implicit method, *International Journal for Numerical Methods in Fluids* 26(7), 751-769.
- [27]Lee, B. H., Park, J. C., Kim, M. H. and Hwang S. C. (2011) Step-by-step improvement of MPS method in simulating violent free-surface motions and impact-loads, *Computer Methods in Applied Mechanics and Engineering* 200(9), 1113-1125.
- [28]Brackbill, J. U., Kothe, D. B. and Zemach, C. (1992) A continuum method for modeling surface tension, *Journal of Computational Physics* 100(2), 335–354.
- [29]Duan, G., Koshizuka, S. and Chen, B. (2015) A contoured continuum surface force model for particle methods, *J Comput Phys* 298, 280–304.

Authors Index

Title	Page No	Title	Page No
Akram, Saima	787	Cui, X.	364
Alberto, Viskovic	868	D' Ayala,, Dina	435
Alejandro, Jimenez	285	Das, Raj	727
Alshomrani, Ali Saleh	566, 822, 874	Dass, Aoop Kumar	847
Antonini, Alessandro	435	Del Monaco, Anna Irene	472
Argyros, Ioannis K.	566, 874	Demartino, Cristoforo	929
Aurilio, Marianna	859	Deng, Di	961
Behl, Ramandeep	566, 822, 874	Dermot, O'Dwyer	285
Beltrán, Pablo R.	454	Di Gangi, Giorgia	929
Bishay, Peter L.	486	Ding, Yi	1
Bogoevska, Simona	128	Dobroka, Mihaly	903
Brzezinski, Dariusz W.	103	Dolatkah, Saeed	168
Cai, Dong-Yi	120	Dong, Xiangwei	42
Calderini, Chiara	418	Dragisic, Maja	752
Caramihai, M.	815	Duan, Shuyong	364, 368, 591
Chakraborty, Debabrata	264	Dumova-Jovanoska, Elena	128
Chakraborty, Debaleena	264	Elias, Jan	448
Chang, Che-Kai	120	Erdönmez, Cengiz	492
Chang, Se-Myong	924	Ernst, Frederik	917
Chatzi, Eleni	128	Espinosa, Guadalupe	29, 454
Chen, Jinjian	386	Faella, Giuseppe	859
Chen, Linfeng	638	Fan, Longling	661
Chen, T.	198, 211	Fan, S.Y.	211
Chen, Zhiyi	238, 629, 650	Ferrari, Rosalba	506
Cheng, Ping	973	Ferrero, Chiara	418
Cheng, Yung-Chang	767	Fu, Pak Hang	255
Choi, Shinbeom	299	Furukawa, Akira	706
Chou, Chun-Mei	586	Gaponenko, Elena	949
Clementi, Francesco	496	Gáspár, Csaba	278
Cocchetti, Giuseppe	506	Gawell, Ewelina	518
Coda, Humberto Breves	219	Gazzani, Valentina	496
		Giddens, Don P.	622

Gil, Juan Miguel	29, 454	Lee, Heow Pueh	138
Goncalves, Eric	603	Lenci, Stefano	496
Gopalan, Harish	394	Li, Ming	721, 897
Guadagnuolo, Mariateresa	859	Li, Mingzhi	151
Guo, Haishan	608	Li, T. S.	908, 939
Guo, Xiaoya	622	Li, Yanhua	897
He, Jiawei	989	Liang, Lihong	116
He, Yanping	151	Liang, Sunbin	238
Hirose, Sohichi	706	Liao, Chencong	386
Höffer, Rüdiger	128	Liao, D.M.	198, 211
Honório, Hermínio T.	372	Lignola, Gian Piero	668
Hsiao, Hsi-Chi	586	Lim, Arthur	394
Hsu, Feng-Chi	120	Lim, Kian-Meng	138
Hu, Junjiao	62	Liotta, Marc'Antonio	929
Huang, Chao	151	Liou, Jin-Long	767
Huang, Lihua	76	Liu, Kang	608
Huang, Yang	973	Liu, Yadong	151
Huang, Yi	1, 12	Liu, Zhiqian	629
Huo, S. H.	364	Liu, S.	591
Iezzi, Fabrizio	19	Lojanica, Vladimir	752
Inoue, Takumi	714	Long, Hao	116
Jayaraman, Krishnan	727	Long, Nguyen Nhut Phi	681
Jen, Ming-Hwa R.	120	Lou, Jing	394
Jiang, Zonglin	270	Lourenço, Paulo B.	418
Jiao, Geng	608	Lozano-Guzmán, Alejandro A.	88
Kalita, Paragmoni	847	Luo, Changtong	270
Kamiński, Marcin	184	Ma, Genshan	622
Katamine, E.	595	Maehara, Akiko	622
Kentel, Behzat B	855	Majchrzak, Ewa	835
Kim, Jongwook	299	Maliska, Clovis R.	372
Kosec, Gregor	887	Malyshev, Dmitry	949
Kozulyn, Alexander A.	546	Melania, Lucci	868
Květoň, Josef	448	Menges, Achim	917
Kwon, Hyun-Jin	924	Miao, Aiqin	536
Lam, Kuan Hou	255	Milani, Gabriele	496
Lao, Seng Kin	255	Mintz, Gary S.	622

