

Structural Integrity 29

Series Editors: José A. F. O. Correia · Abílio M. P. De Jesus

Dimitrios Pavlou · Hojjat Adeli ·

José A. F. O. Correia · Nicholas Fantuzzi ·

Georgios C. Georgiou ·



Knut Erik Giljarhus · Yanyan Sha *Editors*

Advances in Computational Mechanics and Applications

OES 2023

 Springer

Series Editors

José A. F. O. Correia , *Faculty of Engineering, University of Porto, Porto, Portugal*
Abílio M. P. De Jesus , *Faculty of Engineering, University of Porto, Porto, Portugal*

Advisory Editors

Majid Reza Ayatollahi, *School of Mechanical Engineering, Iran University of Science and Technology, Tehran, Iran*

Filippo Berto, *Department of Mechanical and Industrial Engineering, Faculty of Engineering, Norwegian University of Science and Technology, Trondheim, Norway*

Alfonso Fernández-Canteli, *Faculty of Engineering, University of Oviedo, Gijón, Spain*

Matthew Hebdon, *Virginia State University, Virginia Tech, Blacksburg, VA, USA*

Andrei Kotousov, *School of Mechanical Engineering, University of Adelaide, Adelaide, SA, Australia*


Grzegorz Lesiuk, *Faculty of Mechanical Engineering, Wrocław University of Science and Technology, Wrocław, Poland*

Yukitaka Murakami, *Faculty of Engineering, Kyushu University, Higashiku, Fukuoka, Japan*

Hermes Carvalho, *Department of Structural Engineering, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil*

Shun-Peng Zhu, *School of Mechatronics Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, China*

Stéphane Bordas, *University of Luxembourg, Esch-sur-Alzette, Luxembourg*

Nicholas Fantuzzi , *DICAM Department, University of Bologna, Bologna, Bologna, Italy*

Luca Susmel, *Civil Engineering, University of Sheffield, Sheffield, UK*

Subhrajit Dutta, *Department of Civil Engineering, National Institute of Technology Silchar, Silchar, Assam, India*

Pavlo Maruschak, *Ternopil IP National Technical University, Ruska, Ukraine*

Elena Fedorova, *Siberian Federal University, Krasnoyarsk, Russia*

The *Structural Integrity* book series is a high level academic and professional series publishing research on all areas of Structural Integrity. It promotes and expedites the dissemination of new research results and tutorial views in the structural integrity field. The Series publishes research monographs, professional books, handbooks, edited volumes and textbooks with worldwide distribution to engineers, researchers, educators, professionals and libraries. This series is managed by team members of the ESIS/TC12 technical committee. Topics of interested include but are not limited to:

- Structural integrity
- Structural durability
- Degradation and conservation of materials and structures
- Dynamic and seismic structural analysis
- Fatigue and fracture of materials and structures
- Risk analysis and safety of materials and structural mechanics
- Fracture Mechanics
- Damage mechanics
- Analytical and numerical simulation of materials and structures
- Computational mechanics
- Structural design methodology
- Experimental methods applied to structural integrity
- Multiaxial fatigue and complex loading effects of materials and structures
- Fatigue corrosion analysis
- Scale effects in the fatigue analysis of materials and structures
- Fatigue structural integrity
- Structural integrity in railway and highway systems
- Sustainable structural design
- Structural loads characterization
- Structural health monitoring
- Adhesives connections integrity
- Rock and soil structural integrity.

**** Indexing: The books of this series are submitted to Web of Science, Scopus, Google Scholar and Springerlink ****

Springer and the Series Editors welcome book ideas from authors. Potential authors who wish to submit a book proposal should contact Dr. Mayra Castro, Senior Editor, Springer (Heidelberg), e-mail: mayra.castro@springer.com

Dimitrios Pavlou · Hojjat Adeli ·
José A. F. O. Correia · Nicholas Fantuzzi ·
Georgios C. Georgiou · Knut Erik Giljarhus ·
Yanyan Sha
Editors

Advances in Computational Mechanics and Applications

OES 2023


Editors

Dimitrios Pavlou
University of Stavanger
Stavanger, Norway

Hojjat Adeli
Ohio State University
Columbus, OH, USA

José A. F. O. Correia
Faculty of Engineering
University of Porto
Porto, Portugal

Nicholas Fantuzzi
University of Bologna
Bologna, Italy

Georgios C. Georgiou 
Department of Mathematics and Statistics
University of Cyprus
Nicosia, Cyprus

Knut Erik Giljarhus
Faculty of Science and Technology
University of Stavanger
Stavanger, Norway

Yanyan Sha
University of Stavanger
Stavanger, Norway

ISSN 2522-560X

ISSN 2522-5618 (electronic)

Structural Integrity

ISBN 978-3-031-49790-2

ISBN 978-3-031-49791-9 (eBook)

<https://doi.org/10.1007/978-3-031-49791-9>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Preface

Computational methods are the main tools to simulate the response of structures and mechanical systems without making full-scale installations. Novel algorithms and computer codes for data-driven engineering, diagnostic and prognostic methods, and automatic control, have benefited from the developments in machine learning, optimization techniques, and numerical analysis.

Nowadays, structural design, stress and displacement analysis, structural health monitoring, and dynamic response of civil and mechanical engineering systems take place with the aid of commercial software packages. Still, the commercial packages are not able to solve complex real-life problems, and the assumptions in the calculations can usually result in unrealistic solutions.

Improved computational techniques are always necessary for better accuracy and realistic simulations. The present volume aims to contribute with new proposals and methodologies for numerical simulation of engineering structures and mechanical components.

Key topics for this purpose are machine learning methods, optimization techniques, automatic control, data-driven engineering, diagnostic and prognostic methods, structural integrity, and engineering mathematics.

The present volume contains recent *Advances in Computational Mechanics and Applications*. The included works have been developed by renowned researchers and have been presented in the 1st Olympiad in Engineering Science that took place in May 2023, in Olympia, Greece.

All the papers were reviewed by two experts under the strict quality standards of the Olympiad.

The conference was coordinated by the **International Olympiad Committee** consisting of five eminent professors representing the five continents:

Based on the outstanding impact of his work in Engineering, the honorary chairman of the International Olympiad Committee in Engineering Science was Prof. Hojjat Adeli.

Members of each continent:

Europe: Prof. Dimitrios Pavlou, University of Stavanger, Norway (Chairman of the committee)

America: Prof. Hojjat Adeli, Ohio State University, USA (Honorary chairman)

Asia: Prof. Gui Yun Tian, Newcastle University, UK, and University of Electronic Science and Technology of China, China

Africa: Prof. Ahmed G. Radwan, Nile University, Egypt

Australia: Prof. Priyan Mendis, University of Melbourne, Australia

The **Chairs of the Scientific Committees** were renowned researchers worldwide:

- Professor Nicholas Fantuzzi, Mechanics of Solid and Structures (Europe)
- Professor Georgios Georgiou, Fluid Mechanics (Europe)
- Professor José A. F. O. Correia, Materials Science (Europe)
- Professor Yuzhu Li, Climate Resilience and Coastal Engineering (Asia)
- Professor Tiago Fazerer-Ferradosa, Subsea Engineering (Europe)
- Professor Ranjith Dissanayake, Civil Engineering Structures (Asia)
- Professor Jon G. Hall, Expert Systems (Europe)
- Professor Joo-Hyung Kim, Intelligent Systems (Asia)
- Professor Nikolas Xiros, Marine and Ocean Engineering Process, Modelling and Simulation (America)
- Professor Wim De Waele, Structural Integrity of Wind Turbines (Europe)
- Professor Kazuyuki Nakahata, Nondestructive testing and evaluation—NDT/NDE (Asia)
- Professor Xiaoping Huang, Offshore Structures and Ship Design (Asia)
- Professor Stefanos Zaoutsos, Mechanical Behaviour of Viscoelastic Materials (Europe)

The Editors

Contents

Machine Learning

A Combined Machine Learning and Computational Methodology for Optimum Thrust Bearings' Behavior in Mixed Lubrication Regime	3
<i>Konstantinos P. Katsaros and Pantelis G. Nikolakopoulos</i>	

Limited-Data-Driven Machine Learning in Structural Health Diagnosis	18
<i>Yang Xu, Yuequan Bao, and Hui Li</i>	

A Numerical Study on the Early-Stage Performance of 3D Composite PLA/316L Scaffolds in Tissue Engineering	28
<i>George Drakoulas, Theodore Gortsas, Stephanos Tsinopoulos, and Demosthenes Polyzos</i>	

Ensembled Multi-classification Generative Adversarial Network for Condition Monitoring in Streaming Data with Emerging New Classes	45
<i>Yu Wang, Qingbo Wang, and Alexey Vinogradov</i>	

Ultrasonic-Based Stress Identification of a Reinforced Concrete Beam via USR-Net	58
<i>Zheng Gang, Yu Jigang, Gu Zhuangzhuang, and Zhang Benniu</i>	

Determination of the Presence or Absence of Defect for Laser Ultrasonic Visualization Testing Using Transfer Learning	80
<i>Takahiro Saitoh, Yuto Kuwabara, and Tsuyoshi Kato</i>	

Development of Nanomodified Graphene Concrete Using Machine Learning Methods	93
<i>Thusitha Ginigaddara, Thushara Jayasinghe, and Priyan Mendis</i>	

Optimization Techniques

A Multi-objective Optimization Design Framework for Thin-Walled Tubular Deployable Composite Boom for Space Applications	111
<i>Tian-Wei Liu, Jiang-Bo Bai, and Nicholas Fantuzzi</i>	

Highly Efficient Multi-resolution Topology Optimization Based on the Kriging-Interpolation Network	124
<i>Wenliang Qian and Hui Li</i>	

Study on How 3D Printer Process Parameters Can Be Optimized for Use in Orthopedics Fracture Classification and Surgical Planning	143
<i>Yosef W. Adugna, Heinrich Brüggemann, Erik Hafnor, Ane Djuv, Aksel Paulsen, Hanne R. Hagland, and Hirpa G. Lemu</i>	
Metaheuristic Optimization of Compressive Strength and Microstructures for Fused Deposition Modelling Using PLA Material	155
<i>Kedir Hussien, Hirpa G. Lemu, and Amanuel Diriba Tura</i>	
Heat Exchanger Optimization for Thermoelectric Generating Power System	172
<i>Kyeong-ho Shin, Yunho Jung, and Joo-Hyung Kim</i>	
Automatic Control	
A Physicomimetic Approach to Distributed Intelligence and Control of Autonomous Watercraft	187
<i>Nikolaos I. Xiros and Erdem Aktosun</i>	
Distributed Supervisor Control for a Large-Scale Waterway Lock System	201
<i>Fotis N. Koumboulis, Dimitrios G. Fragkoulis, Maria P. Tzamtzi, and Panayotis G. Totomis</i>	
Supervisory Control of Multiple Product Flow for a Flexible Manufacturing System	213
<i>Fotis N. Koumboulis, Dimitrios G. Fragkoulis, Maria P. Tzamtzi, and Dimitrios Papadopoulos</i>	
An Oscillator Model Applied to Power Take-Off and Tuned Control for Renewable Energy	225
<i>Nikolaos I. Xiros and Erdem Aktosun</i>	
Data-Driven Engineering	
Simulation-Based Data-Driven Wind Engineering—Analyzing the Influence of Building Proximity and Skyways on Pedestrian Comfort	241
<i>Knut Erik Teigen Giljarhus and Trond-Ola Hågbo</i>	
Anatomical Model Reconstruction (Solid Modeling) Using a Reverse Engineering Approach	254
<i>Yosef W. Adugna, Navaneethan S. Kurukkal, and Hirpa G. Lemu</i>	

Damage Assessment of Crack Growth Through Numerical Modelling and Rain Flow Counting Method	268
<i>Yohannes Regassa and Hirpa G. Lemu</i>	
Influence of Viscosity on Density-Unstable Fluid-Fluid Displacement in Inclined Eccentric Annuli	280
<i>Maryam Ghorbani, Knut Erik Teigen Giljarhus, and Hans Joakim Skadsem</i>	
Numerical Simulation of Stress in a Rotational Autofrettaged Thick Walled Tube Using the Actual Material Stress-Strain Curve	298
<i>Xiaoping Huang and Dimitrios Pavlou</i>	
Diagnostic-Prognostic Methods and Structural Integrity	
Identification of Unbalance in a Rotating System Using Artificial Neural Networks	311
<i>Ioannis Tselios and Pantelis Nikolakopoulos</i>	
A Finite Element Model of a Steel Truss Bridge Validated with Controlled Load Test Data	327
<i>Kris Hectors, Lien Saelens, Jona Bracke, Hans De Backer, and Wim De Waele</i>	
Comparative Study on Frequency-Based Cable Parameter Identification Algorithms	349
<i>Ceshi Sun, Gang Zheng, Jigang Yu, Peng Chen, and Chuan Zeng</i>	
Simulation of Photoacoustic Wave Generation and Propagation in Fluid-Solid Coupled Media Using Finite Integration Technique	359
<i>Kazuyuki Nakahata, Akihiro Miki, and Taizo Maruyama</i>	
Numerical and Analytical Methods in Engineering Mathematics	
The Singular Function Boundary Integral Method for Solving Three-Dimensional Laplacian Problems with Conical Vertex Singularities	377
<i>Hind Lamsikine, Otmane Souhar, and Georgios C. Georgiou</i>	
Analytical Solutions of Axial Annular Newtonian Flows with Dynamic Wall Slip	385
<i>Meryieme EL Farragui, Otmane Souhar, and Georgios C. Georgiou</i>	

**A Mathematical Approach for Recreation Non-symmetric 2D Railway
Alignments** 401
*Gerardo Casal, Alberte Castro, Duarte Santamarina,
and Miguel E. Vázquez-Méndez*

**Analytical Estimation of Natural Frequencies of Offshore Monopile Wind
Turbines** 421
Hadi Pezeshki, Dimitrios Pavlou, and Sudath C. Siriwardane

Author Index 437

Machine Learning



A Combined Machine Learning and Computational Methodology for Optimum Thrust Bearings' Behavior in Mixed Lubrication Regime

Konstantinos P. Katsaros and Pantelis G. Nikolakopoulos^(✉)

Machine Design Laboratory, Department of Mechanical Engineering and Aeronautics,
University of Patras, 26504 Patras, Greece
pnikolakop@upatras.gr

Abstract. A hydrodynamic thrust bearing could be forced to operate within the mixed lubrication regime under various circumstances. At this state, the tribological characteristics of the bearing could be affected significantly and the developed phenomena would have a severe impact on the performance of the mechanism. Recent technological advances, especially on the field of computer science, have provided tools that enhance and accelerate the modeling of thrust bearings' operation. The aim of this study is to combine numerical analysis and machine learning techniques in order to solve the mixed lubrication problem of a tilting pad thrust bearing providing the optimum selection of lubricant and coating for the given operating conditions. For the hydrodynamic analysis of the bearing the 2-D Reynolds equation is solved numerically with the finite difference (central differences) approach. In order to describe the roughness of the profiles, both the flow factors suggested by N. Patir and H. S. Cheng (1978) and the model of J. A. Greenwood and J. H. Tripp (1970) are taken into consideration. Two lubricants, the SAE 10W40 and the SAE 10W60, are tested and compared for a variety of operating velocities and applied coatings. All the numerical analysis data are gathered and used in order to train machine learning algorithms. Three different ML methods are applied in this investigation: Multi-Variable Quadratic Polynomial Regression, Quadratic SVM and Regression Trees. The R^2 , is calculated and used in order to determine the best fit ML method for the current study.

Keywords: Thrust bearing · Mixed lubrication · Numerical analysis · Machine learning

1 Introduction

Thrust bearings are common machine elements that were designed in order to carry the axial loads that can be developed in rotating machinery. Three main lubrication regimes can be defined based on the fraction of the minimum film thickness and the average surface roughness: the hydrodynamic, the mixed and the boundary lubrication regime. Normally, thrust bearings are designed to operate within the hydrodynamic

lubrication regime where the roughness heights of the pads and the rotor are small compared to the minimum film thickness; thus there is no contact amongst them. In this case only hydrodynamic conditions and loads affect the performance of the bearing. When the roughness of the mechanism is high enough, or when the minimum film thickness decreases at a level that a contact between asperities is possible, then the bearing operates within the mixed lubrication regime. At that point, the applied load is shared between the full-film lubrication and the asperities' contact regions. In the asperities' contact regions, only a small fraction of the total area supports the major fraction of load resulting in very high compressive stresses being induced at the contacting asperities. The prediction of the real contact area, the developed stresses at the points of contact as well as the corresponding friction and wear developed in these contact areas can be very challenging. During the years, many researches have approached the investigation of the mixed lubrication regime in thrust bearings operation both experimentally [1, 2] and analytically [3, 4]. At the same time, studies have also been carried out in order to improve the tribological performance of such mechanisms by proper lubrication characteristics analysis and design [5, 6], coatings [7, 8] as well as textured surface profiles [9, 10].

Nowadays, with the evolution in computer science, there has been a substantial increase of machine learning applications in mechanical engineering. Nevertheless, in Tribology, machine learning applications still constitute a new era of research [11], which is gradually expanding in every direction of the field. To begin with, machine learning algorithms have been applied for bearings' fault diagnosis and residual life estimation [12]. Furthermore, studies have been performed in order to optimize the performance of thrust bearings in various operational conditions by applying machine learning techniques [13, 14]. To add to that, machine learning algorithms have been used to predict mechanical properties [15] as well as friction and wear in lubricated contacts [16, 17]. Finally, machine learning techniques have been applied in order to optimize coating applications on thrust bearings as well as piston rings [18].

The aim of this study is to combine numerical analysis and machine learning techniques in order to solve the mixed lubrication problem of a tilting pad thrust bearing providing the optimum selection of lubricant and coating for the given operating conditions. Both N. Patir and H. S. Cheng (1978) as well as J. A. Greenwood and J. H. Tripp (1970) models are taken into consideration in order to describe the effect of roughness on the lubrication of the pad rotor conjunction. Two lubricants, SAE 10W40 and SAE 10W60, are tested and compared for a variety of operating conditions. All the numerical analysis data are gathered and used in order to train machine learning algorithms. Three different ML methods are applied in this investigation: Multi- Variable Quadratic Polynomial Regression, Quadratic SVM and Regression Trees. The R^2 , is calculated and used in order to determine the best fit ML method for the current study. The combination of numerical and machine learning techniques to solve the mixed lubrication problem of thrust bearings has great novelty value since no similar studies can be found in literature review.

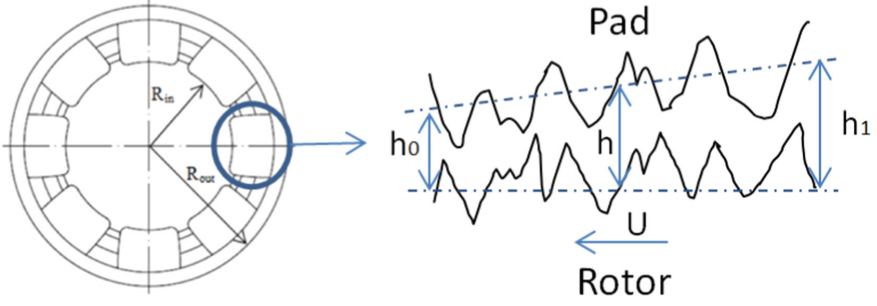


Fig. 1. Pad thrust bearing schematic with roughness

2 Theory

2.1 Lubrication Model

A single pad from the thrust bearing, shown in Fig. 1, is simulated as a center pivoted rectangle. The lubricant's flow inside the pad-rotor conjunction is assumed to be laminar and isothermal with the minimum pressure not reaching the vapor pressure; thus, cavitation effects are not taken into consideration. Furthermore, the lubricant is assumed to be Newtonian and incompressible. The no-slip condition is applied as the lubricant's flow is simulated with the use of the rotor's velocity on the pad-rotor interface [19].

The calculation of the developed pressure distribution is based on the 2-D Reynolds equation. While moving from the hydrodynamic to the mixed lubrication regime the roughness of the surfaces has a considerable impact on the flow of the lubricant as well as on the generated pad's pressure. In order to describe the resistance effect of contact areas on the flow of the lubricant, the transversely oriented surface pattern ($\gamma < 1$) from N. Patir and H.S. Cheng [20] is taken into consideration. As a result the 2-D Reynolds equation is transformed (1) in order to include the pressure flow factors Φ_x (2) and Φ_y (3). Moreover, the film thickness is assumed to be a function of the pad's length direction and is calculated from Eq. (4).

$$\frac{\partial}{\partial x} \left(\Phi_x h^3 \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left(\Phi_y h^3 \frac{\partial p}{\partial y} \right) = 6\mu U \frac{\partial h}{\partial x} \quad (1)$$

$$\Phi_x = 1 - CE^{-rH} \text{ for } \gamma \leq 1 \quad (2)$$

$$\Phi_y(H, \gamma) = \Phi_x \left(H, 1/\gamma \right) \quad (3)$$

$$h = f(x) = h_1 + \frac{x}{B} (h_1 - h_0) \quad (4)$$

The total pad's load carrying capacity is considered to be the sum of the hydrodynamic load of the lubricant and the contact load of the asperities (5). The hydrodynamic load is calculated by integrating the pressure over the pad's surface (6) while the contact load is calculated by Eq. (7) using the stochastic model of Greenwood and Tripp [21].

The $F_{5/2}$ statistical function (8) used for the calculation of the boundary load is considered to be a 5th order polynomial curve as described in the investigation of Teodorescu et al. [22].

$$W_{tot} = W_h + W_c \quad (5)$$

$$W_h = \int_A p dA \quad (6)$$

$$W_c = \frac{16\sqrt{2}}{15} \pi (\zeta \kappa \sigma)^2 \sqrt{\frac{\sigma}{\kappa}} E^* A F_{5/2}(\lambda) \quad (7)$$

$$F_{5/2}(\lambda) = -0.0046\lambda^5 + 0.0574\lambda^4 - 0.2958\lambda^3 + 0.7844\lambda^2 - 1.077\lambda + 0.6167 \quad (8)$$

The terms $\zeta \kappa \sigma$ and σ/κ are parameters related to the roughness of the profiles and calculated according to the study of Arcoumanis et al. [23]. The symbol ζ is used to describe the surface density of the asperity peaks while κ is the average asperity radius. The lambda ratio (9) is the well-known parameter that distinguishes the lubrication regime. In addition, the combined standard deviation of roughness for the rotor-pad conjunction is calculated by Eq. (10).

$$\lambda = h/\sigma \quad (9)$$

$$\sigma = \sqrt{\sigma_p^2 + \sigma_r^2} \quad (10)$$

Similarly, the total friction force (11) developed inside the pad-rotor conjunction is assumed to be the sum of the viscous friction (12) and the boundary friction (13). Equation (14) is used in order to calculate the area of direct asperities' contact. The statistical Gaussian distribution of the asperities F_2 (15) is also assumed to be a 5th order polynomial according to Teodorescu et al. [22]. In addition the corresponding coefficient of friction is calculated by Eq. (16).

$$f_{tot} = f_v + f_b \quad (11)$$

$$f_v = \iint_A \tau dA \quad (12)$$

$$f_b = \tau_0 A_c + \mu_{asp} W_c \quad (13)$$

$$A_c = \pi^2 (\zeta \beta \sigma)^2 A F_2(\lambda) \quad (14)$$

$$F_2(\lambda) = -0.0018\lambda^5 + 0.0281\lambda^4 - 0.1728\lambda^3 + 0.5258\lambda^2 - 0.8043\lambda + 0.5003 \quad (15)$$

$$\mu = \frac{f_{tot}}{W_{tot}} \quad (16)$$

Finally, assuming a minimum film thickness for the simulation, the equilibrium position of the pad is defined by applying a Newton-Raphson iterative procedure to calculate the corresponding inclination value K (17) for which the external load equals to the sum of the hydrodynamic and contact pas load.

$$K = \frac{h_1 - h_0}{h_0} \quad (17)$$

2.2 Viscosity Model

The viscosity of the lubricant is assumed to be constant during the simulation. The concept of the “effective temperature”, as described from Tala-Ighil and Fillon [24], is used in order to approximate the pad’s operating temperature. The Eqs. (18) and (19) are implemented in order to calculate the effective temperature taking into consideration the developed friction force, the rotating velocity, the flow of the lubricant as well as the corresponding lubricant’s properties.

$$T = T_0 + k_e \Delta T \quad (18)$$

$$\Delta T = \frac{2FU}{\frac{l_m}{L} Q \rho \sigma} \quad (19)$$

The Sutherland’s law is then applied in order to model the variation of the oil’s viscosity according to temperature (20), (21). Specific coefficients are calculated as the model is adapted to fit the known dynamic viscosity values for each lubricant.

$$\mu = \mu_v e^A \quad (20)$$

$$A = C_2^\mu \left(\frac{1}{T} + \frac{1}{C_1^\mu} \right) + C_3^\mu \left(\frac{1}{T} + \frac{1}{C_1^\mu} \right) \quad (21)$$

2.3 Numerical Analysis

A grid of 2,500 finite cells is used in order to discretize the fluid control volume. The finite difference (central differences) methodology is applied in order to solve the Reynolds equation at each node and calculate the corresponding nodal pressure. On the pad ambient a constant pressure of p_{am} is applied as a boundary condition. To add to that, no inflow is taken into consideration for the computational domain. The rotor is assumed to be moving with a constant rotational velocity ω , leading to the application of a constant velocity U (22) over the computational grid. Convergence to steady state condition is verified by monitoring the computed nodal pressure. The pressure convergence criterion (23) is applied. Finally, spatial resolution tests and validation of the numerical algorithm are performed with the experimental data for the work of Bielec and Leopard [25]. The

results show a difference of 1% between typical and fine meshes as well as a good agreement between simulation and experimental pad specific load values.

$$U = \omega \times R_{mean} \quad (22)$$

$$Err_{press} = \frac{\sum_1^N |P_i^j - P_{i-1}^j|}{\sum_1^N |P_i^j|} \leq 1 \times 10^{-6} \quad (23)$$

2.4 Machine Learning Techniques

For the purpose of this study, a Multi-Variable Quadratic Polynomial Regression model is developed and applied on the generated dataset. The model is based on the least squares fit methodology, in which the sum of the squares of the residuals needs to be minimized. The model uses two independent predictors: x_{1i} for the sliding velocity and x_{2i} for the coating's young modulus, in order to predict the response variable y_i that represent the coefficient of friction. The model for a set of n -observations is summarized in Eq. (24).

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & x_{11}^2 & x_{11}x_{21} & x_{21}^2 \\ 1 & x_{12} & x_{22} & x_{12}^2 & x_{12}x_{22} & x_{22}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} & x_{1n}^2 & x_{1n}x_{2n} & x_{2n}^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} \quad (24)$$

For comparison purposes, the matlab's Regression Learner tool is used in order to train Quadratic Support Vector Machine and Fine Trees Regression models. The coefficient of determination R^2 (25) is used in order to define the model with the best fit. Finally, the Matlab's standard 5-fold, cross validation procedure was applied for 5 randomly chosen partitions of the original data set. All the models were trained with 80% of the data, while the rest 20% was used for testing. The validation of the models was based on experimental data as shown in [14].

$$R^2 = 1 - \frac{\sum_1^n (y_{num} - \widehat{y}_{fit})^2}{\sum_1^n (y_{num} - \bar{y})^2} \quad (25)$$

3 Results and Discussion

Numerical analysis simulations were performed for sliding velocities of 0.5 m/s to 5 m/s and external applied load on the pad 8450N. Two lubricants were used for the simulations: SAE 10W40 and SAE 10W60. In addition, all simulations were performed for a steel shaft, a plain steel pad as well as 4 different coated pads with TiN, CrN, TiAlN and DLC. The flow factors' parameters were defined from [20], for a transversely oriented surface with $\gamma = 1/3$ and $\lambda = 1.94-1.63 > 1$. All the input parameters of the investigation are summarized in Table 1.

Table 1. Numerical analysis input parameters

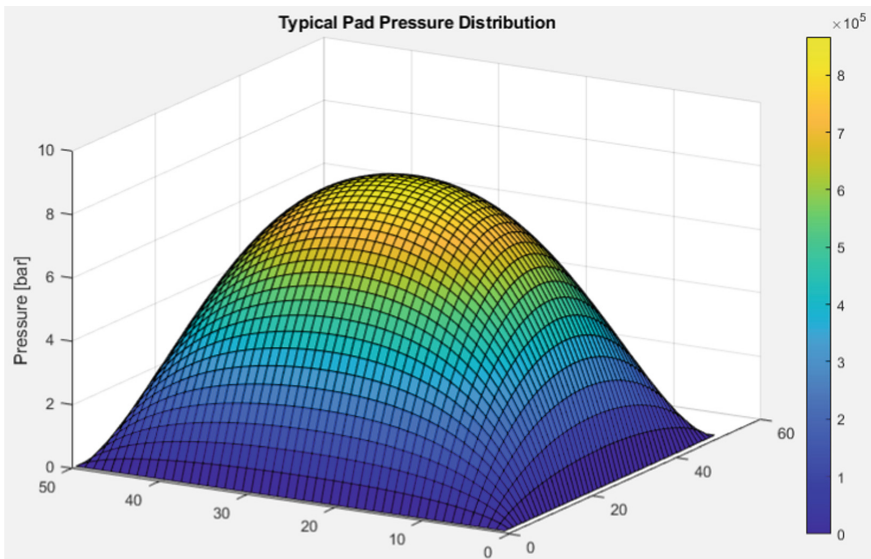
Input parameter	Value	Unit
Pad's length	32	mm
Pad's width	28	mm
Pad's outer radius	62	mm
SAE 10W40 dynamic viscosity @50 °C	0.054	Pasec
SAE 10W40 density @40 °C	917	Kg/m ³
SAE 10W40 specific heat capacity	1980	J/kgK
SAE 10W60 dynamic viscosity @50 °C	0.088	Pasec
SAE 10W60 density @40 °C	841	Kg/m ³
SAE 10W60 specific heat capacity	1901	J/kgK
Lubricant's inlet temperature	323	K
Young's modulus of elasticity steel	200	GPa
Poisson's ratio steel	0.31	–
Surface roughness steel	0.4	μm
Roughness parameter ($\zeta\kappa\sigma$) steel	0.074	–
Asperity slope (σ/κ) Steel	0.0105	–
Young's modulus of elasticity CrN	400	GPa
Poisson's ratio CrN	0.20	–
Surface roughness CrN	0.25	μm
Roughness parameter ($\zeta\kappa\sigma$) CrN	0.065	–
Asperity slope (σ/κ) CrN	0.0101	–
Young's modulus of elasticity TiN	250	GPa
Poisson's ratio TiN	0.25	–
Surface roughness TiN	0.31	μm
Roughness parameter ($\zeta\kappa\sigma$) TiN	0.064	–
Asperity slope (σ/κ) TiN	0.0172	–
Young's modulus of elasticity TiAlN	291	GPa
Poisson's ratio TiAlN	0.22	–
Surface roughness TiAlN	0.059	μm
Roughness parameter ($\zeta\kappa\sigma$) TiAlN	0.053	–
Asperity slope (σ/κ) TiAlN	0.0161	–
Young's modulus of elasticity DLC	300	GPa
Poisson's ratio DLC	0.25	–

(continued)

Table 1. (continued)

Input parameter	Value	Unit
Surface roughness DLC	0.033	μm
Roughness parameter ($\zeta\kappa\sigma$) DLC	0.052	–
Asperity slope (σ/κ) DLC	0.0132	–
Steel shaft's surface roughness	1.7	μm
C_{Φ_x}	1.18	
C_{Φ_y}	0.225	
r_{Φ_x}	0.42	
r_{Φ_y}	1.5	

A typical pressure distribution from the numerical analysis of the pad bearing is shown in Fig. 2. Moreover, the variation of the pressure flow factors Φ_x and Φ_y according to λ is presented in Fig. 3. It is clear that while λ decreases, so does the pressure flow factor Φ_x . The results agree with the ones shown in the paper of Patir and Cheng [20] as well as with the fact that the asperities start interacting with each other, as λ decreases from 1.94 to 1.63, creating obstructions to the normal flow and resulting in drop of the pressure flow.

**Fig. 2.** Typical pad pressure distribution

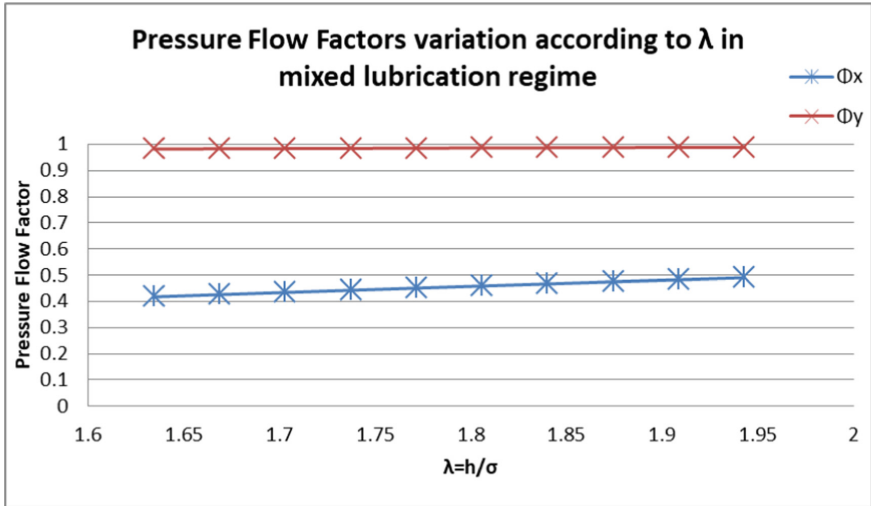


Fig. 3. Variation of the pressure flow factors Φ_x and Φ_y according to λ

The first ML models that were created from the numerical analysis data were the Multi- Variable Quadratic Polynomial Regression models for the two lubricants SAE 10W40 and SAE 10W60 shown in Table 2.

Table 2. Multi-variable quadratic polynomial regression models

Case study	ML model	R^2
SAE 10W40	$y = 0.466 - 0.321 \times 10^{-1}x_1 - 0.210 \times 10^{-2}x_2 + 0.161 \times 10^{-2}x_1^2 + 0.138 \times 10^{-4}x_1x_2 - 0.326 \times 10^{-5}x_2^2$	0.87
SAE 10W60	$y = 0.465 - 0.319 \times 10^{-1}x_1 - 0.209 \times 10^{-2}x_2 + 0.165 \times 10^{-2}x_1^2 + 0.141 \times 10^{-4}x_1x_2 - 0.324 \times 10^{-5}x_2^2$	0.87

Figure 4 is the graphical representation of the Quadratic Polynomial Regression models. It is easily noted that the drop of the sliding velocity leads to rise of the coefficient of friction in both cases. It is true that as the velocity drops so does the minimum film thickness and the amount of lubricant flowing inside the mechanism. As a result, the hydrodynamic characteristics of the bearing and the viscous friction do not have a significant impact on the final performance. On the other hand, the contact of the two surfaces that come closer as the minimum film thickness wanes, gain a leading role on the total amount of friction that is developed within the conjunction. At this point, the material of the surfaces that come in contact has a significant impact on the coefficient of friction. From Fig. 2, it can be seen that rough materials like steel and CrN show larger coefficient of friction values, while the smoother materials do help the mechanism operate with lower friction. In all studies cases the DLC coating shows up

to 45% better performance in terms of low coefficient of friction values. At the same time the two lubricants, although with a notable difference in viscosity, show similar behavior, due to the fact that the viscous forces do not affect the performance of the bearing significantly, as λ drops from 1.94 to 1.63.

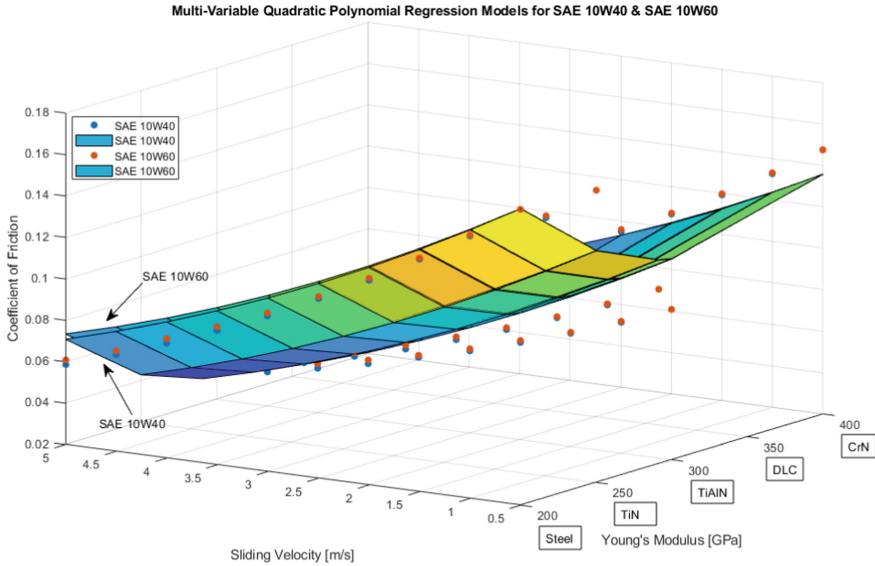


Fig. 4. Multi-variable quadratic polynomial regression models for SAE 10W40 and SAE 10W60

For comparison purposes the same data sets were used in Matlab's Regression Learner application in order to train Quadratic SVM and Fine Tree Regression models. The results of these models are summarized in Table 3 below:

Table 3. Trained machine learning models' R^2 results

Model	R^2
Quadratic SVM SAE 10W40	0.82
Quadratic SVM SAE 10W60	0.85
Fine Tree SAE 10W40	0.61
Fine Tree SAE 10W60	0.56

The coefficient of determination R^2 is the value that shows how good the fit of the models to the numerical data is. The Fine Tree models for both lubricants show the lowest R^2 values and do not have a good fit to the data sets. On the other hand, both the quadratic SVM and the quadratic Polynomial regression models show better fit to the numerical data. The Quadratic Polynomial Regression models with R^2 equal to 0.87 show the best

fit of all. Typical response plots for the quadratic SVM model of SAE10W40 are shown in Figs. 5 and 6.

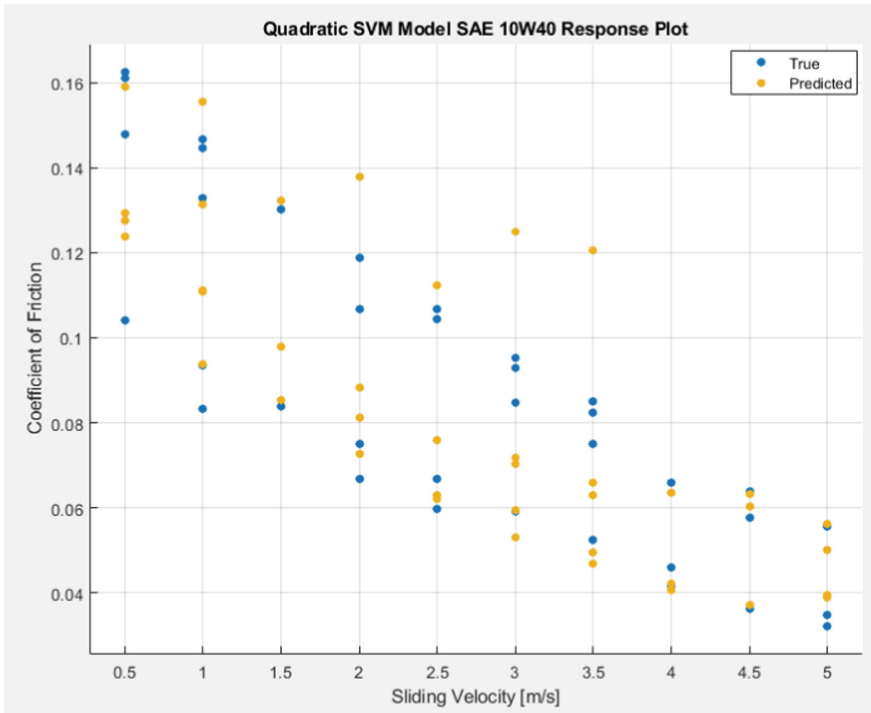


Fig. 5. Quadratic SVM model SAE 10W40 response plot for sliding velocity

A comparison between the Fine Tree and the Quadratic SVM responses is shown in Fig. 7. It is clear that the numerical data (observations) in the case of the SVM model are gathered with a higher density around the perfect prediction line when compared to the Fine Tree model.

Finally, a typical Fine Tree model for SAE 10W40 is shown in Fig. 8.

4 Conclusions

To sum up, the modified 2-D Reynolds equation was used in order to calculate the tribological characteristics of a pivoted pad bearing for different sliding velocities, two multi grade lubricants and four coatings. The Patir and Cheng model of flow factors was applied in order to describe the resistance of a transversely oriented rough surface to the main flow. At the same time, the Greenwood and Tripp model was used in order to calculate the contact loads that were developed within the pad-rotor conjunction while the bearing was simulated to operate in mixed lubrication regime (with λ between 1.94 and 1.63). All the numerical analysis data were used as input to train three different

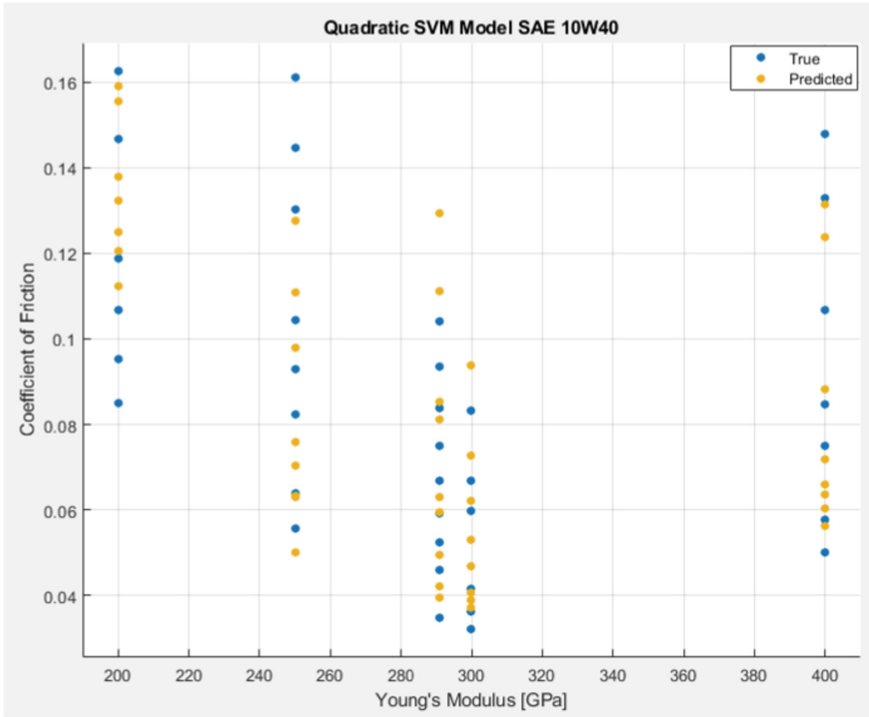


Fig. 6. Quadratic SVM model SAE 10W40 response plot for young's modulus

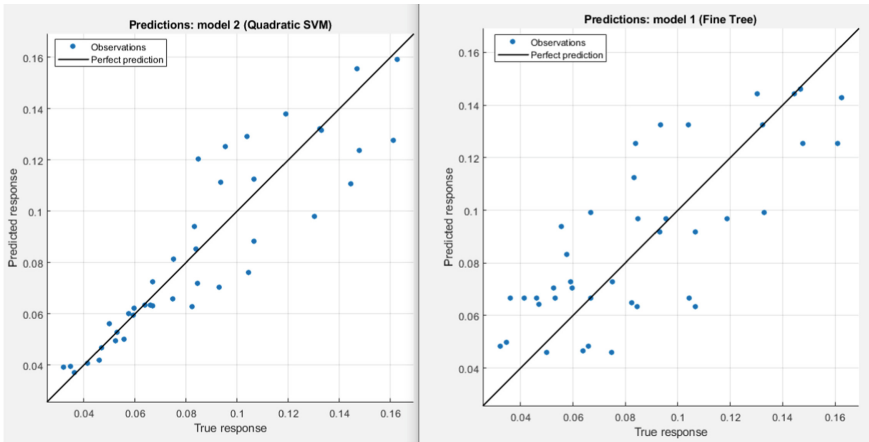


Fig. 7. Predictions to true responses plots-comparison between SVM and fine tree models

machine learning models: a Quadratic Polynomial Regression model, a Quadratic SVM Regression model and a Fine Tree Regression model. The main conclusions that derive from the investigation are:

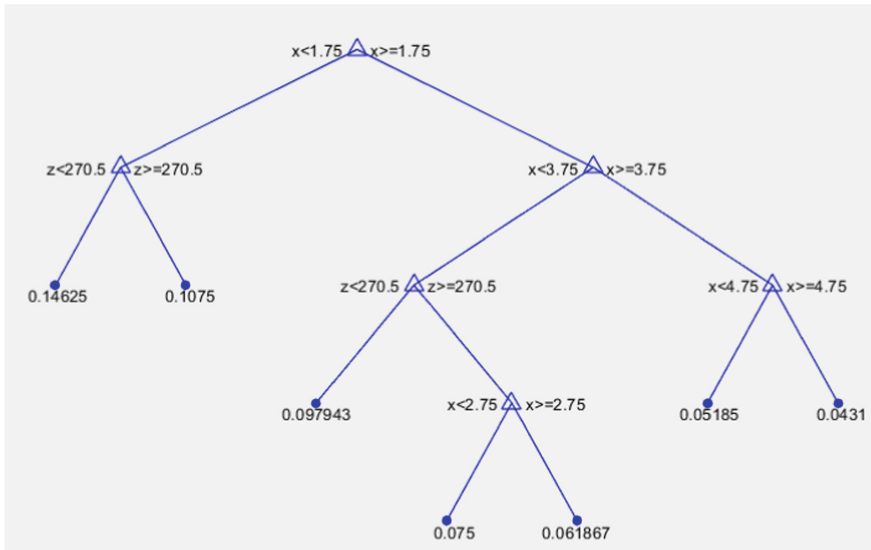


Fig. 8. Typical fine tree ML model for SAE 10W40

- Both lubricants show similar coefficient of friction in mixed lubrication regime because as the speed lowers the viscous loads tend to have low impact on the final performance of the bearing while the contact forces play the major role.
- The DLC coating shows up to 45% less coefficient of friction compared to the rest of the coated and uncoated surfaces.
- The Fine Tree Regression model has the worst fit to the data set.
- The Quadratic Polynomial Regression model has the best fit of all with an R^2 equal to 0.87.

All in all, SAE 10W40 and the DLC coating constitute the best choices from the ones examined for the operation of the bearing within the mixed lubrication regime. To add to that the Quadratic Polynomial Regression model is the best fit for the current investigation.

References

1. Wang, X., Kato, K., Adachi, K., Aizawa, K.: Loads carrying capacity map for the surface texture design of SiC thrust bearing sliding in water. *Tribology International*, 36(3), 189–197, (2003).
2. Bian, Z., McClean, S., Delado, A., Patil, A.: Development of A High-Load Capacity Test Rig to Evaluate the Static Performance of Process Fluid-Lubricated Thrust Bearings. *Journal of Physics: Conference Series* 1909, IOP Publishing (2021).
3. Wang, Y., Wang, Q.J., Lin, C.: Mixed Lubrication of Coupled Journal-Thrust Bearing Systems. *CMES*, vol.3, no4, pp. 517-530, (2002).
4. Zhang, X., Xu, Y., Jackson, R.: A mixed lubrication analysis of a thrust bearing with fractal rough surfaces. *Proceeding of the Institution of Mechanical Engineers Part J, Journal of Engineering Tribology*, vol. 208–210, IMechE (2019).

5. Dadouche, A., Fillon, M., Dmochowski, W.: Performance of a hydrodynamic fixed geometry thrust bearing: comparison between experimental data and numerical results, *Tribol Transactions*, vol.49 (3), pp.419-426, (2006).
6. Anandan, N., Sathianmurugan, P., Mathiazhagan, P.: Load carrying capacity of adiabatically lubricated thrust bearings with various film profiles, *Universal Journal of Mechanical Engineering*, 2(3):94-100, (2014).
7. Ricci, R., Chatterton, S., Pennacchi, P., Vania, A.: Multiphysics Modeling of a Tilting Pad Thrust Bearing: Comparison Between White Metal and Polymeric Layered Pads, *Proceedings of the ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Society of Mechanical Engineers: New York, NY, USA, (2011).
8. Katsaros, K., Bompos, D.A., Nikolakopoulos, P.G., Theodosiades, S.: Thermal-Hydrodynamic Behaviour of Coated Pivoted Pad Thrust Bearings: Comparison between Babbitt, PTFE and DLC, *Lubricants*, 6, 50, (2018).
9. Papadopoulos, C.I., Kaiktsis, L., Fillon, M.: CFD Thermohydrodynamic analysis of 3-D sector-pad thrust bearings with rectangular dimples. *Proceedings of ASME Turbo. Expo 2013: Turbine Technical Conference and Exposition*, GT2013; V07BT30A002; 10 p, June 3-7, San Antonio (2013).
10. Gherca, A., Fatu, A., Hajjam, M., & Maspeyrot, P.: Influence of surface texturing on the hydrodynamic performance of a thrust bearing operating in steady-state and transient lubrication regime, *Tribology International*, 102, 305-318, (2016).
11. Marian, M., Tremmel, S.: Current Trends and Applications of Machine Learning in Tribology-A Review, *Lubricants*, 9,86, (2021).
12. Soualhi, A., Medjaher, K., Zerhouni, N.: Bearing health monitoring based on Hilbert-Huang transform, support vector machine, and regression, *IEEE Transactions on Instrumentation and Measurement*, 64(1), 52-62, (2015).
13. Lostado-Lorza, R., Escribano-García, R., Fernández-Martínez, R., Illera-Cueva, M., Donald, B.J.M.: Combination of the finite element method and data mining techniques to design and optimize bearings, *International Joint Conference SOCO'14-CISIS'14-ICEUTE'14, Advances in Intelligent Systems and Computing*, Vol 299, 165-174, Cham: Springer (2014).
14. Katsaros, K., Nikolakopoulos, P.G.: On the tilting-pad thrust bearings hydrodynamic lubrication under combined numerical and machine learning techniques, *Lubrication Science*, 1-18, (2021).
15. Moon, I.Y., Jeong, H.W., Lee, H.W., Kim, S.J., Oh, Y.S., Jung, J., Oh, S., Kang, S.H.: Predicting High Temperature Flow Stress of Nickel Alloy A230 Based on an Artificial Neural Network, *Metals*, 12, 223, (2022).
16. Bhaumik, S., Pathak, S.D., Dey, S., Datta, S.: Artificial intelligence based design of multiple friction modifiers dispersed castor oil and evaluating its tribological properties, *Tribology International*, 140, (2019).
17. Saridakis, K.M., Nikolakopoulos, P.G., Papadopoulos, C.A., Dentsoras, A.J.: Identification of wear and misalignment on journal bearings using artificial neural networks. *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, 226, 46-56 (2012).
18. Zavos, A., Katsaros, K.P., Nikolakopoulos, P.G.: Optimum Selection of Coated Piston Rings and Thrust Bearings in Mixed Lubrication for Different Lubricants Using Machine Learning, *Coatings*, 12, 704, (2022).
19. Aurelian, F., Patrick, M., Mohamed, H.: Wall slip effects in (elasto) hydrodynamic journal bearings. *Tribology International*, 44 (7-8), 868-877, (2011).
20. Patir, N., Cheng, H.S.: An Average Flow Model for Determining Effects of Three-Dimensional Roughness on Partial Hydrodynamic Lubrication, *Transactions of the ASME*, 12/ vol. 100, (1978).