Mohanna, Aya	741	Raby, Alison	435
Molony, David	622	Ramaglia, Giancarlo	668
Mondal, Hiranmoy	305	Rêbielak, Janusz	347, 356
Monti, Giorgio	929	Ren, Zhen	1001
Morooka, Shigehiro	248	Repka, M.	400
Murthy, K. S R. K.	264	Riveros-Jerez, Carlos Alberto	776
Ng, K. S.	939	Rizzi, Egidio	506
Niu, Ruiping	675	Rodriguez, Rafael	29, 454
Noguchi, Gohki	714	Rokicki, Wieslaw	518
Novak, Pavol	486	Romero, José A.	88
Nowak, Anna	518	Rybak, Larisa	949
Nuamah, Daniel	903	Saitoh, Takahiro	706, 714
Okada, N.	595	Samady, Habib	622
Oliveira, Tiago	332	Sator, L.	400
Ong, Eng Teo	138	Savvas, Dimitrios	322
Onishi, Yuki	804	Sepehrirahnama, Shahrokh	138
Onodera, Takashi	706	Serbenta, Vitas A.	558
Ormos, Tamas	903	Severin, Irina	815
Otremba, Frank	88	Shen, Chien-Hua	586
Panamoottil, Shyam M.	727	Shen, Tsu-Chuan	586
Pappas, Athanasios	435	Sibanda, Precious	305
Peng, Yong	62	Sin, Vai Kuong	255
Pereira, Elvis	332	Siqueira, Tiago Morkis	219
Perminov, Valeriy	527	Skripnyak, Evgeniya G.	546, 558
Poh, Hee Joo	394	Skripnyak, Natalia V.	558
Poiani, Marina	496	Skripnyak, Vladimir A.	546, 558
Popov, Alexey	467	Skripnyak, Vladimir V.	546, 558
Portela, Artur	332	Sladek, Jan	400, 486
Posypkin, Mikhail	949	Sladek, Vladimir	400, 486
Prota, Andrea	668	Slak, Jure	887
Qi, Hu	608	Sokolowski, Damian	184
Qie, Y.H.	364	Somogyi-Molnar, J.	903
Qu, Tian-yang	12	Son, Nguyen Hoai	681
Quan, Nguyen	681	Song, Shi-de	1, 12
Quaranta, Giuseppe	693, 929	Spena, Francesco Russo	668
		Stefanou, George	322

Stefanska, Anna	518	Wong, S. M.	908, 939
Sun, F.	198	Wu, Hao	386
Surh, Hanbum	299	Wu, S.C.	198
Tam, Hou Kuan	255	Xie, Pengpeng	62
Tam, Lap Mou	255	Xu, Di	661
Tang, Dalin	622, 661	Xu, George	394
Tang, Iok Wa	255	Xu, Pengfei	591
Tang, Y.L.	211	Yang, Chun	622, 661
Teng, Z.H.	198	Yang, Zhiquan	76
Tian, Lida	608	Yanov, Dmitri	467
Tin, Le Trung	681	Yao, Jing	661
Trentadue, Francesco	693	Yasmin, Nusrat	787
Turchan, Lukasz	835	Yi, Shengen	62
Uehara, Takuya	700	Yue, Junhong	675, 721
Vailati, Marco	929	Zafar, Fiza	787
Valencia-Cardenas, Maria	776	Zeidan, Dia	603
Catalina		Zelepugin, Alexey S.	467
Valente, Claudio	19	Zelepugin, Sergey A.	467
Vélez, Wilber	332	Zhang, Bu	650
Wan, Decheng	536, 961, 973, 989, 1001, 1013	Zhang, Yufei	675
Wang, Chunqiao	721	Zhang, Z. B.	198
Wang, Jian	386	Zhang, Zhanming	368
Wang, Jianhua	386, 1001	Zhao, Weiwen	989
Wang, Xiao-na	1, 12	Zheng, Jie	622
Wang, Yuefang	76	Zhou, Lilan	638
Wang, Zhe	961	Zhu, Jian	622
Wei, Yueguang	116	Zhu, Xinyao	638
Wen, Xiao	1013		