21. Greenwood, J. A., Tripp, J. H.: The Contact of Two Nominally Flat Rough Surfaces. *Proceedings of the Institution of Mechanical Engineers*, 185(1), 625–633, (1970).
22. Teodorescu, M., Balakrishnan, S., Rahnejat, H.: Integrated tribological analysis within a multi-physics approach to system dynamics, *Tribology and Interface Engineering Series* 48, 725–737, (2005).
23. Arcoumanis, C., Ostovar, P., Mortimer, R.: Mixed Lubrication Modeling of Newtonian and Shear Thinning Liquids in a Piston-Ring Configuration, *Society of Automotive Engineers*: paper No. 972924, Warrendale, PA, USA, (1997).
24. Tala-Ighil, N., Fillon, M.: A numerical investigation of both thermal and texturing surface effects on the journal bearings static characteristics, *Tribology International*, 90, 228–239, (2015).
25. Bielec, M.K., Leopard, A.J.: Tilting Pad Thrust Bearings: Factors Affecting Performance and Improvements with Directed Lubrication, *Proceedings of the Institution of Mechanical Engineers, Conference Proceedings*, 184(12), 93–102, (1969).



Limited-Data-Driven Machine Learning in Structural Health Diagnosis

Yang Xu^{1,2,3} , Yuequan Bao^{1,2,3} , and Hui Li^{1,2,3} 

¹ Key Lab of Smart Prevention and Mitigation of Civil Engineering Disasters of the Ministry of Industry and Information Technology, Harbin 150090, China

xyce@hit.edu.cn

² Key Lab of Structures Dynamics Behavior and Control of the Ministry of Education, Harbin 150090, China

³ Harbin Institute of Technology, Harbin 150090, China

Abstract. Currently, structural health diagnosis has been extensively investigated following a data-driven paradigm with advanced deep learning and computer vision techniques. However, the identification accuracy and generalization ability of data-driven models highly rely on the quality and diversity of the collected data. In contrast, data with specific patterns and concerned characteristics are always in small quality and diversity under real-world scenarios, causing the problem of data incompleteness. This paper established a framework for structural health diagnosis under limited supervision following data, model, and algorithm perspectives to fix the above issue. Firstly, a data augmentation process of random elastic deformation was designed to enrich the feature space using a few structural damage images. Secondly, a novel neural network model was constructed to enhance the nonlinear expression power, feature extraction ability, and recognition accuracy by introducing the subnet inside a single neuron and self-attention module. Thirdly, a task-significance-aware meta-learning optimization algorithm was proposed to learn across various tasks and enhance the generalization ability for structural damage identification. Finally, an unsupervised deep learning method for structural condition assessment was proposed to mine the shared latent space between the source and target domains based on intra- and inter-class probabilistic correlations of quasi-static responses. Real-world applications, including tiny fatigue crack segmentation in steel box girders, multitype structural damage identification for bridge inspection, and condition assessment for long-span cable-stayed bridges, were successively performed to demonstrate the effectiveness of the proposed framework for structural health diagnosis under limited supervision.

Keywords: Structural health monitoring · Machine learning · Computer vision

1 Introduction

For past decades, one of the most commonly used ways for structural health diagnosis was manual inspection, which had the shortcomings of high dependence on subjective judgment and engineering experience, severe unreliability, and low efficiency. Following the paradigm of damage prognosis established by Farrar and Lieven (2007) [1],

structural damage recognition, condition assessment, and reliability evaluation were the most significant issues towards structural health diagnosis. Since the 1990s, structural health monitoring techniques have been widely adopted in large-scale infrastructure using non-destructive testing and vibration-based methods. The measured signals were directly compared with peak values or statistical indices with thresholds regulated by design codes. However, the following challenges remained to be addressed: these techniques required the dense deployment of sensors on bridges and faced the ill-posedness of the reverse problem; the modal parameters were insensitive to minor damage in a local position; the accuracy was influenced by temperature and noise.

With the successful development of artificial intelligence, data-driven methods have been developed for damage detection and condition assessment based on machine learning, deep learning, and computer vision algorithms [2–4]. Recently, vision-based damage detection has been elaborately investigated using image processing techniques [5, 6]. Generally, these methods mainly utilized close-up imaging of structures and only focused on a small area of local damage regions. Moreover, model performances heavily relied on the optimal selection of handcrafted features and critical parameters, thus lacking accuracy and robustness facing large-scale images with complex backgrounds under real-world scenarios [7]. For deep learning-based methods, they were always performed by directly migrating the well-trained model to newly-collected onsite images, thus requiring a massive dataset for training and a large volume of model parameters to ensure the recognition accuracy and robustness under various scenarios. Additionally, the recognition stability on multiscale damages with different morphologies remained challenging.

To address these issues mentioned above, this study established a framework for structural health diagnosis under limited supervision from vision-based damage recognition and deep learning-based condition assessment in Sects. 2 and 4, respectively. Section 5 concluded this paper.

2 Vision-Based Damage Recognition Using Few Images

Although many investigations have been performed for damage recognition from images, the following issues remain to be addressed: (1) the accuracy heavily relies on sufficient images and large network parameters; (2) the sensitivity to minor damage in local positions is limited; (3) the robustness is inadequate on complex coupled damage with various morphological features and disturbances. In this section, a series of recent advances are reported to solve the above issues.

A random elastic deformation (RED) algorithm was proposed to enrich the diversity of damage morphology with only a handful of original images [8], as shown in Fig. 1. Firstly, control nodes (red dots) were equidistantly set on mesh grids of the original image. Random offsets of $(\Delta x, \Delta y)$ were assigned to these control nodes following a uniform distribution (blue arrows). Secondly, offsets of other pixels were calculated using two-dimensional cubic spline interpolation. Thirdly, the pixel value of the sub-pixel was determined by bilinear interpolation. The results indicated that RED could increase the geometric shapes and local microstructures of cracks and add high-order components into the original crack shapes. Therefore, new crack images generated by

RED have significant differences from the original ones, demonstrating that RED could enrich the feature space of structural damage images.

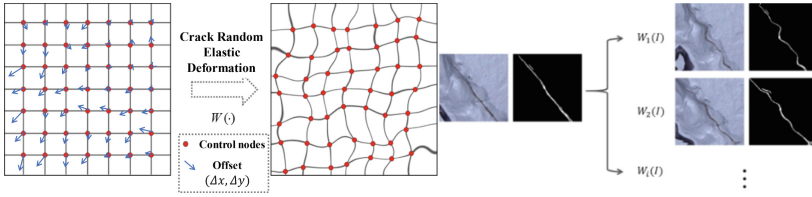


Fig. 1. Random elastic deformation for image augmentation of structural damage

A novel Self-Attention-Self-Adaption (SASA) neuron computing model [8] was proposed to enhance the capability of feature extraction and nonlinear expression power for neural networks:

$$x_j^{l+1} = \sigma \left[\sum_i^N w_{ij}^{l,l+1} x_i^l + b_j^{l+1} + m_j^{l+1} (\alpha^l * X^l, \theta_j^{l+1}) \right]$$

$$\alpha^l = G(\text{softmax}(X^l), \beta), \text{softmax}(X^l)_i = e^{x_i} / \sum_j^{N^l} e^{x_j} \quad (1)$$

where x_i^l denotes the i th neuron in the l th layer, N^l denotes the number of neurons in the l th layer, $w_{ij}^{l,l+1}$ denotes the connecting weight between x_i^l and x_j^{l+1} , b_j^{l+1} denotes the individual bias associated with x_j^{l+1} , and σ denotes the nonlinear activation function. X^l denotes the neurons in the l th layer, α^l denotes the significance vector in the Self-Attention module, and m_j^{l+1} denotes the subnet of multilayer perceptron in the Self-Adaption module associated with the j th neuron in the $(l+1)$ layer and parameterized with θ_j^{l+1} . $*$ denotes the multiplication operator of corresponding elements for two vectors. The gate function G reserves top β elements and assigns the others to zero. θ_j^{l+1} denotes subnet parameters.

Figure 2 shows the schematic of the SASA neuron computing model. The Self-Attention module applied softmax and gate operations to obtain the significance vector. It enabled the neuron to focus on the most significant receptive fields when processing large-scale feature maps, emphasized the saliency of interior neurons inside one layer, and did not introduce additional trainable parameters. The Self-Adaption module was designed as a subnet of multilayer perceptron m_j^{l+1} and implemented by a standard neural network with k equal hidden layers. The interior subnet structure was controlled by the number of hidden layers k and the number of neurons h in each hidden layer. For consistency with the exterior neuron network, the number of neurons h in the hidden layer of the subnet was set proportionate to the number of neurons N^{l+1} in the current layer with a default coefficient γ . It could achieve powerful feature extraction using only a small number of images. The SASA neuron model allowed for the “plug and play” of arbitrary conventional neural networks.

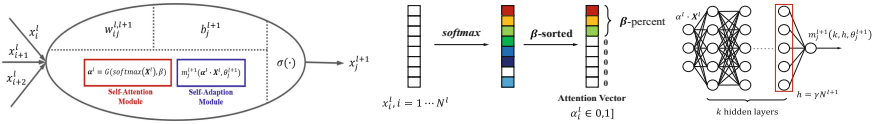


Fig. 2. Schematic of SASA neuron computing model

A case study was performed on the semantic segmentation of distributed tiny fatigue cracks in steel box girders using U-net as the baseline model. Figure 3 shows several comparative results of tiny crack segmentation with and without integrating the SASA neuron in U-net. The results indicated that using the modified U-net integrating with SASA neuron could achieve accurate pixel-level recognition of tiny cracks with complex background interferences. False alarms and crack gaps were effectually suppressed, implying that the SASA neuron could enable the model to focus on the local regions of interest. Based on the image recognition results of tiny fatigue cracks in steel box girder, a hierarchical dynamic Bayesian network was established for fatigue crack propagation modeling considering initial defects [9].

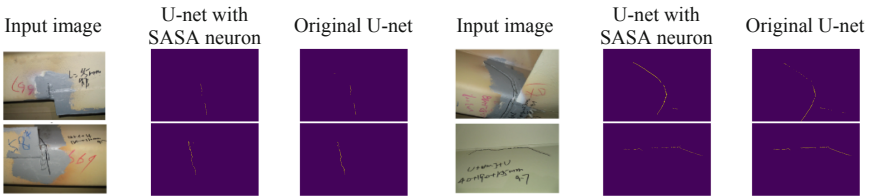


Fig. 3. Comparisons of tiny crack segmentation with/without SASA neuron in U-net

A dual-stage attribute-based few-shot meta learning paradigm was proposed for multitype structural damage identification [10], as shown in Fig. 4. An exterior few-shot meta learning framework was established based on randomly-selected tasks as meta-batches to produce robust classifiers for new damage classes. Support and query subsets comprising only partial damage categories and a few examples were randomly generated from the original image dataset. An interior attribute-based transfer learning model was trained by minimizing the l_2 -norm and angular differences of predicted and ground-truth attribute vectors. Damage attribute acts as the common inter-class knowledge and is transferred among various damage categories instead of using one-hot vector labels for the standard supervised classification. The latter only considered that the class labels were orthogonal and had no connections, and therefore only one class was assigned on the position of maximum softmax probability, causing the misrecognition of coupled damage.

For pixel-wise recognition for various structural damage, a lightweight modified DeepLabv3 + model was established as the interior model [11]. Figure 5 shows the schematic of the modified DeepLabv3 + model for semantic segmentation of multitype structural damage. The backbone network of the original ResNet101 was replaced with the lightweight MobileNetV2. Depthwise separable and dilated convolutions were

used instead of standard convolution to reduce parameter volume. A refined atrous spatial pyramid pooling module was designed following the backbone network to expand the receptive fields of multilevel feature maps using dilated convolutions with various dilation rates. Furthermore, a piecewise loss function based on Focal and Dice losses was designed for different training stages. Several representative results for semantic segmentation of concrete crack, concrete spalling, rebar exposure, and cable corrosion indicated that the established model performed well and was stable facing various structural damage. It could be inferred that the morphological feature and shape texture for various categories of structural damage were automatically captured.

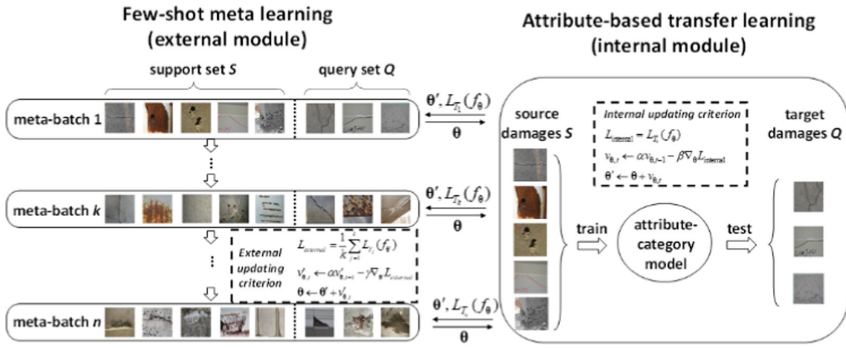


Fig. 4. Dual-stage attribute-based few-shot meta learning for multitype damage classification

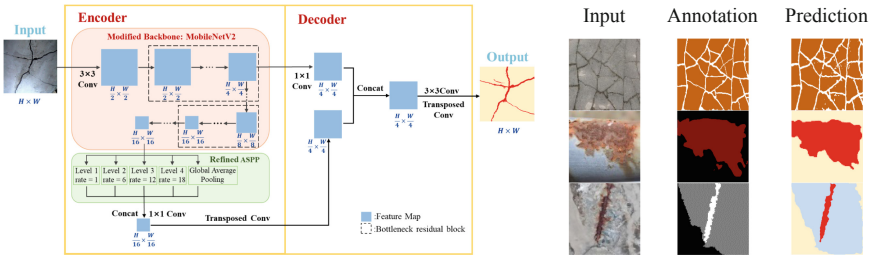


Fig. 5. Lightweight modified DeepLabv3+ model for structural damage segmentation

3 Correlation Pattern Recognition Based Condition Assessment

Considering that the correlation between quasi-static responses subjected to identical external loads is only a function of structural parameters and independent from the external loads, the correlation can therefore be employed as an indicator of the structural condition.

A bi-directional long short-term memory (BiLSTM) model was established to model the temporal correlation between the vertical deflection of girders (GVD) and tension of cables (CT) [12], as shown in Fig. 6. Test results showed that the bridge was under normal

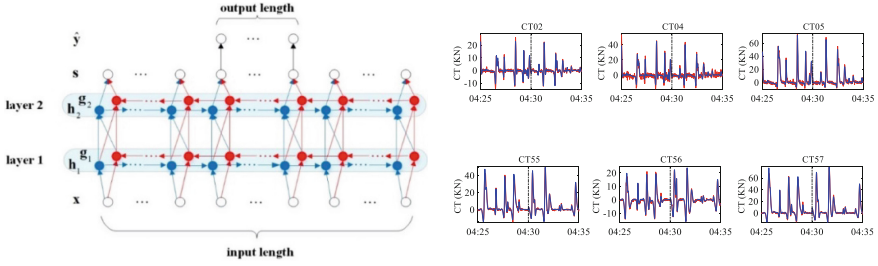


Fig. 6. Schematic of BiLSTM for temporal correlation modeling under normal conditions

conditions and that the average root mean square error (RMSE) and relative RMSE between the predicted and ground-truth CTs were 1.83 kN and 3.19%, respectively.

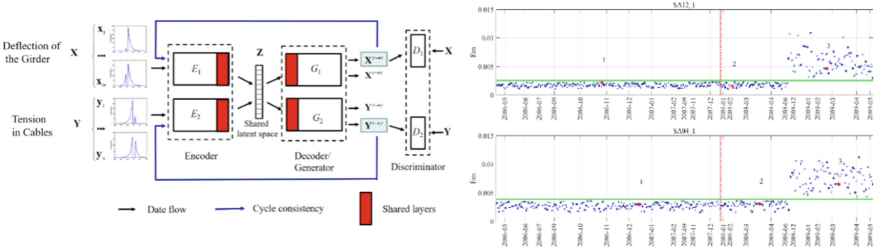


Fig. 7. Probabilistic correlation modeling and Wasserstein distance variation

A deep learning network comprising two variational autoencoders (VAEs) and two generative adversarial networks (GANs) was established to model the probabilistic correlations of quasi-static responses of bridges [13], as shown in Fig. 7. VAEs were designed to model intra-class correlations among either GVDs or CTs, and GANs were designed to model inter-class correlations between GVDs and CTs. The input and output were marginal probability density functions (PDFs) of the quasi-static responses, and they were obtained in the same time window under identical vehicle loads and structural parameters. The Wasserstein distance between the predicted and ground-truth PDFs of tension in the cables was used as an indicator of the structural condition. The results showed that the Wasserstein distance was very sensitive to damage and presented noticeable variations when the damage of the stay cable occurred.

4 Vision-Based Structural Seismic Assessment for Buildings

Recently, remote sensing satellites [14], unmanned aerial vehicles (UAVs) [15], and smartphones have been extensively utilized in non-contact post-earthquake inspection at different scales with cutting-edge computer vision and machine learning techniques. In this section, a computer-vision-based coarse-to-fine seismic assessment framework was established to localize dense buildings in urban areas, classify collapsed and non-collapsed states, recognize multi-type surface damage on structural components, and evaluate seismic performances.

A Transformer-CNN deep learning architecture was designed for semantic segmentation of dense buildings and binary classification of collapsed states using large-scale remote sensing satellite images [16]. It consisted of a Swin Transformer encoder, multi-stage feature fusion module, and UPerNet decoder to extract global correlations and local features of dense buildings synchronously, as shown in Fig. 8.

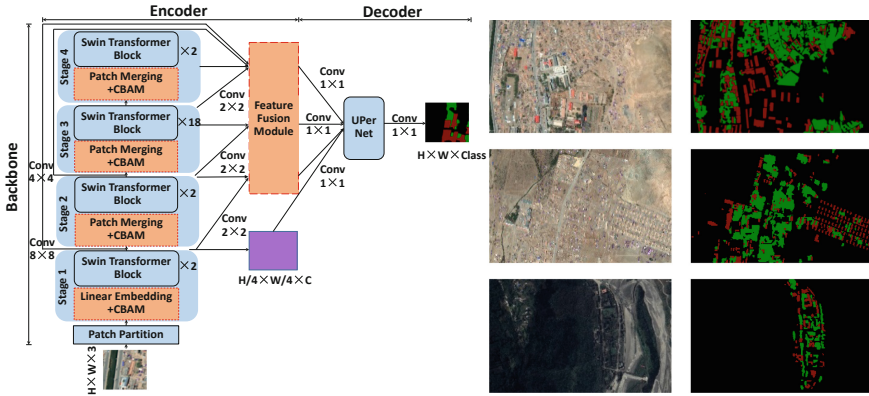


Fig. 8. Improved Swin Transformer for dense building segmentation and state classification

A multi-task learning strategy was proposed to simultaneously recognize multi-type structural components (column, beam, wall), seismic damage (concrete crack, spalling, and rebar exposure), and multi-level damage states (minor, moderate, major) using medium-scale UAV images [17]. It contained a CNN-based encoder-decoder backbone with skip-connection modules and multi-head segmentation subnetworks for different tasks, as shown in Fig. 9.

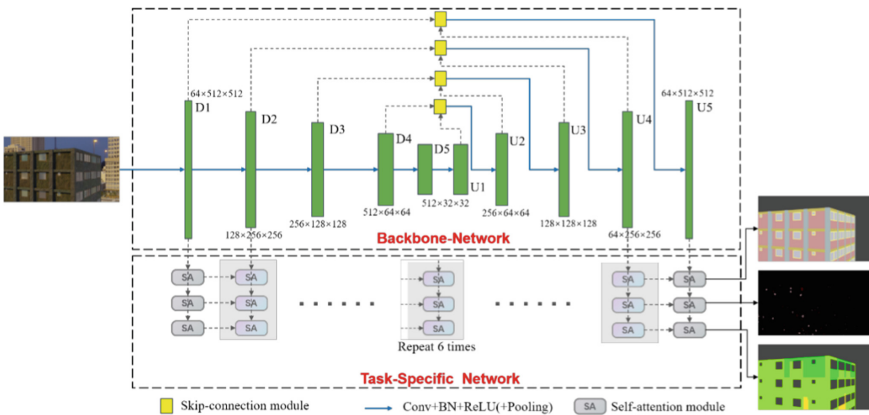
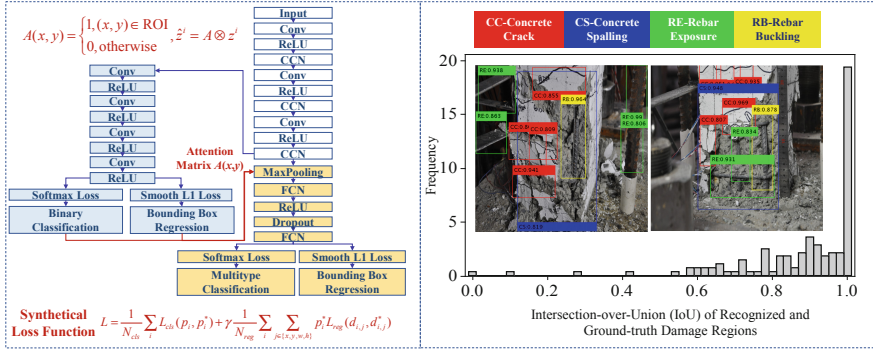
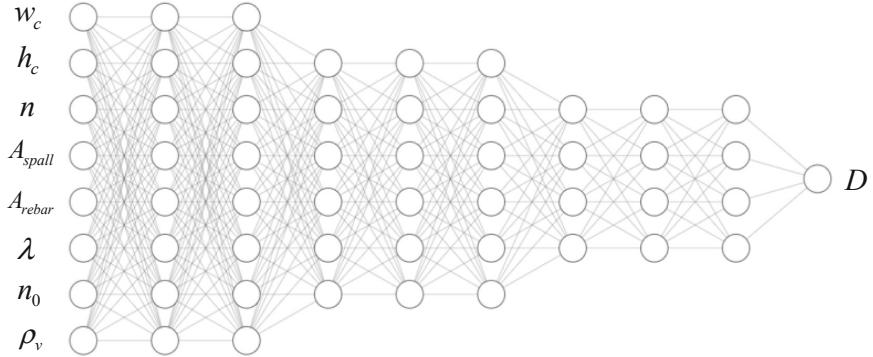


Fig. 9. Multi-task learning semantic segmentation of structural component, damage, and state

An earthquake engineering knowledge-enhanced machine learning method was established for seismic damage assessment of structural components [18], as shown in Fig. 10. A machine learning neural network was established to quantify the seismic damage index of structural components using damage-related parameters (lengths, areas, and numbers of concrete crack, spalling, and rebar exposure) and design-related parameters (axial compression ratio, shear span ratio, and volumetric stirrup ratio) as inputs. A seismic damage indicator with an explicit bound of [0,1] could be obtained to reflect the nonlinear accumulation of seismic damage.



(a) Region-based object detection for multi-type seismic damage from images



(b) Deep neural network for seismic damage index regression

Fig. 10. Seismic damage quantification using quasi-static experimental data and images

5 Conclusions

This study established a limited-data-driven machine learning framework for structural health diagnosis. The main conclusions were summarized as follows.

- (1) A data augmentation algorithm of random elastic deformation was designed to enrich the feature space using a few structural damage images. A novel neural network model was designed to enhance the nonlinear expression power, feature extraction

ability, and recognition accuracy by introducing the self-attention and subnet modules inside a standard neuron. A task-significance-aware meta-learning optimization algorithm was proposed to learn across various tasks and enhance the generalization ability for multitype structural damage identification.

- (2) Two deep learning networks were established to mine the shared latent space between the source and target domains based on intra-class and inter-class temporal and probabilistic correlations between two kinds of quasi-static responses for structural condition assessment.
- (3) A computer-vision-based coarse-to-fine seismic assessment framework was established to localize dense buildings in urban areas, classify collapsed and non-collapsed states, recognize multi-type surface damage on structural components, and evaluate seismic performances. A series of deep learning models were designed for the localization, classification, and quantification of dense buildings, deterioration states, and damage index using large-scale remote sensing satellite images, medium-scale UAV images, near-field surface images, and quasi-static experimental data.
- (4) Real-world applications, including distributed tiny fatigue crack segmentation in steel box girders, multitype structural damage classification and segmentation for bridge inspection, condition assessment for long-span cable-stayed bridges, and vision-based structural seismic assessment for buildings were performed to demonstrate the effectiveness of the proposed limited-data-driven machine learning methods for structural health diagnosis.

Acknowledgement. Financial support for this study was provided by the National Key R&D Program of China [Grant No. 2021YFF0501003], National Natural Science Foundation of China [Grant Nos. 52192661, 51921006, and 52008138], China Postdoctoral Science Foundation [Grant Nos. BX20190102 and 2019M661286], and Heilongjiang Province Postdoctoral Science Foundation [Grant Nos. LBH-TZ2016 and LBH-Z19064], Heilongjiang Province Natural Science Foundation [Grant No. LH2022E070], Fundamental Research Funds for the Central Universities [Grant No. HIT.NSRIF202334], and China University Innovation Fund - A New Generation of Information Technology Innovation Project [Grant No. 2022IT187].

References

1. Farrar, C., Lieven, N.: Damage prognosis: the future of structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365(1851), 623-632 (2007).
2. Bao, Y., Li, H.: Machine learning paradigm for structural health monitoring. *Structural Health Monitoring* 20(4), 1353-1372 (2021).
3. Bao, Y., Chen, Z., Wei, S., Tang, Z., Xu, Y., Li H.: The state of the art of data science and engineering in structural health monitoring. *Engineering* 5(2), 234-242 (2019).
4. Xu, Y., Qian, W., Li, N., Li, H.: Typical advances of artificial intelligence in civil engineering. *Advances in Structural Engineering* 25(16): 3405-3424 (2022).
5. Qiao, W., Zhao, Y., Xu, Y., Lei, Y., Wang, Y., Yu, S., Li H.: Deep learning-based pixel-level rock fragment recognition during tunnel excavation using instance segmentation model. *Tunnelling and Underground Space Technology* 115, 104072 (2021).

6. Zhao, J., Hu, F., Xu, Y., Zuo, W., Zhong, J., Li H.: Structure-PoseNet for identification of dense displacement and three-dimensional poses of structures using a monocular camera. *Computer-aided Civil and Infrastructure Engineering* 37(6), 704-725 (2021).
7. Spencer, Jr B.F., Hoskere, V., Narazaki, Y.: Advances in computer vision-based civil infrastructure inspection and monitoring. *Engineering* 5(2), 199-222 (2019).
8. Zhao, J., Hu, F., Qiao, W., Zhai, W., Xu, Y., Bao, Y., Li H.: A modified U-net for crack segmentation by self-attention-self-adaption neuron and random elastic deformation. *Smart Structures and Systems* 29(1), 1-16 (2022).
9. Xu, Y., Zhu, B., Zhang, Z., & Chen, J.: Hierarchical dynamic Bayesian network-based fatigue crack propagation modeling considering initial defects. *Sensors* 22(18), 6777 (2022).
10. Xu, Y., Bao, Y., Zhang, Y., Li H.: Attribute-based structural damage identification by few-shot meta learning with inter-class knowledge transfer. *Structural Health Monitoring* 20(4), 1494-1517 (2020).
11. Xu, Y., Fan Y., Li, H.: Lightweight semantic segmentation of complex structural damage recognition for actual bridges. *Structural Health Monitoring-An International Journal* 14759217221147015: 1-20 (2023).
12. Tian, Y., Xu, Y., Zhang, D., Li H.: Relationship modeling between vehicle-induced girder vertical deflection and cable tension by bilstm using field monitoring data of a cable-stayed bridge. *Structural Control and Health Monitoring* 28(2), e2667 (2021).
13. Xu, Y., Tian, Y., Li, H.: Unsupervised deep learning method for bridge condition assessment based on intra- and inter-class probabilistic correlations of quasi-static responses. *Structural Health Monitoring-An International Journal* 22(1): 14759217221103016 (2022).
14. Wang, Y., Cui, L., Zhang, C., Chen, W., Xu, Y., Zhang, Q.: A two-stage seismic damage assessment method for small, dense, and imbalanced buildings in remote sensing images. *Remote Sensing* 14(4), 1012 (2022).
15. Wang, Y., Jing, X., Chen, W., Li, H., Xu, Y., Zhang, Q.: Geometry-informed deep learning-based structural component segmentation of post-earthquake buildings. *Mechanical System and Signal Processing* 188, 110028 (2023).
16. Cui, L., Jing, X., Wang, Y., Huan, Y., Xu, Y., Zhang, Q.: Improved swin transformer-based semantic segmentation of postearthquake dense buildings in urban areas using remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 16, 369-385 (2022).
17. Xu, Y., Qiao, W., Zhao, J., Zhang, Q., Li, H.: Vision-based multi-level synthetical evaluation of seismic damage for RC structural components: a multi-task learning approach. *Earthquake Engineering and Engineering Vibration*. DOI: <https://doi.org/https://doi.org/10.1007/s11803-023-2153-4> (2023).
18. Xu, Y., Wei, S., Bao, Y., Li, H.: Automatic seismic damage identification of reinforced concrete columns from images by a region-based deep convolutional neural network. *Structural Control and Health Monitoring* 26(3), e2313 (2019).



A Numerical Study on the Early-Stage Performance of 3D Composite PLA/316L Scaffolds in Tissue Engineering

George Drakoulas¹, Theodore Gortsas^{1,2}, Stephanos Tsinopoulos²,
and Demosthenes Polyzos¹✉

¹ Department of Mechanical Engineering and Aeronautics, University of Patras, 26500 Patras, Greece

g.drakoulas@upnet.gr, gortsas@upatras.gr, polyzos@mech.upatras.gr

² Department of Mechanical Engineering, University of Peloponnese, 26334 Patras, Greece
stsinop@uop.gr

Abstract. In recent years, 3D-printed biocompatible and biodegradable scaffolds used for tissue engineering have attracted great attention as an effective approach for bone repair in critical-size fractures. The surface strains imposed by the loading of the scaffold, the shear wall stresses due to the flow of the interstitial fluid, the material properties, and the geometry of a fused filament fabricated scaffold as well as the final pore size, are some of the crucial parameters that significantly influence the behavior of the cells on the scaffold. In this work, the early-stage performance of a rectangular and four-layer orthogonal scaffold with isometric pores made by polylactic acid PLA/316L with 5% stainless steel particle content, is numerically investigated. The criterion of cell attachment is accomplished via a mechanoregulatory model based on the surface strains and the shear wall stresses of the scaffold imposed by a three-point bending loading and an interstitial flow, respectively. The boundary element method is employed to compute the strain distribution on the scaffold's surface and then a physics-informed neural network is trained to provide the deformations in the domain given the material properties and the boundary conditions. The element-based finite volume method is utilized to solve the Navier-Stokes equations and evaluate the shear stresses imposed by the fluid flow on the scaffold surface. The predicted surface strain and wall shear stress distribution provide useful insight into the correlation between the structure of the scaffold, the loading type, the cell viability, and differentiation.

Keywords: Fused filament fabricated scaffolds · Scaffold-aided bone healing · Mechanoregulatory computational model · Boundary element method · Physics informed neural networks

1 Introduction

In the last two decades, a plethora of computational and experimental studies dealing with the design of biocompatible 3D printing scaffolds suitable for bone repair in severe fractures, have appeared in the literature [1–4]. However, despite the intensive work

a clear answer on which is the proper scaffold's structure that enhances cell viability and promotes cellular proliferation and differentiation has not been provided so far. Although well-tailored *in vitro* and *in vivo* experiments need to be carried out to answer the above question, computational models remain a robust tool for the design of those experiments due to their remarkable potential to explain and predict experimental results in scaffold-aided bone healing processes [5–8].

For long bones, the most popular computational model used for fracture healing predictions is by Prendergast [9], known as the mechanobioregulatory model, where the octahedral shear strain in the solid phase of the bone and the fluid velocity in the interstitial fluid phase are used as regulators of the tissue differentiation process [10–12]. Analogously, the mechanobioregulatory computational model of Prendergast has been exploited by many investigators for providing scaffold-aided bone regeneration predictions [13–16]. However, octahedral shear strain and interstitial fluid velocity concern cell differentiation regulators valid for the bone fracture biological environment and not for the initial cell proliferation and differentiation at the surface of a scaffold structure. At this early stage, all cells being attached at the surface of the scaffold experience only the surface deformation and the wall shear stress imposed by the loading of the scaffold and the surrounding interstitial flow, respectively. Therefore, Hendrikson et al. [15, 16] proposed the octahedral volumetric strains and interstitial fluid velocity be replaced by the new regulators corresponding to surface strains and shear wall stress, respectively.

Most of the scaffolds tested so far for tissue regeneration concern architectures with cylindrical fibers made of biocompatible polymers. All those 3D fused filament fabricated scaffolds are usually made up of layers containing parallel struts with the same or different orientation for each layer, with the most common being that with a $0/90^0$ stacking sequence. However, such scaffolds appear problems associated with strut bending between two supports, disturbances of the shape and the volume of the pores, and degradation of the mechanical performance of the scaffold. To resolve this issue, multi-layer deposition of polymeric or particulate composite polymeric filaments has been adopted in the fabrication of rectangular and staggered $0/90$ scaffolds [17–23]. Along with the filament deposition, strut's material properties, and scaffold geometry, the pore size is an important parameter that influences drastically the scaffold-aided bone regeneration [24]. As it is mentioned in all cited works, scaffolds should exhibit mechanical properties that approach those of a physiological bone and adequate permeability to facilitate cell adhesion and proliferation, nutrient transport, and vascularization. Large, interconnected pores promote permeability but degrade the mechanical properties of the scaffold. Thus, scaffolds with multi-layer particulate composite polymeric filaments provide a compromised solution between mechanical and permeability requirements.

In the context of this work, the mechanoregulatory model of Hendrikson et al. [15, 16] is employed to explore the possible location of the early-stage cell attachment in a four-layer orthogonal scaffold with isometric pores made of polylactic acid PLA/316L with 5% stainless steel particle content when subjected to three-point bending. For the evaluation of surface strains and shear wall stresses, a hybrid computational mechanics framework is proposed based on conventional numerical simulation and Physics-Informed Neural Networks (PINNs) predictions. More precisely, the Boundary Element Method (BEM) is utilized to evaluate the displacement and the surface strain distributions

at the surface points of the scaffold [25–27]. A PINN framework is further implemented, as a reduced order model (ROM), taking as input the surface displacement of the scaffold provided by the BEM boundary solutions and the material properties, to compute the displacement field in the internal domain of the scaffold [28, 29]. Moreover, the Element-based Finite Volume Method (EbFVM) [30] is utilized to simulate the Navier Stokes equations and calculate the velocity and wall shear stresses imposed by the interstitial blood flow at the surface of the scaffold. The goal of the present work is to show that conventional numerical methods like the BEM and EbFVM in conjunction with PINN algorithms are able, through mechanoregulatory models, to provide a robust numerical tool for scaffold-aided bone regeneration simulations.

The paper is structured as follows: Sect. 2 presents the geometry, the material composition, and the structure of the scaffold studied in the present work. Section 3 addresses in brief the numerical modeling of the problem via the BEM and the PINN, as well as the solution to the CFD problem with the EbFVM. Section 4 provides the numerical results dealing with the displacements, and the surface strain distribution, as well as the velocity, and wall shear stresses on the scaffold surface. These results are introduced in the mechanoregulatory model of Hendrikson et al. [15, 16] to provide predictions for the early-stage viability and differentiation of the cells attached at the surface of the scaffold. Finally, the main conclusions of the present work and thoughts about future steps are provided in Sect. 5.

2 Scaffold Design

In this work, an orthogonal scaffold with a rectangular, four-layer strut design has been studied. The scaffold is composed of 9 filament rows, with a $0^\circ/90^\circ$ stacking sequence. Each strut is of width $w = 0.142$ mm and height $h = 0.125$ mm. The length of the scaffold is $L_s = 3.375$ mm, its width is $D_s = 1.875$ mm, and its height equals $H_s = 1.125$ mm. The distance between the struts is equal to 0.238 mm (Fig. 1i). The scaffold porosity percentage, defined by the volume of the scaffold and the total volume of a solid cube that has the same outline as the size of the designed scaffold, equals 65% [31].

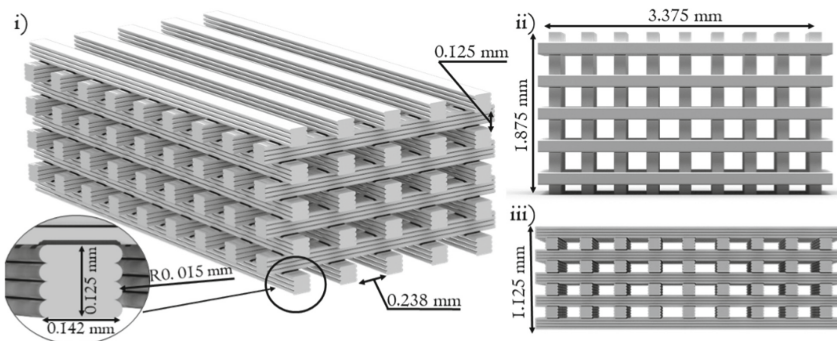


Fig. 1. Geometric representation of the rectangular and four-layer scaffold. The geometric details of the designed scaffold are indicated in i) isometric, ii) upper side, and iii) left side view.

The composite material of the scaffold is PLA enhanced with steel nanoparticles. PLA is a biodegradable polymer suitable for fabricating porous scaffolds in bone tissue engineering [20]. To improve the mechanical behavior of the PLA, steel particles are added into the PLA matrix, with specific content, and construct composite scaffolds. In the present work, the designed scaffold is made of PLA/316L with stainless steel particle content of 5 vol%. We consider that the steel particle number is large and randomly distributed and therefore, the composite material is assumed to be isotropic and linearly elastic. This is a realistic assumption since the cooling of each filament prevents the movement of the metallic powder inside and the four-layer composition of each strut ensures the uniform distribution of the steel particles in the internal space [22]. The Young modulus of the PLA when the steel powder is close to 5%, equals $E = 0.28$ GPa [32], while the Poisson ratio for the PLA is $\nu = 0.37$ [22].

This particular type of scaffold was chosen for the following reasons: (i) Due to the multi-layer fabrication of the rectangular struts and their reinforcement with steel particles, compromises pore size, mechanical performance, and blood flow requirements. (ii) The PLA/iron scaffold appears remarkable resistance against defect formation and layer delamination [22]. (iii) A PLA scaffold with a tetragonal structure presents the best mechanical behavior with quicker tissue growth compared to other scaffolds with different compositions and geometry [33]. (iv) The considered scaffold appears lower surface-to-volume ratio, compared to the corresponding one-layer rectangular scaffold, a fact that promotes better cell migration, proliferation, and differentiation [24].

3 Numerical Modeling

In this section, the computational tools and the problems solved to numerically study the designed scaffold are described. In the first subsection, a brief description of the BEM is provided, while in the second subsection, the PINN-based ROM framework employed for the fast computation of the displacements in the interior of the scaffold is presented. Finally, in the third subsection, the CFD tools, used to study the interaction of the scaffold with a fluid flow field, are briefly described.

3.1 Boundary Element Method (BEM)

In this work, the BEM is utilized for all the calculations regarding the structural behavior of the scaffold. The BEM is a robust numerical tool for solving linear elastic problems [26, 34]. Two well-known advantages of the BEM are the accurate evaluation of stresses and the dimensionality reduction of the problem since only the boundary of the domain needs to be discretized. In this section, the machinery for the numerical solution of elastostatic problems with the BEM is described.

The equilibrium equation for an isotropic and homogeneous linear elastic body is given as follows,

$$\mu \partial_{ii} u_i(\mathbf{x}) + (\lambda + \mu) \partial_j \partial_j u_i(\mathbf{x}) = 0 \quad (1)$$

where $u_i(\mathbf{x})$ are the Cartesian components of the displacement vector at a point \mathbf{x} of the elastic body with volume V and external boundary S , ∂_i denotes differentiation with respect to Cartesian coordinate x_i , while λ and μ are the Lamé parameters.

A well-posed boundary value problem for linear elasticity, requires appropriate boundary conditions for the displacement and traction vectors \mathbf{u} and \mathbf{t} , respectively, at the boundary S of the body, i.e.,

$$\begin{aligned} u_i(\mathbf{x}) &= \bar{u}_i(\mathbf{x}), \quad \mathbf{x} \in S_u \\ t_i(\mathbf{x}) &= \bar{t}_i(\mathbf{x}), \quad \mathbf{x} \in S_t \end{aligned} \quad (2)$$

where \bar{u}_i, \bar{t}_i denote prescribed displacement and traction values with $S_u \cup S_t = S$ and $S_u \cap S_t = \emptyset$. The solution of Eq. (1) is represented in an integral form as follows [24],

$$C_{ij}(\mathbf{x})u_j(\mathbf{x}) + \int_S T_{ij}^*(\mathbf{x}, \mathbf{y})u_j(\mathbf{y})dS_{\mathbf{y}} = \int_S U_{ij}^*(\mathbf{x}, \mathbf{y})t_j(\mathbf{y})dS_{\mathbf{y}} \quad (3)$$

where \mathbf{x} denotes the collocation point and \mathbf{y} denotes the field point, while $C_{ij}(\mathbf{x})$ is the jump tensor obtaining the values $C_{ij}(\mathbf{x}) = \delta_{ij}$ when $\mathbf{x} \in \Omega$, and $C_{ij}(\mathbf{x}) = \frac{1}{2}\delta_{ij}$ when \mathbf{x} belongs to a smooth part of the boundary S , with δ_{ij} denoting the components of the identity matrix. The expressions for the kernel tensor functions, $U_{ij}^*(\mathbf{x}, \mathbf{y})$ and $T_{ij}^*(\mathbf{x}, \mathbf{y})$, appearing in Eq. 3 can be found for 3D elastic problems in [34, 35].

In the BEM formulation, integral Eq. (3) is utilized along with the boundary conditions (2) for the solution of the problem. Assuming that the boundary of the body is discretized into N_E boundary elements, with each element e defined by $N_F^{(e)}$ nodes, Eq. (3) obtains the following form for a collocation point \mathbf{x}_k :

$$\mathbf{C}_k \mathbf{u}_k + \sum_{e=1}^{N_E} \sum_{i=1}^{N_F^{(e)}} \mathbf{H}_{kei} \cdot \mathbf{u}_i = \sum_{e=1}^{N_E} \sum_{i=1}^{N_F^{(e)}} \mathbf{G}_{kei} \cdot \mathbf{t}_i \quad (4)$$

where $\mathbf{u}_k = \mathbf{u}(\mathbf{x}_k)$, and the matrices \mathbf{H}_{kei} and \mathbf{G}_{kei} contain the integrals appearing in Eq. (3). A challenge for the application of BEM is the computation of the boundary integrals, which exhibit singular behavior when $\mathbf{x} \rightarrow \mathbf{y}$ [34]. Collocating Eq. (4) for all the nodes of the boundary the following linear system of equations can be obtained:

$$\mathbf{H} \cdot \mathbf{u} = \mathbf{G} \cdot \mathbf{t} \quad (5)$$

where the vectors \mathbf{u} and \mathbf{t} contain all the displacements and tractions for the nodes of the boundary. By applying the boundary conditions of the problem, a linear system of equations $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ is obtained, where the matrix \mathbf{A} is dense and non-symmetric.

In the conventional BEM, the fully populated matrix \mathbf{A} requires $O(N^2)$ operations for its buildup and $O(N^3)$ operations for the solution of the final linear system of equations through typical LU-decomposition solvers. The dense nature of the \mathbf{A} matrix has traditionally confined the BEM to small-scale problems, where the degrees of freedom N cannot be more than a few hundred of thousand. To overcome this limitation, various accelerated BEM implementations have been proposed, employing hierarchical matrices and Adaptive Cross Approximation (ACA) techniques [26, 36–39], combined with an iterative Generalized Minimum Residual (GMRES) solver. These BEM implementations, partition the geometry into clusters and utilize algebraic or hybrid algorithms

to compress the submatrices of \mathbf{A} , which contain the values of the integrals corresponding to certain cluster combinations of collocation points and boundary elements. Using these techniques, the assembly time for the matrix \mathbf{A} , the required memory for its storage, as well as the solution time of the final linear system of equations can be significantly reduced.

In a post-processing step, using Eq. (3) with $C_{ij} = \delta_{ij}$ the displacement of an interior point \mathbf{x} can be calculated, using the previously computed boundary values of the tractions and the displacements. Furthermore, for a point \mathbf{x} , lying on a smooth part of the boundary, the strain can be calculated through the displacement gradient computed with the following integral equation [35],

$$\frac{1}{2} \partial_k u_i(\mathbf{x}) + \int_S P_{kij}^*(\mathbf{x}, \mathbf{y}) u_j(\mathbf{y}) dS_{\mathbf{y}} = \int_S Q_{kij}^*(\mathbf{x}, \mathbf{y}) t_j(\mathbf{y}) dS_{\mathbf{y}} \quad (6)$$

which can be derived by applying the gradient operator to Eq. (3). Having computed the displacement gradient, the strain is given as $\varepsilon_{ij}(\mathbf{x}) = \frac{1}{2} (\partial_i u_j(\mathbf{x}) + \partial_j u_i(\mathbf{x}))$. The mathematical expressions for the tensor kernels $P_{kij}^*(\mathbf{x}, \mathbf{y})$ and $Q_{kij}^*(\mathbf{x}, \mathbf{y})$ are given in [34, 35].

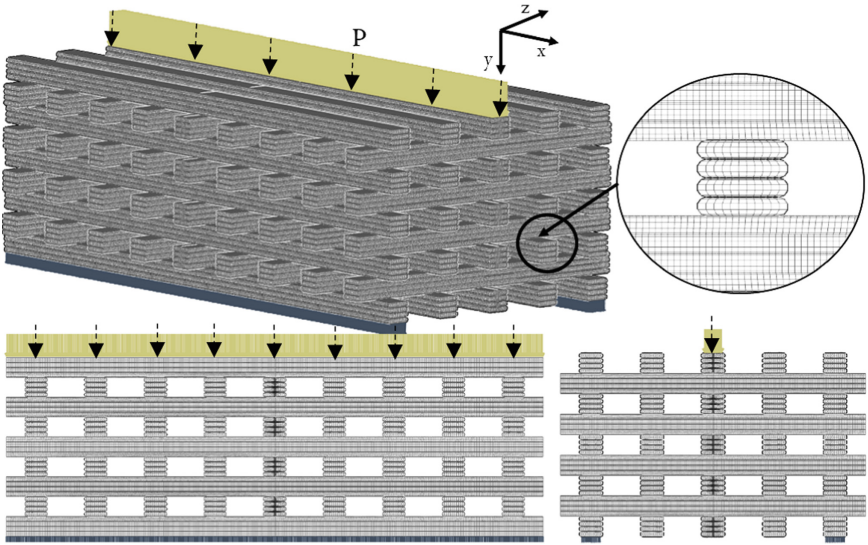


Fig. 2. Boundary conditions and mesh grid of the designed scaffold for the numerical solution of the structural problem.

To simulate a three-point bending loading scenario of the scaffold structure, a vertical compressive load is imposed on the upper, central face of the scaffold with magnitude $p = 5$ MPa, as it is shown in Fig. (2). The bottom left and right faces of the scaffold are kept fixed. For the BEM simulation, an in-house developed software using fast solvers for integral equations is employed [26, 39].

3.2 Physics-Informed Neural Networks (PINNs)

In recent years, ROMs [40] or deep learning-based solvers, such as PINNs [41] are commonly employed to obtain approximate solutions for parameterized, large-scale PDEs in a limited amount of time. A PINN is a meshless method for solving PDEs that utilizes an artificial neural network (ANN) to approximate in an unsupervised or semi-supervised manner the solution of a boundary value problem [42]. The method integrates the information from the physical laws of the boundary value problem, by including the PDEs that described the problem in the loss function during the training of the ANN.

In general, an ANN, also known as a multi-layer perceptron, approximates the functional relation among a set of inputs and outputs. It contains several layers with numerous neurons. The first and last layers represent the input and output layers, respectively and the intermediate layers are called hidden layers. When the ANN contains more than one hidden layer, it is called a deep neural network. Each layer performs a transformation by combining a linear mapping, with the application of a non-linear activation function. Based on a well-known extension of the universal approximation theorem, an ANN consisting of an arbitrary number of hidden layers, each of them containing an arbitrary number of neurons, can approximate a function with respect to some accuracy.

In the frame of this work, a multilayer ANN is employed for the fast computation of the displacements at the interior of the scaffold, taking as input the coordinates, the material properties, and the pressure load combined in a single vector $\mathbf{x}_I = (x, y, z, E, \nu, p)$ while predicting the displacement vector $\tilde{\mathbf{u}} = (\tilde{u}, \tilde{v}, \tilde{w})$. The ANN approximation can be expressed as follows,

$$\mathbf{u}(\mathbf{x}_I) \approx \tilde{\mathbf{u}}(\mathbf{x}_I) = \varphi_f(\mathbf{x}_I; \mathbf{w}_f, \mathbf{b}_f) \quad (7)$$

where $\mathbf{w}_f, \mathbf{b}_f$ are the vectors containing the weight and bias parameters, and φ is a (non)-linear mapping $N : \mathbb{R}^6 \rightarrow \mathbb{R}^3$. The entries of the weight matrices and bias vectors constitute the parameters of the neural network $[\mathbf{w}_f, \mathbf{b}_f]$, which are computed using optimization techniques so that the neural network approximates the functional relation of interest with respect to certain criteria imposed by a predefined loss function. The derivatives of the ANN parameters can be calculated with automatic differentiation using well-known libraries (e.g., PyTorch, TensorFlow).

In the context of this work, to include the equations related to the physics of the problem in the training of the PINN, the PDE along with the boundary conditions are incorporated into the loss function. More specifically, a PINN framework is implemented to predict the solution of the homogenous Navier-Cauchy equations (Eqs. 1, 2), and provide the displacement distribution in the interior of the scaffold geometry (Fig. 4). The developed PINN framework is further used as ROM since the ANN is trained by including

the material properties and the bending load to the input vector. The loss function that includes the governing equations and the boundary conditions can be expressed as [28, 29, 43, 44],

$$L = L_{PDE} + L_{BC} \quad (8)$$

Regarding the loss of the boundary conditions, we get the numerical solution of the variable \mathbf{u} on the boundaries from the BEM solutions. Therefore the L_{BC} term obtains the following form:

$$L_{BC} = \frac{1}{N_{bc}} \sum_{i=1}^{N_{bc}} [(\tilde{u} - u^{BEM})^2 + (\tilde{v} - v^{BEM})^2 + (\tilde{w} - w^{BEM})^2]_{(i)} \quad (9)$$

where u^{BEM} , v^{BEM} , w^{BEM} are the data provided by the high-fidelity BEM solutions on the boundaries and N_{bc} is the number of the training points uniformly sampled from the mesh on the scaffold surface. While the loss derived from the PDE, L_{PDE} is written as follows (Fig. 3),

$$L_{PDE} = \frac{1}{N_p} \sum_{i=1}^{N_p} \left[\left(\mu \left(\frac{\partial^2 \tilde{u}}{\partial x^2} + \frac{\partial^2 \tilde{u}}{\partial y^2} + \frac{\partial^2 \tilde{u}}{\partial z^2} \right) + (\lambda + \mu) \left(\frac{\partial^2 \tilde{u}}{\partial x^2} + \frac{\partial^2 \tilde{v}}{\partial y \partial x} + \frac{\partial^2 \tilde{w}}{\partial z \partial x} \right) \right)^2 + \right. \\ \left. \left(\mu \left(\frac{\partial^2 \tilde{v}}{\partial x^2} + \frac{\partial^2 \tilde{v}}{\partial y^2} + \frac{\partial^2 \tilde{v}}{\partial z^2} \right) + (\lambda + \mu) \left(\frac{\partial^2 \tilde{u}}{\partial x \partial y} + \frac{\partial^2 \tilde{v}}{\partial y^2} + \frac{\partial^2 \tilde{w}}{\partial z \partial y} \right) \right)^2 + \right. \\ \left. \left(\mu \left(\frac{\partial^2 \tilde{w}}{\partial x^2} + \frac{\partial^2 \tilde{w}}{\partial y^2} + \frac{\partial^2 \tilde{w}}{\partial z^2} \right) + (\lambda + \mu) \left(\frac{\partial^2 \tilde{u}}{\partial x \partial z} + \frac{\partial^2 \tilde{v}}{\partial y \partial z} + \frac{\partial^2 \tilde{w}}{\partial z^2} \right) \right)^2 \right]_{(i)} \quad (10)$$

where N_p is the number of the training points uniformly sampled from the scaffold mesh. Where the notation (i) denotes that the function is applied for each point i .

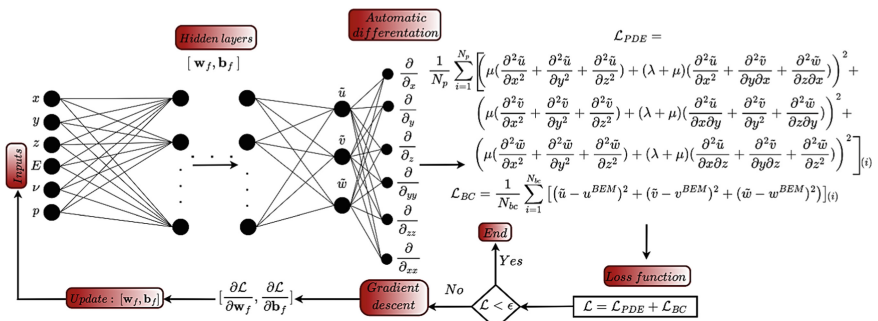


Fig. 3. Schematic diagram of the PINN framework for the Navier-Cauchy equations. The PINN is used as a ROM, taking as inputs the coordinates, the material properties, and the boundary conditions, derived through the BEM, aiming to predict the deformation in the scaffold domain.

The adaptive moment estimation (ADAM) optimizer, a version of stochastic gradient descent is utilized for the optimization process of the ANN. The PINN is trained

for 10000 epochs, with a batch size of 2000 and a learning rate equal to 0.001. All the parameters of the ANN are tuned using a hypermeter optimization analysis. Regarding the ANN architecture, 10 hidden layers are employed, each including 30 neurons, selected in an ad-hoc manner, while the \tanh is used as an activation function. Regarding the computational times, the BEM needs 30 min to compute a high-fidelity solution and provide the boundary results to the PINN framework. The PINN requires 1 h of training using an Intel Core i7 @ 2.20GHz, NVIDIA Ge-Force GTX 1050 Ti GPU computer.

3.3 Element-Based Finite Volume Method (EbFVM)

As a second step, we examine the interaction of the scaffold with a fluid flow to further elucidate the mechanism of the scaffold design in the early-stage cell differentiation procedure. The objective is to compute the distribution of the wall shear stress on the scaffold surface, as this is a quantity that is often linked to the differentiation of the cells on the scaffold [16, 45]. To this end, high-fidelity CFD simulations are performed with the EbFVM [30], using the flow solver available in the commercial software SIEMENS Simcenter 3D.

The steady-state incompressible fluid flow within a volume $\Omega \subset R^d$, satisfies the Navier-Stokes equations,

$$\rho \mathbf{v}(\mathbf{x}) \cdot \nabla \mathbf{v}(\mathbf{x}) + \nabla p(\mathbf{x}) - \mu \nabla^2 \mathbf{v}(\mathbf{x}) = \mathbf{0}, \mathbf{x} \in \Omega \quad (11)$$

$$\nabla \cdot \mathbf{v}(\mathbf{x}) = 0, \mathbf{x} \in \Omega \quad (12)$$

where $\mathbf{v}(\mathbf{x}) : \Omega \rightarrow R^d$, is the velocity field, ρ is the fluid density, $p(\mathbf{x}) : \Omega \rightarrow R$ is the pressure field, and $\mu \in R^+$ is the dynamic viscosity.

We consider a Newtonian fluid with density $\rho = 1000 \text{ kg/m}^3$ and dynamic viscosity $\mu = 0.001 \text{ Pa} \cdot \text{s}$. The laminar inlet flow is perpendicular to the top plane of the scaffold, with a magnitude $v_{in} = 3 \text{ mm/s}$. Symmetry conditions are imposed on the z -axis to decrease the computational time, and zero pressure on the outlet face, while no-slip condition (zero velocity) is applied to the remaining boundary faces (Fig. 4).

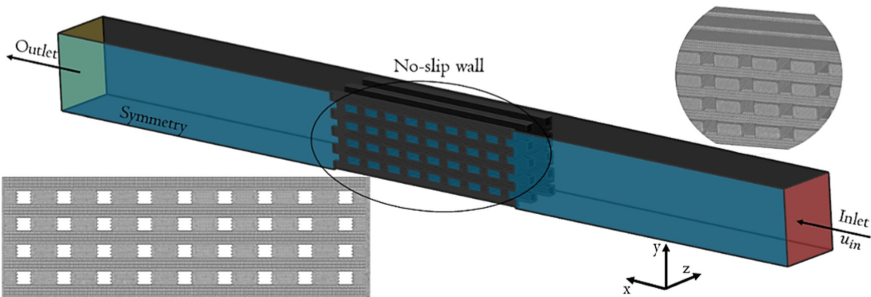


Fig. 4. Mesh grid of the designed scaffold and boundary conditions for the computational fluid dynamics problem.

4 Numerical Results

In this section, the numerical results of the computational mechanics simulations including the CSM and CFD solutions using the BEM, the PINN, and the EbFVM around the scaffold geometry are presented.

4.1 Computational Structural Mechanics (CSM) Analysis

In this subsection, the main results are presented for the structural behavior of the scaffold. Regarding the utilized mesh grid, it consists of 1,010,880 isoparametric, linear, quadrilateral elements and 1,010,200 nodes on the scaffold surface (Fig. 2). The mean element size of the discretization is equal to 0.01mm and has been derived after a convergence analysis. The resulting number of degrees of freedom for this problem is about 3 million and cannot be solved with the conventional BEM. To overcome this problem a BEM implementation based on Hierarchical Matrices and ACA has been employed [26, 39]. For the solution of the final linear system of equations, the GMRES iterative solver has been utilized.

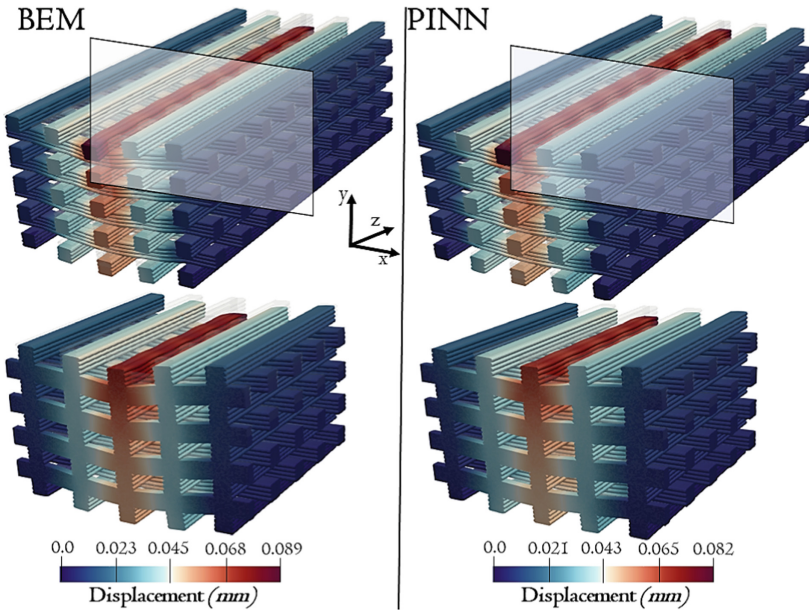


Fig. 5. Displacement magnitude obtained through the BEM and the PINN on the scaffold geometry (top) and the half (bottom). The results are scaled x2 to better visualize the results.

In Fig. 5 the magnitude of the displacement computed through the BEM simulations, as well as the PINN results obtained using the BEM boundary solutions, are provided. The PINN framework is utilized as a ROM aiming to provide the solution in the scaffold domain given the material properties and the boundary conditions. In particular, the high-fidelity solution of the BEM on the scaffold surface is given as a boundary condition

to the PINN framework. To compare the BEM and PINN results in the domain with the BEM, based on the surface mesh we generate a volume mesh for the scaffold. In a second step, the displacements at the interior nodes are obtained with the BEM in a post-processing step using Eq. (3), since the displacements and tractions at the boundary of the domain have already been computed. Similarly, the PINN is utilized for the computation of the displacements on the interior nodes of the scaffold. The results for both methods, presented in Fig. 5, are in very good agreement. As expected, the larger values of the displacement are obtained on the top middle strut of the filament row, located on the region of the bending load.

As soon as the surface displacements and the tractions are known, using Eq. (6) the strains are evaluated with BEM for the boundary points. The computed strains are then transformed, to calculate the surface shear strain on each element face defined as follows [16],

$$\gamma_{srf} = \frac{1}{2}(\varepsilon_1 - \varepsilon_2) \quad (13)$$

where ε_1 and ε_2 are the principal strains on the element surface. To compute the principal strains on the surface of each element the strain tensor is rotated with respect to a coordinate system $(\mathbf{n}, \mathbf{s}_1, \mathbf{s}_2)$ where \mathbf{n} is the unit normal vector of each element, and $\mathbf{s}_1, \mathbf{s}_2$ are the unit vectors on the plane of each element. The orthogonal projection of the rotated strain tensor on the surface of the element is the 2×2 part of the rotated matrix with components corresponding to the unit vectors \mathbf{s}_1 and \mathbf{s}_2 . The obtained eigenvalues of this matrix are the values of ε_1 and ε_2 used in Eq. (13).

The calculated surface shear strains are presented in Fig. 6. It is shown that significant surface strains are obtained as expected in the middle strut of each filament row due to the increased deformations in that region. The results reveal a more localized surface strain distribution, at the crossings of the fibers with a peak value close to 0.1. This is to be expected, due to the non-uniform, bending loading condition.

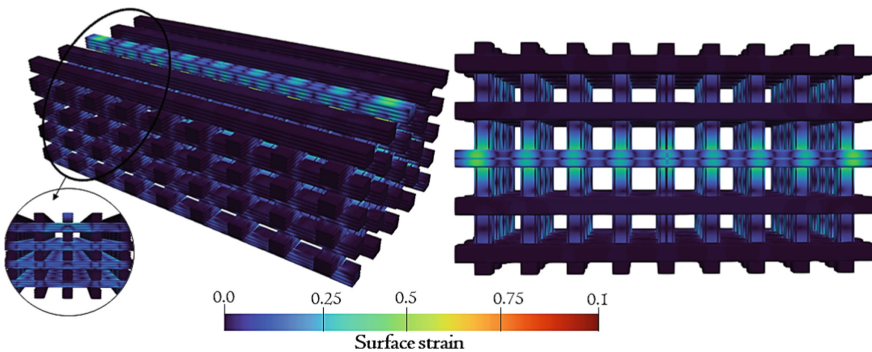


Fig. 6. Surface strain distribution on the scaffold geometry obtained by the BEM solution.

4.2 Computational Fluid Dynamics (CFD) Analysis

In this section, the velocity, and wall shear stress distributions imposed by the fluid flow on the scaffold surface are presented. A fine mesh grid has been employed for the CFD analysis, including 26,219,072 tetrahedral solid elements, resulting in 4,620,051 nodes. A local refinement is used near the scaffold area, with the mean element size being equal to 0.01 mm (Fig. 4).

In Fig. 7, the magnitude of the velocity contour in the domain is presented. It is observed that the maximum velocity is approximately six times higher than the inlet velocity, indicating the acceleration of the blood flow due to the variability in the inlet area and the scaffold region [46]. Observing the streamlines contour, leads us to consider that the geometry and the inlet velocity, significantly affect the blood flow within the scaffold.

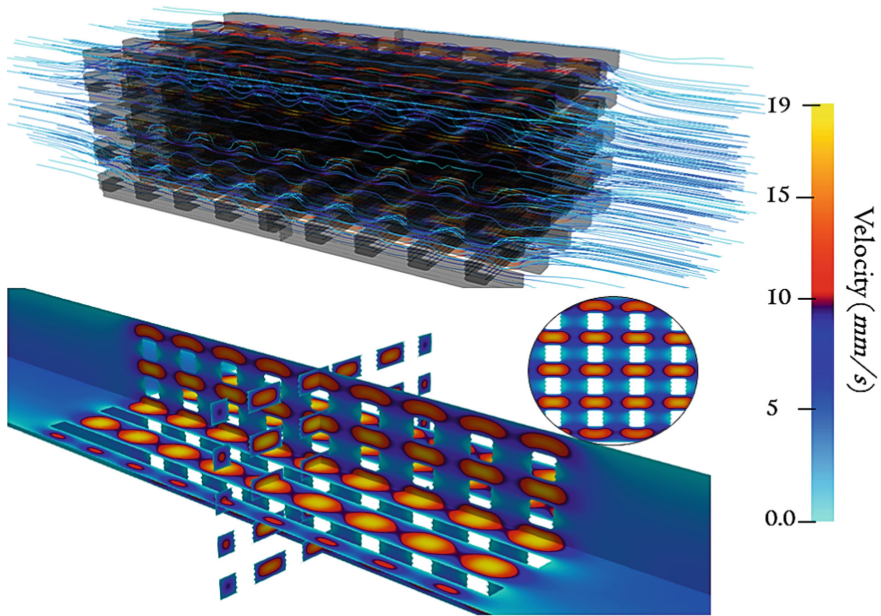


Fig. 7. Velocity contour on the scaffold geometry obtained through the CFD analysis.

In Fig. 8 the wall shear stresses of the fluid flow on the scaffold geometry are displayed. High values are observed on the top faces of each strut along the filament row, and side faces of the scaffold. The pick values are obtained on those faces, which are parallel to the fluid flow, and further present the largest velocity values (Fig. 7) [47]. The wall shear stress distribution is uniform, with a maximum value close to 0.9 Pa.

4.3 Mechanoregulatory Model for the Early-Stage Cell Differentiation

The cell differentiation on the scaffold surface can be computed based on the following mechanoregulatory model [16] which considers both the surface shear strain, and the

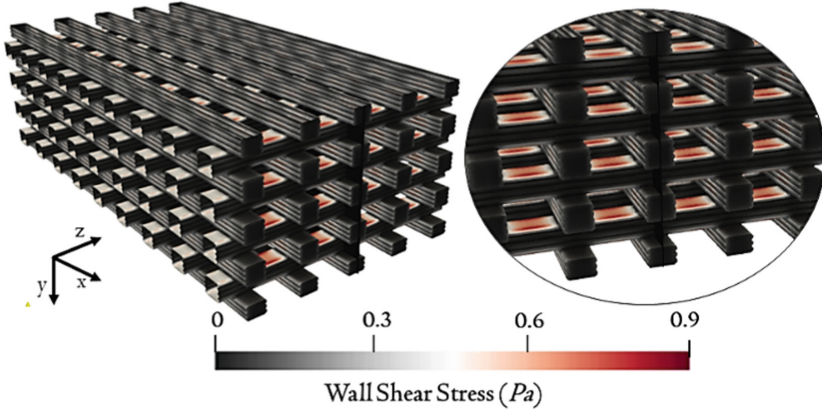


Fig. 8. Wall shear stress contour on the scaffold geometry obtained through the CFD analysis.

fluid wall shear stress as follows,

$$S = \frac{\gamma_{srf}}{\alpha} + \frac{\tau}{b} \tag{14}$$

where S is the cell stimulus, γ_{srf} is the surface strain calculated in Ch. 4.1, τ is the wall shear stress obtained in Ch. 4.2, while $a = 0.0375$ and $b = 0.01 Pa$, are constants influencing the behavior of the model [16] (Fig. 9). Based on the cell stimulus values the cell types on the surface of the scaffold are characterized [16]. In particular, the cell stimulus S values thresholds are $0 \leq S < 0.001$ for resorption, revealing that in this case the cell stimulus is too low and cannot induce any differentiation to the cell, $0.001 \leq S < 1$ for bone generation, $1 \leq S < 3$ for cartilage, $3 \leq S < 6$ for fibrous tissue, and $S \geq 6$ for necrosis, indicating that the stimuli are too intense to induce any differentiation.

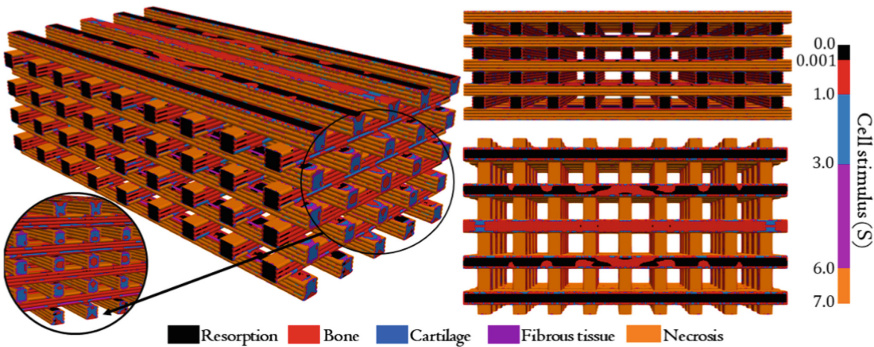


Fig. 9. Cell differentiation stimulus distribution on the scaffold surface.

In Fig. 9, it is observed the highest values of the cell stimulus are observed on the regions of the pick wall shear stress. Therefore, the utilized cell differentiation prediction

model [16], based on a combination of the surface shear stress and the wall shear stress, reveals that for the current loading setup, the fluid flow affects more the cell stimulus value than the bending load. In particular, based on the obtained values of the surface shear strain and the fluid wall shear stress, for a distribution of cells on the scaffold, about 0.67% of the cells would correspond to resorption, 45.56% to bone, 9.58% to cartilage, 6.60% to fibrous tissue and 37.56% to necrosis.

5 Conclusions

In the present work, an innovative, hybrid methodology based on numerical simulation and deep learning is developed, to study the mechanical behavior of a composite scaffold made of PLA/316L with 5% stainless steel particle content. The BEM was used to evaluate the structural behavior of the scaffold when a bending load is applied and compute the displacement and the surface shear strain on the scaffold. It is shown that the BEM is an ideal methodology for the calculation of the surface strain for a linear elastic material. Furthermore, a PINN framework is developed as a ROM to provide the solution given the materials properties and the boundary conditions. It is proven that the PINN is a promising methodology for constructing reduced order models which explicitly incorporate the physics of the problem, while its training can be achieved by utilizing information obtained from numerical methods such as the BEM. In terms of accuracy, both the BEM and PINN results of the displacements in the scaffold domain are in good agreement.

Regarding the CFD calculations, the obtained wall shear stress has higher values in the areas parallel to the fluid flow. The predicted CFD and CSM results are used to exploit the efficiency of the scaffold design by calculating the cell differentiation stimulus and obtaining the regions of possible tissue development. It is shown that for the developed numerical approach and the designed geometry, most of the scaffold surface is predicted in the bone generation while the minority belongs to resorption.

Future steps, amongst others, include the implementation of ROMs for the calculation of the surface strains and the wall shear stresses for various material properties, pores size, and boundary conditions.

Acknowledgments. This research work is supported by the Hellenic Foundation for Research and Innovation (HFRI) under the “First Call for HFRI. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant” (Project Number: 2060).

References

1. Velasco, M.A., Narváez-Tovar, C.A., Garzón-Alvarado, D.A.: Design, Materials, and Mechanobiology of Biodegradable Scaffolds for Bone Tissue Engineering. *Biomed Res Int.*, 729076 (2015).
2. Yin, S., Zhang, W., Zhang, Z., Jiang, X.: Recent Advances in Scaffold Design and Material for Vascularized Tissue-Engineered Bone Regeneration. *Adv. Healthcare Mater.* 8, 1801433 (2019).

3. Szymczyk-Ziółkowska, P., Labowska, M.B., Detyna, J., Michalak, I., Gruber, P.: A review of fabrication polymer scaffolds for biomedical applications using additive manufacturing techniques. *Biocybernetics and Biomedical Engineering* 40, 624–638 (2020).
4. Roque, R., Barbosa, G.F., Guastaldi, A.C.: Design and 3D bioprinting of interconnected porous scaffolds for bone regeneration: An additive manufacturing approach. *Journal of Manufacturing Processes* 64, 655–663 (2021).
5. Lacroix, D., Planell, J.A., Prendergast, P.J.: Computer-Aided Design and Finite-Element Modelling of Biomaterial Scaffolds for Bone Tissue Engineering. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 367, 1895 (2009).
6. Thavorniyutikarn, B., Chantarapanich, N., Sittthiseripratip, K., Thouas, G.A. Chen, Q.: Bone tissue engineering scaffolding: computer-aided scaffolding techniques. *Prog Biomater* 3, 61–102 (2014).
7. Post, J.N., Loerakker, S., Merks, R.M.H., Carlier, A. Implementing Computational Modeling in Tissue Engineering: Where Disciplines Meet. *Tissue Engineering Part A* 28, Numbers 11 and 12 (2022).
8. Mustafa, N.S., Akhmal, N.H., Izman, S., Ab Talib, M. H., Shaiful, A.I.M., Omar, M.N.B., Yahaya, N.Z., Illias, S.: Application of Computational Method in Designing a Unit Cell of Bone Tissue Engineering Scaffold: A Review. *Polymers* 13, 1584 (2021).
9. Prendergast P.J.: Finite element models in tissue mechanics and orthopedic implants design. *Clin. Biomech.* 12, 343–366 (1997).
10. Lacroix, D., Simulation of tissue differentiation during fracture healing. Ph.D., University of Dublin (2001).
11. Checa, S., Prendergast, P.J.: Effect of cell seeding and mechanical loading on vascularization and tissue formation inside a scaffold: A mechano-biological model using a lattice approach to simulate cell activity. *Journal of Biomechanics* 43, 961–968 (2010).
12. Grivas, K.N., Vavva, M.G., Carlier, A., Polyzos, D., Geris, L., Van Oosterwyck, H., Fotiadis, D.I.: Effect of ultrasound on bone fracture healing: A computational mechanobioregulatory model. *Journal of the Acoustical Society of America* 145 (2), 1048–1059 (2019).
13. Olivares, A, Marsal, E, Planell, J.A., Lacroix, D. Finite element study of scaffold architecture design and culture conditions for tissue engineering. *Biomaterials* 30, 6142–6149 (2009).
14. Boccaccio, A., Uva, A.E., Fiorentino, M., Lamberti, L., Monno, G.: A Mechanobiology-based Algorithm to Optimize the Microstructure Geometry of Bone Tissue Scaffolds. *International Journal of Biological Sciences* 12(1), 1–7 (2016).
15. Hendrikson, W.J., van Blitterswijk, C.A., Verdonschot, N., Moroni, L., Rouwkema, J.: Modeling mechanical signals on the surface of microCT and CAD based rapid prototype scaffold models to predict (early stage) tissue development. *Biotechnol. Bioeng.* 111, 1864–1875 (2014).
16. Hendrikson, W.J., Deegan, A.J., Yang, Y., van Blitterswijk, C.A., Verdonschot, N., Moroni L., Rouwkema, J.: Influence of Additive Manufactured Scaffold Architecture on the Distribution of Surface Strains and Fluid Flow Shear Stresses and Expected Osteochondral Cell Differentiation. *Frontiers in Bioengineering and Biotechnology* 5:6, <https://doi.org/10.3389/fbioe.2017.00006> (2017).
17. Nyberg, E., O’Sullivan, A., Grayson, W.: ScafSLICR: A MATLAB-based slicing algorithm to enable 3D-printing of tissue engineering scaffolds with heterogeneous porous microarchitecture. *PLoS ONE*, doi: <https://doi.org/10.1371/journal.pone.0225007> (2019).
18. Cubo-Mateo, N., Rodríguez-Lorenzo, L.M.: Design of Thermoplastic 3D-Printed Scaffolds for Bone Tissue Engineering: Influence of Parameters of “Hidden” Importance in the Physical Properties of Scaffolds. *Polymers* 12, 1546 (2020).
19. Khogalia, E.H., Choo, H.L., Yap, W.H.: Performance of Triply Periodic Minimal Surface Lattice Structures Under Compressive Loading for Tissue Engineering Applications. In: *AIP Conference Proceedings* 2233, 020012 (2020); <https://doi.org/10.1063/5.0001631>.

20. Baptista, R., Guedes, M.: Morphological and mechanical characterization of 3D printed PLA scaffolds with controlled porosity for trabecular bone tissue replacement. *Materials Science & Engineering C* 118, 111528 (2021).
21. Baptista, R., Guedes, M.: Porosity and pore design influence on fatigue behavior of 3D printed scaffolds for trabecular bone replacement. *Journal of the Mechanical Behavior of Biomedical Materials* 117, 104378 (2021).
22. Jiang, D., Ning, F., Wang, Y.: Additive manufacturing of biodegradable iron-based particle reinforced polylactic acid composite scaffolds for tissue engineering. *Journal of Materials Processing Tech.* 289, 116952 (2021).
23. Gortsas, T.V., Tsinopoulos, S.V., Polyzos, E., Pyl, L., Fotiadis, D.I., Polyzos D. BEM evaluation of surface octahedral strains and internal strain gradients in 3D-printed scaffolds used for bone tissue regeneration. *Journal of the Mechanical Behavior of Biomedical Materials* 125, 104919 (2022).
24. Perier-Metz, C., Cipitria, A., Hutmacher, D.W., Duda, G.N., Checa, S.: An in silico model predicts the impact of scaffold design in large bone defect regeneration. *Acta Biomaterialia* 145, 329–341 (2022).
25. Polyzos, D., Tsinopoulos, S.V., Beskos, D.E.: Static and dynamic boundary element analysis in incompressible linear elasticity. *European Journal of Mechanics – A/Solids* 17, 515–536 (1998).
26. Gortsas, T.V., Tsinopoulos, S.V., Polyzos, D.: An advanced ACA/BEM for solving 2D large-scale problems with multi-connected domains. *CMES-Computer Modeling in Engineering & Science* 107 (4), 321–343 (2015).
27. Rodopoulos, D., Gortsas, T.V., Tsinopoulos, S.V., Polyzos, D.: Numerical evaluation of strain gradients in classical elasticity through the Boundary Element Method, *European Journal of Mechanics / A Solids* 86, 104178 (2021).
28. Haghighat, E., Raissi, M., Moure, A., Gomez, H., Juanes, R.: A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. *Comput. Methods Appl. Eng.* 379, 113741 (2021).
29. Henkes, A., Wessels, H., Mahnken, R.: Physics informed neural networks for continuum micromechanics, arXiv preprint [arXiv:2110.07374v2](https://arxiv.org/abs/2110.07374v2) (2022).
30. Honorio, H.T., Maliska, C.R., Ferronato, M., Janna C.: A stabilized element-based finite volume method for poroelastic problems, *Journal of Computational Physics* 364, 49-72 (2018).
31. Rosa, N., Pouca, M.V., Torres, P.M.C., Olhero, S.M., Jorge, R.N., Parente, M.: Influence of structural features in the performance of bioceramic-based composite scaffolds for bone engineering applications: A prediction study. *Journal of Manufacturing Processes* 90, 391-405 (2023).
32. Jiang, D., Ning, F.: Fused Filament Fabrication of Biodegradable PLA/316L Composite Scaffolds: Effects of Metal Particle Content, *Procedia Manufacturing* 48, 755-762 (2020).
33. Liang, X., Gao, J., Xu, W., Wang, X., Shen, Y., Tang, J., Cui, S., Yang, X., Liu, Q., Yu, L., Ding, J.: Structural mechanics of 3D-printed poly(lactic acid) scaffolds with tetragonal, hexagonal and wheel-like designs. *Biofabrication*, 11, 035009, (2019).
34. Aliabadi, M.H.: *The Boundary Element Method (Volume 2): Application in Solids and Structures*. 1st edn. Wiley, Chichester (2002).
35. Polyzos, D., Tsinopoulos, S.V., Beskos, D.E.: Static and dynamic boundary element analysis incompressible linear elasticity. *European Journal of Mechanics – A/Solids* 17(3), 515–536 (1998).
36. Benedetti, I., Aliabadi, M.H., Davi, G.: A fast 3D dual boundary element method based on hierarchical matrices. *International Journal of Solids and Structures* 45(7-8), 2355-2376 (2008).

37. Haider, A.M., Schanz, M.: Adaptive Cross Approximation for BEM in Elasticity. *Journal of Theoretical and Computational Acoustics* 27(1), 1850060 (2019).
38. Bebendorf, M., Grzhibovskis, R.: Accelerating Galerkin BEM for linear elasticity using adaptive cross approximation. *Mathematical Methods in the Applied Sciences* 29(14), 1721-1747 (2006).
39. Gortsas, T.V., Tsinopoulos, S.V., Polyzos, D.: An accelerated boundary element method via cross approximation of integral kernels for large-scale cathodic protection problems. *Computer-Aided Civil and Infrastructure Engineering* 37(7), 848–863 (2022).
40. Drakoulas, G. I., Gortsas T. V., Bourantas G. C., Burganos V. N., Polyzos D.: *FastSVD-ML-ROM*: A reduced-order modeling framework based on machine learning for real-time applications. arXiv preprint [arXiv:2207.11842](https://arxiv.org/abs/2207.11842) (2022).
41. Le-Duc, T., Nguyen-Xuan, H., Lee., J.: A finite-element-informed neural network for parametric simulation in structural mechanics. *Finite Elements in Analysis and Design* 217, 103904 (2023).
42. Xu, C., Cao, B.T., Yuan, Y., Meschke, G.: Transfer learning based physics-informed neural networks for solving inverse problems in engineering structures under different loading scenarios. *Comput. Methods Appl. Mech. Eng.* 405, 115852 (2023).
43. Gu, Y., Zhang, C., Golub, M.V.: Physics-informed neural networks for analysis of 2D thin-walled structures. *Eng. Anal. Bound. Elem.* 145, 161-172 (2022).
44. Kamali, A., Sarabian, M., Laskari, K.: Elasticity imaging using physics-informed neural networks: Spatial discovery of elastic modulus and Poisson's ratio. *Acta Biomaterialia* 155, 400-409 (2023).
45. Liu, L., Wang, S., Liu, J., Deng, F., Li, Z., Hao, Y.: Architectural design of Ti6Al4V scaffold controls the osteogenic volume and application area of the scaffold, *Journal of Materials Research and Technology* 9(6), 15849-15861 (2020).
46. Omar, A.M., Hassan, M.H.; Daskalakis, E., Ates, G., Bright, C.J., Xu, Z., Powell, E.J., Mirhanage, W., Bartolo, P.J.D.S.: Geometry-Based Computational Fluid Dynamic Model for Predicting the Biological Behavior of Bone Tissue Engineering Scaffolds. *J. Funct. Biomater.* 13, 104 (2023).
47. Ali, D., Sen, S.: Permeability and fluid flow-induced wall shear stress of bone tissue scaffolds: Computational fluid dynamic analysis using Newtonian and non-Newtonian blood flow models. *Computers in Biology and Medicine* 99, 201-208 (2018).



Ensembled Multi-classification Generative Adversarial Network for Condition Monitoring in Streaming Data with Emerging New Classes

Yu Wang¹ (✉), Qingbo Wang¹, and Alexey Vinogradov²

¹ Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway

yuwa@ntnu.no

² Institute of Light Alloys, Kumamoto University, Kumamoto, Japan

Abstract. Condition monitoring (CM) process can be viewed as the problem of streaming classification with emerging new classes (SENC). There are four fundamental challenges we are faced with: (1) timely detection of emerging new classes; (2) model update to adapt to new classes; (3) pattern recognition of these already known classes with high accuracy and (4) distinguishing between different new classes. Although intelligent surrogate models like deep learning have achieved remarkable success in the field of CM, SENC still remains a thorny challenge that has seldom been studied. This paper presents an approach to address the challenge of SENC using an ensembled multi-classification generative adversarial network (EMC-GAN). The proposed method includes a novel deep network architecture called MC-GAN that integrates the tasks of novelty detection and multi-classification into a single framework. To address issues of model stability, an efficient history-state ensemble (HSE) method that does not require additional training costs to generate multiple base models is introduced. Experimental validation is conducted on four simulated SENC tasks using benchmark data, and the results have shown the effectiveness of the proposed approach.

Keywords: Streaming data with emerging new classes (SENC) · Condition monitoring (CM) · History-state ensemble (HSE) · Generative adversarial network · Novelty detection

1 Introduction

Condition monitoring (CM) of rotating machinery is a continuous process of identifying changes that are indicative of the developing failures. The main objectives include: (1) timely detection of the breakpoint when the monitored machine deviates from its normal operating condition, (2) accurate pattern recognition for multiple failure types, and (3) quantification of the damage development rate. With the improvement in computational ability, intelligent surrogate models based on deep learning have achieved remarkable success in multiple fault diagnosis [1–4], early fault detection [5–7] and remaining useful life prediction [8], etc. Traditional CM frameworks attempt to tackle these tasks

separately, and their limitations are standing out from the surroundings. Most existing surrogate models are built on an ideal assumption that when we perform CM on a particular machine, failures encountered in the real streaming data must be included in the categories of training data. Nevertheless, this requirement is hard to meet in real-world applications because the signals acquired from sensors are fed into the CM system over time, one cannot access the whole dataset to train the surrogate models. A common solution is to utilize historical datasets to train the models, however in practice, it is rather difficult to obtain a representative set of actual fault data, not to mention the problem complexity due to numerous possible failure types and their locations. It is impossible to create an all-round dataset that contains all failure signals. Therefore, the surrogate models must have the ability to handle these unseen failure types in streaming data while identifying these already-known faults, a scenario known as streaming classification with emerging new classes (SENC), where most state-of-art approaches fail.

The challenges faced in traditional SENC problem are decomposed as follows [9]: (1) timely detection of emerging new classes that are not included in the training dataset; (2) automatically model update in order to adapt to new classes; (3) multi-classification of the already known classes, and (4) discriminating different new classes. The surrogate models are required to preserve the historical health condition memories of the machine, which contributes to maintenance efficiency. Due to the growing need to reduce reliance on expert knowledge and manpower, it is required that (5) the surrogate models are capable to extract features and make the data-driven decisions automatically. To this end, a novel SENC framework is proposed to address the above challenges. The main contributions of this paper are briefly summarized as:

- (1) A new architecture called MC-GAN is proposed to tackle the SENC problem by integrating novelty detection into the multi-classification framework.
- (2) The MC-GAN architecture is further improved by incorporating a novel ensemble technique known as the “history-state ensemble” (HSE) method, which significantly enhances the model’s performance without requiring additional training resources.
- (3) The effectiveness of the proposed framework is validated through four simulated SENC tasks using benchmark vibration signals.

2 Methodology

2.1 Notations

Given a training dataset $T = \{(x_i, y_i)\}_{i=1}^m$, where $x_i \in R^d$ is the i -th training instance and y_i represents the corresponding label or class which belongs to $Y = \{1, 2, \dots, c\}$, c is the total number of known classes. The task is to build an initial multi-classification model based on the given dataset T so as to process the stream data $S = \{(\tilde{x}_t, \tilde{y}_t)\}_{t=1}^{\infty}$, where $\tilde{x}_t \in R^d$ is a streaming instance at time t , and the corresponding class $\tilde{y}_t \in \bar{Y} = \{1, 2, \dots, c, c + 1, c + 2, \dots, c + K\}$, K is the number of all new classes.

2.2 Basic Structure of the Proposed MC-GAN

Figure 1 illustrates the main structure of the proposed model, comprising a Generator and a Discriminator that are inherited from traditional GAN. Similarly, the Generator is employed to produce the fake data that are close to the input real data, and the Discriminator is trained to distinguish between the real and fake data. As the original Discriminator in traditional GAN is designed as a binary classifier, its network architecture is required to be modified to accommodate multi-classification tasks. As depicted in Fig. 1, the designed Discriminator's architecture consists of two distinct channels, which are employed to learn implicit features from different aspects of the signals.

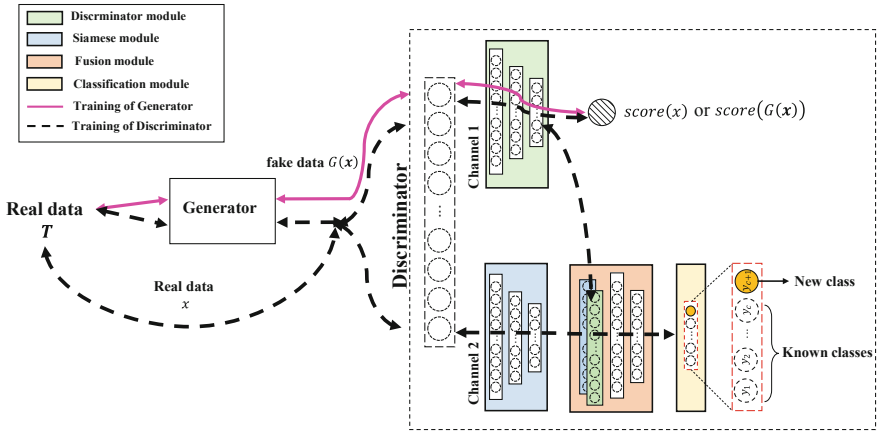


Fig. 1. Basic structure of the proposed MC-GAN.

The ‘channel 1’ contains a ‘Discriminator module’, which is used to learn the similarities of all real data from different categories that can distinguish them from the fake data. Therefore, the module does not distinguish between different classes of real data. It takes in all input instances and outputs a ‘score’ for each of them that indicates whether the instance belongs to fake or real. A linear fully connected layer is employed to compute the score, expressed as:

$$score(\mathbf{x}) = \omega_s D_M(\mathbf{x}) + b_s \quad (1)$$

where $\mathbf{x} \in T$, D_M represents the Discriminator module, which is a neural network block consisting of multiple convolutional layers and pooling layers.

The ‘channel 2’ comprises three models including:

- (1) A Siamese module, S_M , that also takes in all input instances. But different from the previous D_M , it is designed to learn the difference between different classes of real data.
- (2) A Fusion module, $F_M(D_M, S_M)$, that receives the extracted features, generally the last hidden layer of D_M and S_M .

- (3) A Classification module, C_M , that receives the fused feature from F_M . A SoftMax layer is utilized with the number of hidden neurons of $c + 1$, where c is the number of all known classes of training instances, and an additional hidden neuron is added as the detector of emerging new classes.

$$C_M(y_i|\mathbf{x}) = \frac{\exp(\omega_{C_M}^k F_M)}{\sum_{j=1}^{c+1} \exp(\omega_{C_M}^j F_M)} \quad (2)$$

where $y_i \in Y \cup \{c + 1\}$. The equation outputs the probability of \mathbf{x} belongs to label y_i , the $c + 1$ represents the label of a new class that does not belong to the training dataset T .

Detection of a new class. The basic idea of MC-GAN is to take the fake data produced by Generator as a new class to train the entire network. Given a training dataset T with c known classes, in the feed-forward phase, the input real data that are randomly sampled from T are fed into the Generator to produce fake data. The fake data $G(\mathbf{x})$ is considered as a new class labeled as ‘ $c + 1$ ’. Then, both fake and real data are fed into the modified Discriminator. The MC-GAN can directly output the predicted probability that an instance belongs to each known class or a new class. The training of MC-GAN is detailed in **Algorithm 1**.

Training of Generator. The data flow during the training of Generator is depicted by magenta line in Fig. 1. The Generator is optimized by minimizing the following loss function:

$$L(G) = -E[\text{score}(G(\mathbf{x}))] + \gamma \cdot \|\mathbf{x} - G(\mathbf{x})\|_1 \quad (3)$$

where γ is a penalty coefficient, the second term measures the L1 distance between the real data \mathbf{x} and generated data $G(\mathbf{x})$ [10].

Training of Discriminator. The discriminator is optimized by minimizing the following combined loss function:

$$L(D) = \rho \times L_d + (1 - \rho) \times L_c \quad (4)$$

where $\rho \in (0, 1)$ in formula (4) is a trade-off parameter, L_d and L_c represents the loss functions of the ‘channel 1’ and ‘channel 2’, respectively, which are defined as:

$$L_d = E[\text{score}(G(\mathbf{x}))] - E[\text{score}(\mathbf{x})] + \beta E\left[\left(\nabla_{\mathbf{x}'} \text{score}(\mathbf{x}')_2 - 1\right)^2\right] \quad (5)$$

$$L_c = E\left[\left(C_M(\tilde{\mathbf{x}}) - y\right)^2\right] \quad (6)$$

where the L_d adopts the Wasserstein distance (WD) as a measure of the distance between real and generated data distribution; the details of gp-WGAN are presented in [11, 12]. $\mathbf{x}' = \alpha \cdot \mathbf{x} + (1 - \alpha) \cdot G(\mathbf{x})$ in formula (3) represents the interpolation of \mathbf{x} and $G(\mathbf{x})$, where $\alpha \in (0, 1)$. $\tilde{\mathbf{x}} = F_M(D_M(\mathbf{x}), S_M(\mathbf{x}))$ represents the output of Fusion module .

Algorithm 1 Training of MC-GAN using RMSprop stochastic gradient descent. We use default value of Adam hyper-parameters: $lr = 0.0001$, $\alpha = 0.99$.

Input: Training dataset $T = \{(\mathbf{x}_i, y_i)\}_{i=1}^c$; Initial Discriminator weights ω^0 ; Initial Generator weights φ^0 ; Initial training hyper-parameters, including, epoch number N , batch size m , critic number n_{critic} , penalty coefficient β , and balance coefficient ρ .

Output: Well-trained ω^N and φ^N .

```

1: for  $epoch = 1, \dots, N$  do
2:   for  $critic = 1, \dots, n_{critic}$  do
3:     for  $batch = 1, \dots, m$  do
4:       Sample a batch of real data  $\mathbf{x} \sim T$ 
5:       Generate fake data  $G(\mathbf{x})$ 
6:        $L_d \leftarrow Eq. (5)$ 
7:        $L_c \leftarrow Eq. (6)$ 
8:        $L(D) \leftarrow Eq. (4)$ 
9:        $\omega \leftarrow \text{RMSprop}(\nabla_{\omega} E(L(D)), \omega, lr, \alpha)$ , optimize the Discriminator.
10:    End for
11:  End for
12:  Sample a batch of real data  $\mathbf{x} \sim T$ 
13:  Generate fake data  $G(\mathbf{x})$ 
14:   $L(G) \leftarrow Eq. (3)$ 
15:   $\varphi \leftarrow \text{RMSprop}(\nabla_{\varphi} E(L(G)), \varphi, lr, \alpha)$ , optimize the Generator.
16: End for

```

2.3 History-State Ensemble (HSE)

As the Generator and the Discriminator are competing with each other during training, their performance tends to fluctuate, which makes it challenging to assess whether the Discriminator has been sufficiently trained for the SENC task. To address the challenge, a novel ensemble technique referred to as history-state ensemble (HSE) method is introduced in this work [13]. The HSE method operates on the assumption that neural networks can generate several local optima or ‘history-states’ during training, which can serve as base models for ensemble learning. To generate multiple base models, only the historical weights of the network need to be retained, as shown in Fig. 2. As a result, the time cost of this procedure is negligible.

In order to minimize redundancy among the base models, we introduce two parameters: θ_1 and θ_2 . The former parameter, θ_1 , specifies the number of epochs required for model warm-up, while the latter parameter, θ_2 , sets the update frequency between two consecutive base models. Besides, to effectively leverage the HSE, an important conclusion drawn from our previous work is applied, which pointed out that the performance of MT-GAN is strongly related to the loss value L_c , as presented in formula (6). Therefore, the model selection is crucial. A simple model selection approach is adopted in this paper. By sorting the L_c values corresponding to each base models, these base models with smaller L_c are selected. With the selected base models,

$$Y_i = \frac{1}{m} \sum_{j=1}^m P_j(y = i | \mathbf{x}) \quad (7)$$

where $P(y = i|\mathbf{x})$ denotes the predicted probability that a given instance \mathbf{x} belongs to label $i \in \{1, 2, \dots, c, c + 1\}$. m represents the number of base models, and P_j denotes the predicted probability given by the j -th base model.

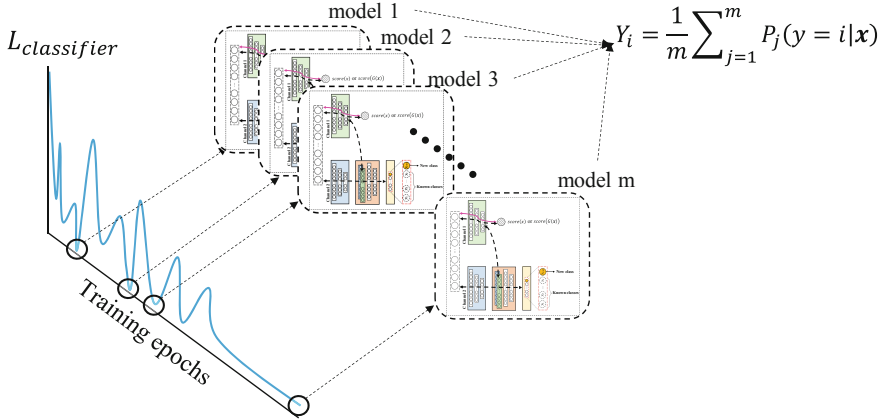


Fig. 2. Illustration of ensemble MC-GAN based on HSE.

2.4 The Proposed SENC Framework

The proposed workflow is shown in Fig. 3. In order to detect multiple new classes, the model is required to update itself over time. Like traditional SENC models, we apply a fix-sized buffer to store the instances from the detected new class. Given an initial training dataset $T = \{(\mathbf{x}_i, y_i)\}_{i=1}^m$ with c known classes, a streaming dataset $S = \{(\tilde{\mathbf{x}}_t, \tilde{y}_t)\}_{t=1}^{\infty}$ with K new classes and a fix-sized buffer B , the primary procedures of the proposed workflow are summarized as follows:

- (1) The initial training dataset T is normalized to train the initial MC-GAN.
- (2) Choose fix-sized Discriminator snapshots with the smallest L_c as base models to construct the EMC-GAN.
- (3) Each instance $\tilde{\mathbf{x}}_t$ from the streaming dataset S is fed into the ensemble Discriminator sequentially. If the predicted label of $\tilde{\mathbf{x}}_t$ is $c + 1$, then feed the instance into the buffer B .
- (4) If the buffer is full, then merge the instances in buffer into the original T , and re-normalized the new training dataset.
- (5) Re-train the MC-GAN with updated T and repeat steps (2–4) until the last instances in S .

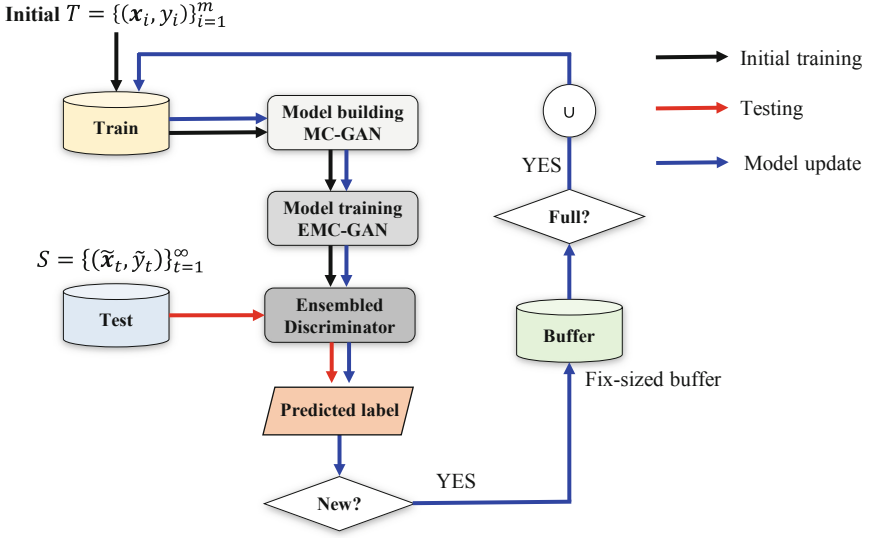


Fig. 3. The workflow of classification in SENC with EMC-GAN.

3 Experimental Validation

3.1 Dataset Description

We used a prevalent benchmark dataset collected by Case Western Reserve University (CWRU) Bearing Data Center [14] to evaluate the proposed method. The CWRU Bearing dataset contains a collection of vibration signals acquired from ball bearings mounted on a motor. Different levels of artificial damages were seeded on the three different locations of the bearings, namely the inner raceway (I), outer raceway (O), and rolling elements or balls (B). The seeded fault sizes were 0.007, 0.014, and 0.021 inches in diameter. Together with health bearing, a total of ten categories are used in this work, including Health (H), inner race faults (I7, I14, I21), outer race faults (O7, O14, O21) and rolling element faults (B7, B14, B21). The vibration signals of the utilized faulted bearings were recorded for motor loads of 1 to 3 horsepower. The sampling frequency is 12kHz. Each category contains 750 samples with dimension of 1024.

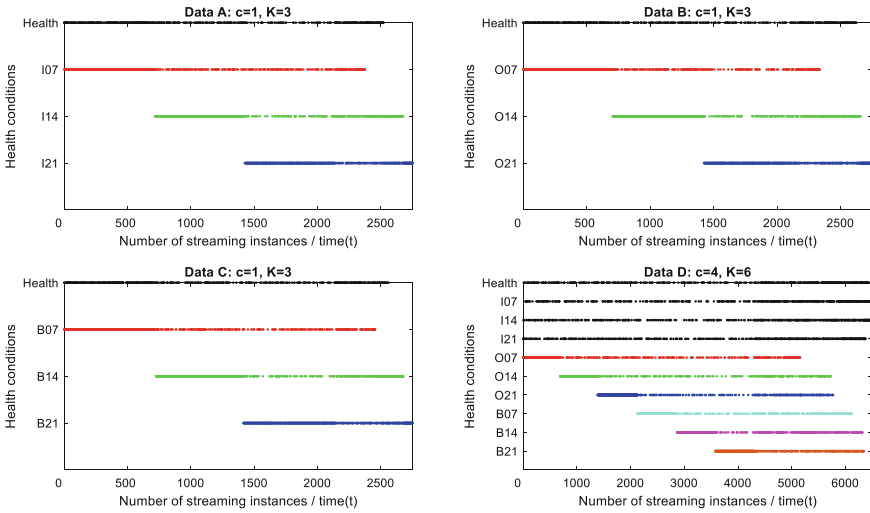
Please note that the first paragraph of a section or subsection is not indented.

Table 1 illustrates the four datasets that were generated to evaluate the proposed method. The Data A-C solely comprise of one known class, namely 'Health,' in the initial training dataset T , while the test set S encompasses four fault conditions, including three new classes that were not present in the training dataset. Data A-C were employed to investigate the model's ability to detect and differentiate between faults in the inner race, outer race, and rolling elements of the bearings, respectively. On the other hand, Data set D further increases the complexity of the problem by introducing four known classes and six emerging new classes, featuring distinct fault locations and sizes in the training and test datasets.

Table 1. The utilized datasets.

Datasets	Training set		Test set	
	Fault condition	Number	Fault condition	Number
A	H	250	H, I07, I14, I21	$500*1 + 750*3$
B	H	250	H, O07, O14, O21	$500*1 + 750*3$
C	H	250	H, B07, B14, B21	$500*1 + 750*3$
D	H, I07, I14, I21	$250*4$	H, I07, I14, I21 O07, O14, O21, B07, B14, B21	$500*4 + 750*6$

The simulated SENC tasks using the utilized datasets are visualized in Fig. 4. The health conditions designated by the color black indicate the initial known classes, while several new classes emerge at different times, represented by different colors. As shown in Fig. 3, instances will be fed into the model one by one, and to evaluate the performance of the model on ‘known classes’, samples from known classes or previously detected new classes will reappear with a certain probability.

**Fig. 4.** Visualization of the utilized datasets.

3.2 Results and Discussion

The proposed method is implemented using Pytorch. The RMSprop stochastic gradient descent is utilized for training the model. The learning rate is set at 0.0001, and the model is trained for 300 epochs on dataset A-C and 500 epochs on dataset D. The batch size is set at 20, and the trade-off parameter ρ is set to 0.5. The buffer size for model updates

is set at 200. For ensemble learning, the warm-up epochs (θ_1) is set to 100, stride (θ_2) is set to 1, and the number of base models is set at 5 for dataset A-C and 15 for dataset D.

Figure 5 displays the comprehensive outcomes of the probed methods using confusion matrix. Based on the observations, the following points can be concluded: (1) The proposed method proved to be highly effective in detecting all types of emerging unseen instances, exhibiting the excellent classification accuracy on the four datasets that were evaluated. (2) The proposed method also demonstrated remarkable performance in differentiating between various emerging new classes, with an impressive number of instances being correctly classified. This indicates that the proposed method has the potential to perform reliably in diverse real-world scenarios where emerging new classes are present.

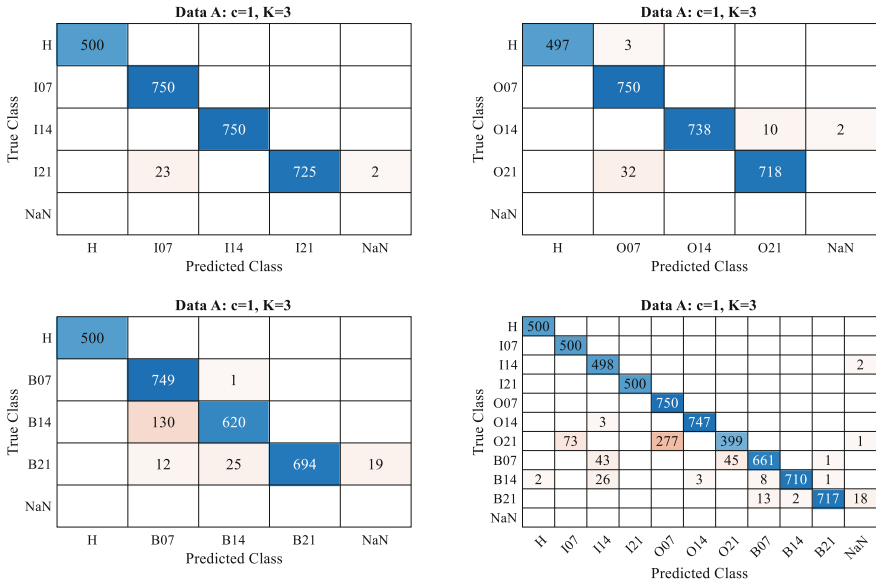


Fig. 5. The confusion matrix of the proposed method.

For comparison, three state-of-art SENC approaches are introduced:

- (1) SENNE [15], a cluster-based ensemble method which is built based on k-Nearest neighbor ensembles and can be used for both new class detection and known class classification after training.
- (2) KNNENS [9], another cluster-based ensemble method derived from SENNE and incorporates a reliable model retirement mechanism.
- (3) SENCForest [16], which incorporates a completely-random tree technique for identifying newly emerging classes by generating an ‘outlying’ anomaly subregion. Additionally, to establish a comprehensive framework for SENC, multi-classification is performed by capturing the class labels within the leaves of the previously trained trees.

As comparative methods are based on shallow learning algorithms that exhibit inadequate performance on raw mechanical signals, 15 features extracted from both the time and frequency domains were utilized as their input. These features encompass several statistics such as mean, variance, root mean square (RMS), skewness, kurtosis, shape factor, crest factor, impulse factor, margin factor, entropy, energy entropy, power, mean frequency, root mean square frequency, and the standard deviation of frequency. A grid search approach is employed to identify the optimal combination of key parameters that yield the best performance of the competing methods. Determining the decision boundaries is heavily dependent on the threshold value [9]. Nevertheless, the selection of suitable thresholds is contingent on the particularities of each dataset. Therefore, the threshold value is calibrated to guarantee that no less than 98% of the training data can be accurately identified as known.

Initially, we examine the collective performance of the evaluated methods over time compared with the above competitor. By utilizing classification accuracy as the metric, we compute the streaming accuracy (SA) as follows:

$$SA(t) = n_t/N_t \quad (8)$$

where N_t denotes the total number of instances from the start to time t , n_t^j represents the number of instances being correctly classified as the corresponding labels.

Figure 6 presents the SA of the probed method on the employed four datasets, the best performance of the probed methods is chosen for exhibition. The following observations are worth noting: (1) Initially, all methods perform well on data A–C during testing, but around the 250th instance, the performance of SENNE and SENCForest begins to decline. Meanwhile, KNNENS and the proposed EMC-GAN maintain excellent performance on dataset A and B. However, (2) KNNENS’s performance starts to decrease around the 1500th instance on datasets C and D, and the proposed EMC-GAN maintains the highest accuracy at the end of the test. (3) Although the proposed EMC-GAN also experiences a decrease in performance around the 800th instance on data C and 1500th instance on data D, indicating some incorrect predictions at that time, it subsequently recovers and exhibits an increase in performance.

The model update information of the probed methods is illustrated in Fig. 7. The baseline signifies the precise time when the model should be updated to detect all new classes. If the probed methods update the model for a new class before the baseline, it implies that some known instances are being erroneously classified as new. On the other hand, if the probed methods update the model for a new class after the baseline, it suggests that some new instances are being inaccurately classified as known.

The following points are noteworthy: (1) The frequent updates of SENNE and SENCForest, as compared to the baselines, suggest that these methods have detected numerous false new classes. This observation elucidates why these methods exhibit a decline in performance after the 250th instance on data A–C. (2) The KNNENS only update twice for dataset C, however, there are three new classes, which means that the model has poor performance to detect the third new classes. The same behavior can also be observed from data D, the KNNENS only update three times for this dataset given six new classes. In contrast, the proposed EMC-GAN successfully detected and distinguished all the emerging new classes of the four datasets.

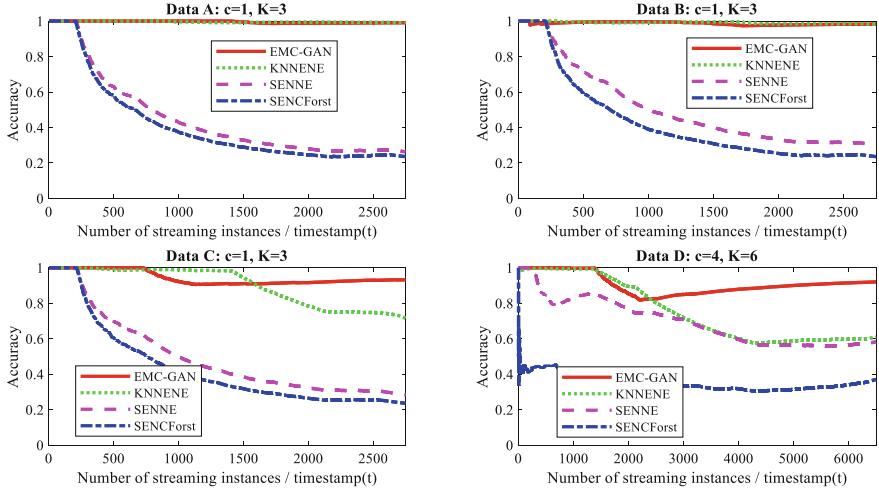


Fig. 6. The streaming accuracy (SA) of the probed methods.

Table 2 presents the average accuracy and standard deviation of the probed methods. It can be seen that the EMC-GAN achieves over 90% accuracy on the four datasets. However, the SENNE and SENCForest perform poor because they tend to provide many false new classes after model update.

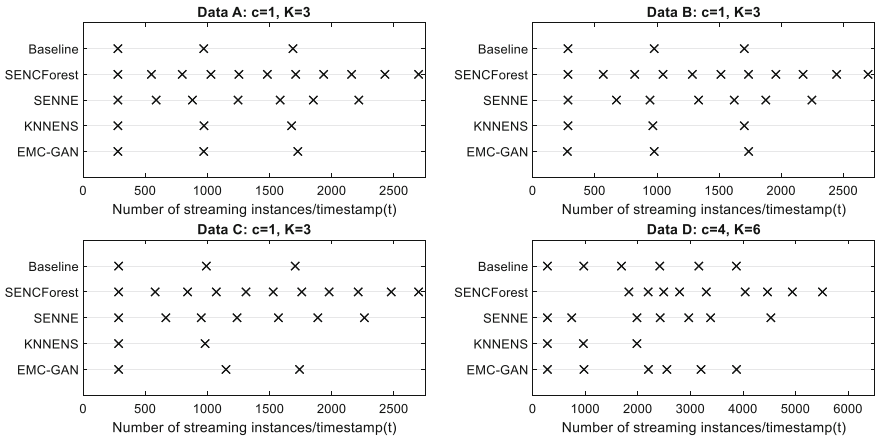


Fig. 7. The comprehensive illustration of the model update timestamp and frequency of the probed methods.

Table 2. Comparison of the average accuracy/standard deviation among probed methods.

Methods	Datasets			
	A	B	C	D
EMC-GAN	95.12%/0.06	94.21%/0.0574	92.21%/0.0086	91.72%/0.0069
KNNENS	94.49%/0.0974	89.66%/0.1178	79.45%/0.0653	55.17%/0.0479
SENNE	18.81%/0.0818	21.52%/0.0938	19.81%/0.0831	39.74%/0.2295
SENCForest	22.44%/0.021	22.47%/0.0207	23.17%/0.0257	33.17%/0.0114

4 Conclusion

In this paper, a novel framework, namely EMC-GAN, is proposed to tackle the challenges associated with SENC. The effectiveness of the proposed method is investigated through four simulated SENC tasks based on well-established benchmark datasets. The proposed method offers several advantages over traditional SENC models, including: (1) being threshold-free, the EMC-GAN approach integrates novelty detection procedures with the multi-classification task and can accurately predict whether a signal belongs to one of the known classes or a new class; (2) the proposed EMC-GAN method can process raw signals directly without requiring any manual feature extraction which is in line with the modern pursuit of intelligent model; and (3) the EMC-GAN model achieves high classification accuracy between different emerging new classes. Overall, the proposed EMC-GAN approach has the potential to significantly enhance the accuracy and efficiency of SENC systems, enabling more reliable and effective monitoring of mechanical systems.

References

1. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of Machine Learning to Machine Fault Diagnosis: A Review and Roadmap. *Mech. Syst. Signal Process.* **2020**, *138*, doi:<https://doi.org/10.1016/j.ymsp.2019.106587>.
2. Zhang, Y.; Ji, J.C.; Ren, Z.; Ni, Q.; Wen, B. Multi-Sensor Open-Set Cross-Domain Intelligent Diagnostics for Rotating Machinery under Variable Operating Conditions. *Mech. Syst. Signal Process.* **2023**, *191*, 110172, doi:<https://doi.org/10.1016/j.ymsp.2023.110172>.
3. Zhao, X.; Yao, J.; Deng, W.; Ding, P.; Zhuang, J.; Liu, Z. Multiscale Deep Graph Convolutional Networks for Intelligent Fault Diagnosis of Rotor-Bearing System Under Fluctuating Working Conditions. *IEEE Trans. Ind. Inform.* **2023**, *19*, 166–176, doi:<https://doi.org/10.1109/TII.2022.3161674>.
4. Li, X.; Jiang, H.; Wang, R.; Niu, M. Rolling Bearing Fault Diagnosis Using Optimal Ensemble Deep Transfer Network. *Knowl.-Based Syst.* **2021**, *213*, 106695, doi:<https://doi.org/10.1016/j.knsys.2020.106695>.
5. Xia, X.; Pan, X.; Li, N.; He, X.; Ma, L.; Zhang, X.; Ding, N. GAN-Based Anomaly Detection: A Review. *Neurocomputing* **2022**, *493*, 497–535, doi:<https://doi.org/10.1016/j.neucom.2021.12.093>.

6. Song, W.; Shen, W.; Gao, L.; Li, X. An Early Fault Detection Method of Rotating Machines Based on Unsupervised Sequence Segmentation Convolutional Neural Network. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–12, doi:<https://doi.org/10.1109/TIM.2021.3132989>.
7. Zhou, H.; Lei, Z.; Zio, E.; Wen, G.; Liu, Z.; Su, Y.; Chen, X. Conditional Feature Disentanglement Learning for Anomaly Detection in Machines Operating under Time-Varying Conditions. *Mech. Syst. Signal Process.* **2023**, *191*, 110139, doi:<https://doi.org/10.1016/j.ymsp.2023.110139>.
8. Pan, Y.; Wu, T.; Jing, Y.; Han, Z.; Lei, Y. Remaining Useful Life Prediction of Lubrication Oil by Integrating Multi-Source Knowledge and Multi-Indicator Data. *Mech. Syst. Signal Process.* **2023**, *191*, 110174, doi:<https://doi.org/10.1016/j.ymsp.2023.110174>.
9. Zhang, J.; Wang, T.; Ng, W.W.Y.; Pedrycz, W. KNNENS: A k-Nearest Neighbor Ensemble-Based Method for Incremental Learning Under Data Stream With Emerging New Classes. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, 1–8, doi:<https://doi.org/10.1109/TNNLS.2022.3149991>.
10. Isola, P.; Zhu, J.-Y.; Zhou, T.; Efros, A.A. Image-to-Image Translation with Conditional Adversarial Networks 2018.
11. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein GAN 2017.
12. Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A.C. Improved Training of Wasserstein Gans. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, doi:Improved training of wasserstein gans.
13. Wang, Y.; Vinogradov, A. Improve the Performance of Convolutional GAN Using History-State Ensemble for Unsupervised Early Fault Detection with Acoustic Emission Signals 2023.
14. Zhao, Z.; Li, T.; Wu, J.; Sun, C.; Wang, S.; Yan, R.; Chen, X. Deep Learning Algorithms for Rotating Machinery Intelligent Diagnosis: An Open Source Benchmark Study. *ISA Trans.* **2020**, *107*, 224–255, doi:<https://doi.org/10.1016/j.isatra.2020.08.010>.
15. Cai, X.-Q.; Zhao, P.; Ting, K.-M.; Mu, X.; Jiang, Y. Nearest Neighbor Ensembles: An Effective Method for Difficult Problems in Streaming Classification with Emerging New Classes. In Proceedings of the 2019 IEEE International Conference on Data Mining (ICDM); 2019; pp. 970–975.
16. Mu, X.; Ting, K.M.; Zhou, Z.-H. Classification Under Streaming Emerging New Classes: A Solution Using Completely-Random Trees. *IEEE Trans. Knowl. Data Eng.* **2017**, *29*, 1605–1618, doi:<https://doi.org/10.1109/TKDE.2017.2691702>.



Ultrasonic-Based Stress Identification of a Reinforced Concrete Beam via USR-Net

Zheng Gang^{1,2}(✉), Yu Jigang¹, Gu Zhuangzhuang³, and Zhang Benniu^{1,2}

¹ Chongqing Jiaotong University, Nanan District, Chongqing, China

zhenggang@cqjtu.edu.cn

² State Key Laboratory of Mountain Bridge and Tunnel Engineering, Chongqing, China

³ University of South Carolina, Columbia, USA

Abstract. In recent two decades, though great progress has been made in stress identification of reinforced concrete structures using the ultrasonic-based coda wave interference technics, the state of art techniques in this field seems to be capable of considering only a single loading pattern. To address this issue, in this paper, a USR-Net based on DL algorithm is proposed for stress identification of reinforced concrete beams under two different load patterns. It is featured by a successive use of S-transform and Residual Blocks, with the latter's conventional "square" convolution kernel replaced by a "flat" one, which gives play to the pictorial characteristics of the time-frequency matrix from the former for a more efficient and accurate stress recognition. The effectiveness of USR-Net is validated on experimental tests of two concrete I-beams under respectively a three-point bending and an eccentric compression load. Compared with traditional non-DL ultrasonic based algorithms, the USR-Net is more adaptive for different load patterns and robust to noises, providing a potentially novel solution for stress identification of practical concrete structures.

Keywords: Ultrasonic coda wave · Concrete stress · Deep learning · Residual block

1 Introduction

The most frequently employed approach for evaluating stress in reinforced concrete structures involves measuring changes in strain within structural members. But, the installation and removal of strain gauges can be a time-consuming process, especially for larger structures such as bridges. In addition, strain gauges can only measure stress changes caused by loading/unloading, or by thermal expansion and contraction-induced creep. They cannot measure the actual stress of structural components, such as permanent load stress [1]. Core drilling and slot cutting are currently widely used techniques for measuring the total stress of concrete structures. However, these techniques rely on the assumption that concrete is linearly elastic, and that stress distribution along depth is uniformly, which does not correspond to the actual situation [2]. Moreover, these methods damage concrete members and are limited in their ability to provide accurate and reliable measurements.

In recent two decades, with the development of ultrasonic non-destructive (NDT) techniques, researchers have begun investigating the use of linear and nonlinear acoustic waves for direct stress measurement in concrete members, like coda wave measurement. Coda waves are secondary waves [3] that are generated by the scattering of primary waves within a medium. One of the primary advantages of using coda waves is their high sensitivity to small changes in mediums, allowing for precise measurements [4]. Also, coda wave measurement has the advantage of real-time monitoring of stress changes over time [5]. This is useful for identifying changes in stress caused by environmental loads or changes in permanent loads.

Previous works have devoted lots of efforts on applying coda wave to measure stress, Taylor series expansion and coda wave interferometry were utilized for coda wave measurement. Snieder [6] first proposed the concept of coda wave interference, demonstrating its sensitivity to medium changes and utilizing it to quantitatively measure changes in coda wave velocity. Payan [7, 8] used coda wave interference to study the functional relationship between the relative change rate of concrete wave velocity and stress, and extracted the third-order nonlinear coefficient of concrete materials. Ali Hafiz [9] used Coda Wave Comparison (CWC) technique to measure changes in the internal stress that are not observable on the surface of concrete. Clau [10] links essential key parameters computed from the coda wave signals to concrete structural state variables.

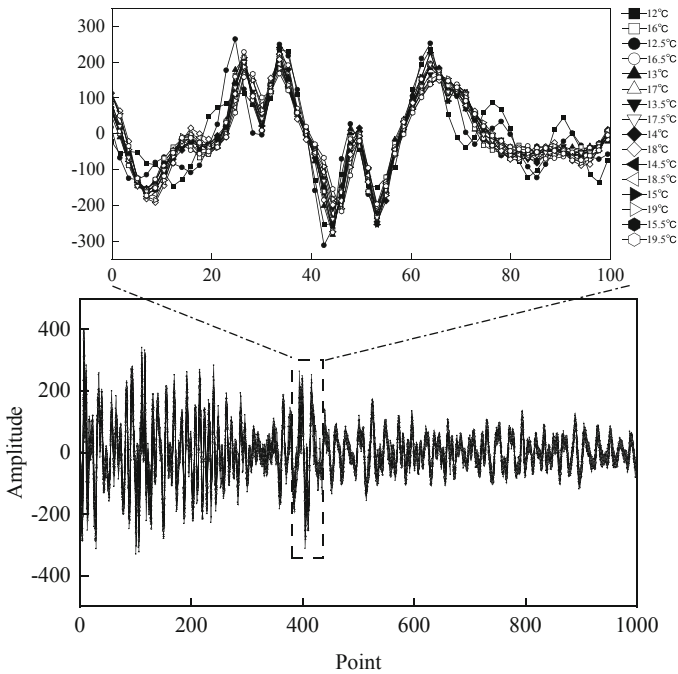


Fig. 1. Coda wave signal under same stress from different temperature.

However, these algorithms have 4 main limitations as below. (1) The amplitude of coda wave signals is typically small, making them susceptible to random noise (such as electronic noise, environmental noise, sensor noise, harmonic distortion, attenuation and reflection, etc.) and decreasing the signal quality, ultimately affecting the accuracy of stress measurement [11]. (2) Measurement results from traditional algorithms are influenced by the measurement environment, such as temperature (see Fig. 1), humidity, and noise [12, 13]. (3) These techniques require preprocessing of input data to extract relevant features, such as frequency, pulse-wave velocity, amplitude and spectrum area [1, 14, 15], that is related to the performance of the instrument and test environment. The stress recognition method of a specific parameter has poor resistance to interference, and the misjudgment rate will increase once the raw signal is polluted by noise.

To address these problems, based on the signal analysis of S-transform, innovating and improving the coda wave measurement stress method by introducing the residual Block, this study proposes USR-Net (ultrasonic-based S-transform with Residual Block Net), a deep learning model for ultrasonic-based concrete stress identification, outperforms the state-of-the-art method on identification accuracy. By extracting the time-domain and frequency-domain features of the coda wave signal comprehensively, it can effectively provide accurate stress prediction results under various loading patterns in a constantly changing environment. USR-Net provides an effective, accurate, and adaptive stress measurement method for construction monitoring, health detection, etc.

Comparing with traditional coda wave measurement methods (non-DL methods) that depend on the single time domain signal or frequency domain signal, the S-transform in the top of USR-Net can convert the raw coda wave signal into the time-frequency domain [16, 17]. By adding filters and threshold processing after S-transform to remove time-frequency components below a certain amplitude threshold, it removes random noise and improves signal quality [18, 19]. In addition, the time-frequency spectrogram output by S-transform can be analyzed and interpreted to extract frequency features [20, 21], instantaneous frequency, amplitude information, which eliminates the limitation of traditional coda wave measurement algorithms that are limited to specific signal parameters and solves the problem of weak anti-interference ability of non-DL methods. To solve the problem of interference from environmental changes on the stability of coda wave signals, the USR-Net uses multi-layer Residual Blocks to enhance feature extraction capabilities and improve the key feature extraction capabilities of coda wave signals under various environments [22]. Furthermore, the introduction of Residual Blocks improves the classification accuracy and robustness of the USR-Net [23]. Therefore, the USR-Net improves the resistance of coda wave measurement to random noise and environmental interference, as well as the ability to extract key features. This method can quickly and accurately predict the stress of concrete structures, and the real-time characteristics of USR-Net have application prospects in real-time monitoring. This study highlights the potential of deep learning algorithms in monitoring the health of concrete structures and contributes to the development of more precise and effective stress measurement techniques.

2 USR-Net

2.1 Overview

USR-Net aims to identify accurate concrete stress from the coda wave signals collected under various stress of two load patterns by solving these challenges. This could enable many new applications in stress measurement, cracks detection in practical engineering under challenging environment, like continuous changing temperature and humidity. As Fig. 2 shows, USR-Net consists of two modules: (1) S-transform-based Time-frequency Analyzer and Filter (STA and STF), which analyzes and visualizes signals in both time and frequency domains, enabling more accurate representation of the time-frequency characteristics. It can filter and denoise the raw signal to improve signal quality and reliability. (2) Feature Extractor with Residual Blocks and Classifier, which enable adaptively learn features from the visual output of STA and accurately predict stress. In the USR-Net as Fig. 2, we firstly input coda wave signal into STF to filter and denoise, and then utilize STA to analyze the denoised signal and obtain the time-frequency spectrogram image (TF-image). Finally, the Feature Extractor will extract the features from the TF-image and the Classifier classify these features and output the predicted result.

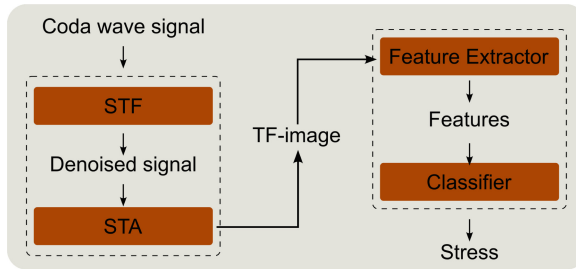


Fig. 2. USR-Net overview.

2.2 STF and STA

The core objective of this module is to denoise the raw coda wave signal and transform the denoised signal to the TF-image. The STF utilizes the multiscale and multidirectional decomposition property of the S-transform to decompose the signal into sub-signals with different scales and directions. The signal is then denoised by adjusting the threshold of the sub-signals. The STA utilizes the superior properties of the S-transform in both the time and frequency domains to transform denoised signal into the time-frequency domain for analysis.

Basic Theory of STF STF is a method based on S-transform used for filtering signal, we refer to some practices of previous works [16, 24, 25], such as threshold filtering method to filter the decomposed signal and inverse S-transform to reconstruct the filtered signal. In the process of signal processing, we find that these methods have a good effect on the removal of random Gaussian noise. The implementation of STF mainly involves the following steps:

Nomenclature

f	frequency	Hz
j	imaginary part	
N	signal length	
t	time	s
$h(t)$	signal on time	
$h'(t)$	denoised signal	
X	signal on frequency	
τ	time location	s
S	S-transform result	
S'	nverse S-transform result	

Step 1: Compute the S-transform.

Input a coda wave signal, compute its S-transform amplitude spectrum at different scales and orientations according to specified parameters. The S-transform is a time-frequency analysis technique proposed by Stockwell et.al [24, 26] and defined as Eq. 1.

$$S(\tau, f) = \int_{-\infty}^{+\infty} h(t) \underbrace{\frac{|f|}{\sqrt{2\pi}} e^{-(\tau-t)^2 f^2 / 2}}_{\text{Window}} e^{-j2\pi ft} dt \quad (1)$$

Step 2: Thresholding.

Thresholding on the S-transform amplitude spectrum, setting values below a certain threshold to zero, thus reducing the impact of noise. The function of thresholding as Eq. 2.

$$S'(\tau, f) = \begin{cases} S(\tau, f), & |S(\tau, f)| > \text{threshold} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

The threshold is determined by the following formula (48):

$$\text{threshold} = \sqrt{2 \log_2(N)}$$

Step 3: Inverse S-transform.

An inverse S-transform on the thresholded S-transform amplitude spectrum to obtain the filtered signal. The formula for inverse S-transform is Eq. 3.

$$h'(t) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} S'(\tau, f) d\tau \right\} e^{j2\pi ft} df \quad (3)$$

Basic Theory of STA STA module includes two parts: DOST (Discrete Orthogonal S-transform) and Rearrange-DOST (Rearranged Discrete Orthogonal S-Transform). The DOST [26–28] is a type of Discrete Orthogonal S-transform used for decomposing signals in the time-frequency domain and extracting local features of the signal.

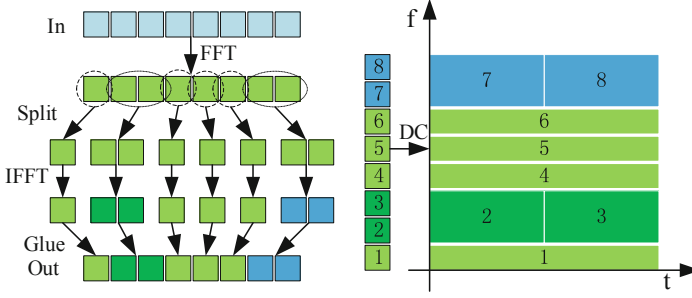


Fig. 3. The processing of DOST (left) and rearrange-DOST (right).

Rearrange-DOST is an operation that rearranges the results of the S-transform to reduce the interference of noise signals. Algorithm 1 shows the processing steps [29].

Algorithm 1 S-transform-based Time-frequency Analyzer

- 1: **Input:** denoised coda wave signals by STF W , sampling number n , signal length N , sampling interval t .
 - 2: **Output:** Time-frequency spectrogram image Datasets TD .
 - 3: **Initialize:** a 4D array TD .
 - 4: **for** each signal x **in** W **do**
 - 5: Process x by normalized and centered FFT, save it to x' ,

$$X_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N}, k = 0, 1, \dots, N-1.$$
 - 6: Calculate the bandwidth partitioning for the dost $dostbw$,

$$dostbw = \frac{2\pi}{N \Delta_t \min(|\Delta S|)}.$$
 - 7: Split the x' according to $dostbw$ and process by normalized and centered IFFT according to $dostbw$.

$$\text{Calculate } X_k = x'(dostbw), x_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k e^{i2\pi kn/N}, n = 0, 1, \dots, N-1.$$
 Save the result to the DOST coefficient $DOST$.
 - 8: Rearrange the DOST coefficients $DOST$ in a matrix form $rearrangeDOST$,

$$rearrangeDOST = DOST(dostbw, N./dostbw).$$
 - 9: Show the time-frequency spectrogram image according to $rearrangeDOST$ and save the RGB matrix to TD .
 - 10: **end for.**
 - 11: **return** TD .
-

As Fig. 3 shows [30], firstly, input the denoised signal and process it by normalized and centered FFT. Normalization is to make the transformation between the forward FFT and inverse FFT symmetric, while also ensuring the equivalence of unit energy in the time and frequency domains. Centering refers to making the result of FFT symmetric from the zero point in the frequency domain, making it more convenient for representing

real numbers. Secondly, calculate the bandwidth partitioning parameters (*dostbw*) for the DOST [26], for DOST analysis, a dyadic partitioning strategy is commonly used to divide the S-plane into a series of frequency bands with equal bandwidth, where each band's bandwidth is equal to the *dostbw*. This allows for selecting a series of rectangular regions on the S-plane and performing DOST analysis on the signal within each region to obtain time-frequency domain information. Specifically, the S-plane is divided into a group of square regions centered at the origin, with boundaries formed by dyadic partitioning from negative infinity to positive infinity, i.e., $[-2^k, 2^k] \times [-2^k, 2^k]$, where k is an integer. The signal within each region is analyzed with DOST to obtain the corresponding time-frequency domain information. Split the frequency-domain signal obtained through normalized and centered FFT to some segments according to *dostbw*, and use normalized and centered IFFT processing for each segment of the frequency-domain signal to obtain the DOST coefficients (DC). Thirdly, rearrange the DOST coefficient according to the frequency or time sequence as a pixel value of a time-frequency spectrum graph that can be represented in different colors [31]. After time-frequency domain spectral graph pixel value of all time is calculated, by composing all the pixel values into a matrix, the time-frequency domain spectral graph is obtained. It can be graphically displayed in order to visually observe the change of the signal in the time-frequency domain [27].

In general, DOST can provide more comprehensive time-frequency domain information, while rearrange DOST processing can transform it into a more easily understandable frequency-time distribution via TF-image. The TF-image can display transient and periodic information of the signal, which is an essential tool commonly used in signal processing and analysis. We analyze the coda wave signals from Stress = 1 to 10 under the same environment via STA and STF, the TF-images are shown in Fig. 4.

2.3 Feature Extractor with Residual Blocks and Classifier

From Fig. 4, the TF-images under various stresses perform significant difference, which can be seen as the key features that represent corresponding stress. We assume that the test environment is same due to the test interval of short period is short. Thus, these differences are caused only by stress changes. It is difficult to quantify this difference in time domain and frequency domain at the same time for previous works, since they acquire the features only from time domain signal or frequency domain signal. To address the limitation, we design the Feature Extractor with Residual Blocks (RB) and Classifier, which aim to extract key features from the time-frequency spectrogram and identify stress. Specially, the Feature Extractor with Residual Blocks is stacked by a set of RBs. The shortcut connection of RB [32, 33] adds the input of the Residual Block to its output, which allows the network to learn the residual mapping between the two, that makes contribution to keep detailed local features and extract fine global features [34]. The Classifier is responsible for classifying the features produced from the Feature Extractor and predicting the probability of corresponding label.

In this study, Feature Extractor with Residual Blocks includes the first convolution layer, four RBs and global average layer (GAP) as Fig. 5 shows [35]. The TF-images (size: $224 \times 224 \times 3$, RGB image) will be input into the first convolution layer for feature extraction and the number of channels will be briefly expanded from 3 to 64. Then the features will be passed into four RBs. With layers deeper, the size of feature map is

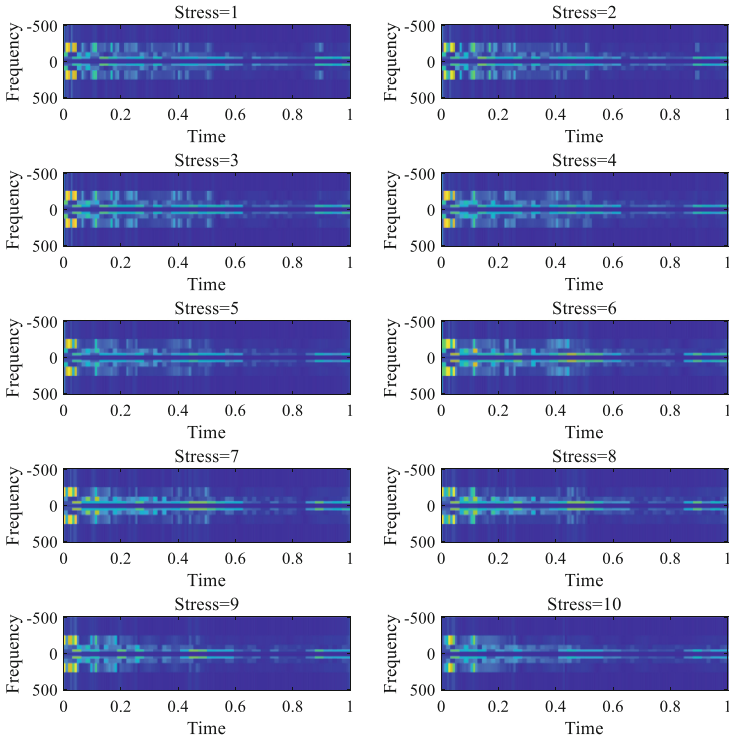


Fig. 4. The TF-images of three-point bending experiment under the same environment (the image size is 224*224*3).

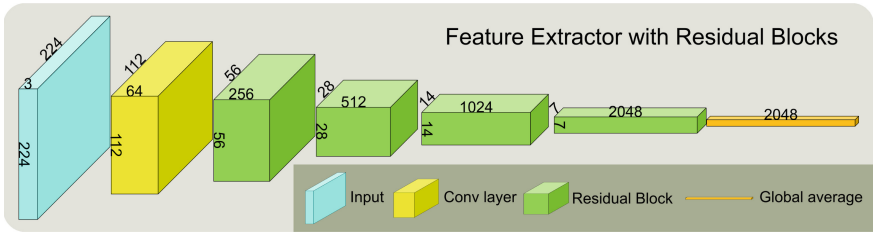


Fig. 5. The architecture of Feature Extractor.

reduced, and the channel number will be expanded. The parameters of Feature extractor and Classifier are shown in Table 1 [36, 39]. Reducing the size of feature map while extracting global features can accelerate the training process. Also, the reduced feature maps enable convolution filters to see a bigger region, which allows model to capture global features of TF-images [37]. Finally, we use 2 linear layers for classifying the extracted features into 20 stress levels.

Table 1. The parameter of Feature Extractor with residual blocks and classifier.

Layer name	Parameter	Output size
Conv	$7 \times 7, 64, \text{stride } 2$	$112 \times 112 \times 64$
Residual Block-1	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$56 \times 56 \times 256$
Residual Block-2	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$28 \times 28 \times 512$
Residual Block-3	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$14 \times 14 \times 1024$
Residual Block-4	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$7 \times 7 \times 2048$
Global average pooling	$1 \times 1 \times 2048$	
Feature classifier	20-d fc, SoftMax	

Residual Block Residual Block (RB) is the most vital element in the Feature Extractor. RB is based on the idea of skipping Blocks of convolutional layers by using shortcut connections. These shortcuts are useful for optimizing trainable parameters in back-propagation to avoid the vanishing/exploding gradients problem [38], which can help to improve the performance of extracting features.

As Fig. 6 shows, each RB consists of convolutional (Conv) layers and batch normalizations (BN) layers, Relu activation function and one shortcut connection. When a 256-d data flow is input into the RB, the number of channels will be compressed from 256 to 64 and expended from 64 to 256 in the last Conv layer so that the global features generated from the three Conv layers and the local features transmitted by the shortcut connection have the same number of channels, which is easy to add. The RB can be defined as Eq. 4.

$$Y = F(X, \{W_i\}) + X \quad (4)$$

Here X and Y are the input and output vectors of the layers considered. The function $F(X, \{W_i\})$ represents the residual mapping to be learned.

Classifier The Classifier is composed of fully connected layers, these fully connected layers perform a weighted sum of the input features and output the predicted class probabilities [40]. As Fig. 7 shows, features generated from Feature Extractor will be input to two hidden layers to calculate the weight. In the last layer of Classifier, the SoftMax function is employed to map features into possibilities of classes. Specially,

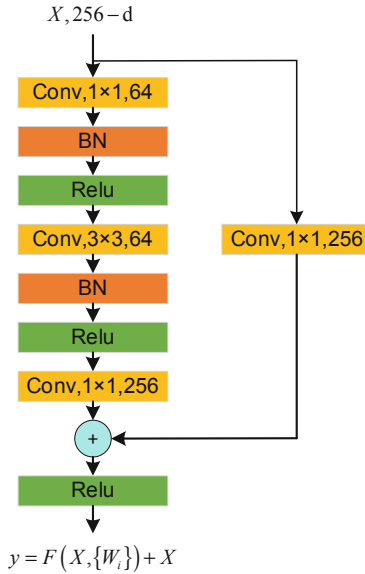


Fig. 6. The architecture of an RB example.

the SoftMax function converts the output values of multiple classes into probability distributions in the range [0, 1]. SoftMax function can be defined as Eq. 5.

$$\text{SoftMax}(z_i) = \frac{e^{z_i}}{\sum_{c=1}^c e^{z_c}} \tag{5}$$

where z_i is the output value of the i th node, and c is the number of output nodes, that is also the number of classifications.

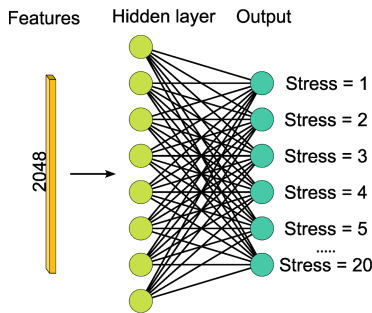


Fig. 7 The architecture of classifier.

Loss function In the iterations, the cross-entropy loss function is used to calculate the loss between true label and predicted label, and the loss value will be back-propagated to the network for updating the parameters, which leverages the model to correctly

extract local and global features from TF-images and get higher prediction accuracy. The cross-entropy loss function is defined as Eq. 6.

$$Loss = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{i,j} \log(p_{i,j}) \quad (6)$$

where $Loss$ represents the loss function, N is the sample numbers, C is the class numbers, $y_{i,j}$ presents the true label of sample i belongs to class j , $p_{i,j}$ presents the predicted label of sample i belongs to class j .

Optimizer The Adam optimizer is chosen in this study and defined as Eq. 7.

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t \\ v_t &= \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \\ \theta_{t+1} &= \theta_t - \frac{\eta}{\sqrt{v_t} + \varepsilon} m_t \end{aligned} \quad (7)$$

where, m_t represents first moment estimation, v_t represents second moment estimation, θ_t represents network parameters, g_t represents gradient, β_1 and β_2 represent exponential decay rate, ε represent represents a very small number, to prevent the case of dividing by zero.

3 Experiment

To validate the feasibility of USR-Net and address these limitations like random noise, continuous environment changes, measuring positions, and loading patterns, we design the two loading patterns experiments of reinforced concrete I-beam under non-controlled experimental environment conditions, collecting the coda wave signals of various stress values.

3.1 Experiment Design

Most coda wave measurement experiment objects are standard concrete square or cylindrical test Blocks, and the loading patterns are mostly uniaxial compression [1], but the concrete structure in the practical engineering of the section form is more diverse and the force form is more complex. Therefore, the previous experiment work is difficult to be applied in practical engineering. In order to improve the application feasibility of USR-Net in practical engineering, we choose the full-scale [41] model to observe the real response of concrete structure and the test is performed using real loads and boundary conditions. As the Fig. 8 shows, we design three-point bending and eccentric compression experiments with two loading patterns. The pulser on the one end transmits pulse wave through the I-beam under various stress and receivers on another end collect the coda waves. In this experiment, by changing the number of steel plate and the pressure value of hydraulic jack, we simulate the stationary stress changes. Besides, to test the universality, it costs plenty of time to conduct this experiment for collecting more data, which has been fully influenced by the environment changes.

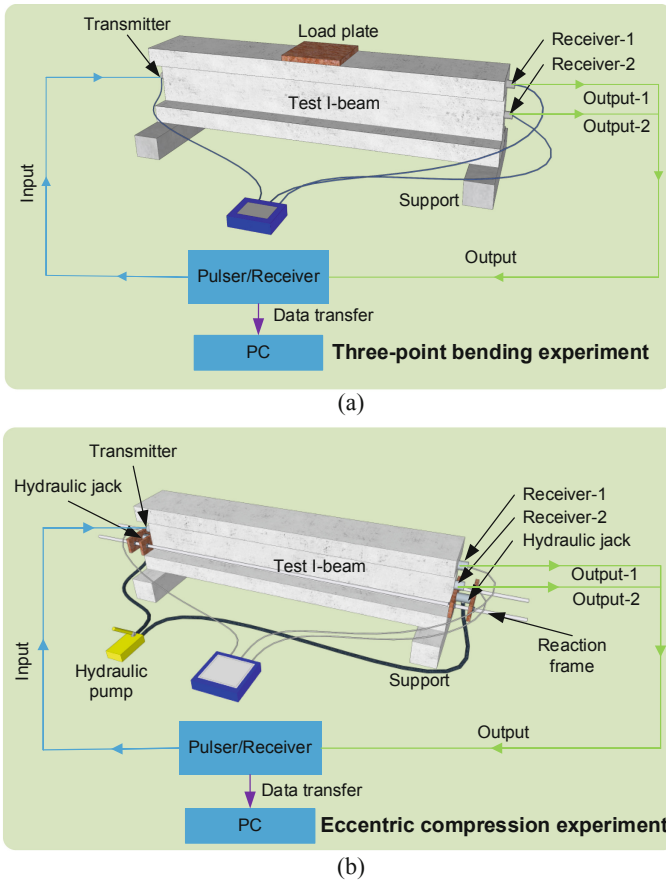


Fig. 8. Instrument layout. (a) Three-point bending ultrasonic experiment. (b) Eccentric compression ultrasonic experiment.

The response of concrete structure varies with different positions of section. In order to address the problems that traditional coda wave measurement only arrange the transducer in a specific position, we arrange the transducer of receiver-2 on the neutral axis of I-beam section and receiver-1 on the compressed and stretched part like Fig. 8. In the three-point bending experiment, the loading position is arranged the middle of I-beam to simulate the pure bending stress, and each load plate weighs 27.5 ± 0.4 kg. Similarly, we set the loading position in the down quarter of section. The loads are applied to the I-beam by the co-action of reaction frame and hydraulic jacks. Each load change is about 0.3 KN, which make sure loading smoothly changes.

3.2 Experiment Materials and Instruments

The full-scale I-beam model (the size as Fig. 9) is composed of C30-strength concrete with a mix ratio as presented in Table 2. The ultrasonic pulser and receiver device was a

nonmetallic transmittance instrument and transducer is the sandwich p-wave transducers (50 kHz).

Table 2. Mix ratio of C30 concrete.

Cement	Fine aggregate	Aggregate	Water	Fly ash	Mineral
1.00	2.50	3.85	0.55	0.22	0.16

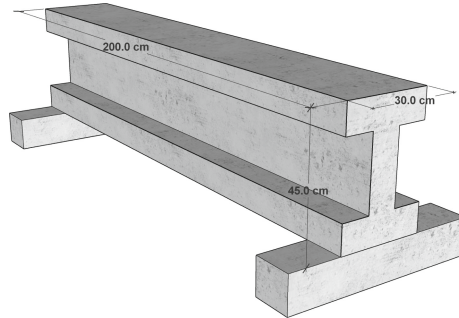


Fig. 9. Sizes of test I-beam.

Due to the signal quality related to the instrument parameters, it is essential to investigate and configure the instruments. Our team conducted an ultrasonic test of T-beam under three-point bending stress and optimized the parameters of the nonmetallic ultrasonic transmitter during the experiment before 2019. For this experiment, the parameters of non-metallic ultrasonic transmitter were selected through debugging, based on the experiences documented in the aforementioned literature. Table 3 summarizes the parameters associated with the nonmetallic ultrasonic transmitter, ensuring that the signal contains as much information as possible and transmittance receives a relatively stable signal.

Table 3. Instrument parameters.

Trigger delay [μ s]	Sampling interval [μ s]	Pulse width [μ s]	Gain [dB]	Sampling length
9999	16	5	200	1024

4 Result and Discussion

To evaluate the effectiveness of USR-Net, 656000 signal samples, which consists of 20 classes from two I-beams of the three-point bending and eccentric compression experiment, are applied to predict stress labels. After STA and STF processing, TF-images

are generated for deep learning mode. To be specific, we use 524800 TF-images for training, 65600 TF-images for validation, and 65600 TF-images for testing. To establish the related labels with actual stress values, we perform the finite element analysis of three-point bending loading experiment and eccentric compression experiment under various loadings and acquire the maximal combined stress values (see Table 4). Then, the stress labels are encoded to one-hot format to facilitate deep learning model training process. The experimental results demonstrate high prediction accuracy on our dataset. To further explore the performance of USR-Net, four indexes based on confusion matrix are calculated. Besides, we compare our model with popular non-DL algorithms, including Singular Value Decomposition (SVD), Support Vector Machine (SVM), Developed Support Vector Machine (DSVM), and Coda Wave Interferometry (CWI).

All samples were collected over a period of two years under changing environments, which can certainly verify the ability of resisting environmental interference.

Table 4. The stress of mid-span position via the finite element.

label	Max combined stress (KPa)	Shear stress (KPa)
0	118.5	63.5
1	144.1	71.6
2	169.7	79.7
3	195.3	87.8
4	220.9	95.9
5	246.5	104
6	272.1	112.1
7	297.7	120.2
8	323.3	128.3
9	348.9	136.4
10	374.5	144.5
11	400.1	152.6
12	425.7	160.7
13	451.3	168.8
14	476.9	176.9
15	502.5	185
16	528.1	193.1
17	553.7	201.2
18	579.3	209.3
19	604.9	217.4
20	630.5	225.5

4.1 Experimental Result

We train the deep learning model, Features Extractor and Classifier, with TF-images produce by applying STA and STF on signal samples. As Fig. 10 shows, with iteration increases, the accuracy of eccentric compression experiment (E-train) and three-point bending experiment (T-train) increases and achieves 99.99% accuracy on two loading mode. Even though there exist some fluctuation during training, for example, The loss for both E-train and T-train increases dramatically around the 20th iteration, the prediction accuracy converges to stable value.

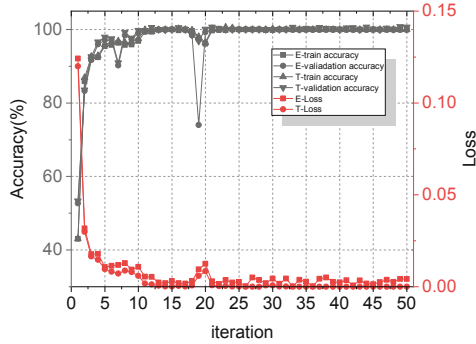


Fig. 10. The classification accuracy (left axis) and loss value (right axis) in the iteration of two experiments, where E represents the eccentric compression experiment and T represents the three-point bending experiment.

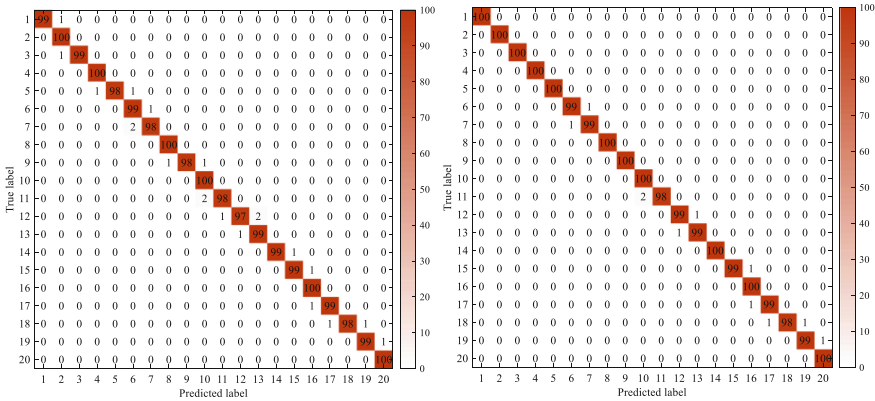


Fig. 11. The confusion matrix of three-point bending (left) and eccentric compression stress identification accuracy (right).

However, accuracy is not enough to reliably reflect the performance of the USR-Net. To Further evaluate the result, we build confusion matrix to calculate four indexes: accuracy, precision, recall and specificity. Accuracy represents the proportion of all correctly judged results of classification model in the total observed value. Precision

represents the proportion of the model prediction that is correct in all the results where the model prediction is positive. Recall represents the proportion of the model prediction that is correct in all the results where the true value is positive. Specificity represents the proportion of all results in which the true value is negative that the model's prediction is correct. Figure 11 demonstrates the confusion matrix of our model. According to the definition of four indexes as Eqs. 8–11, we can calculate accuracy, recall, precision, and specificity (see Fig. 12).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (9)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (10)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (11)$$

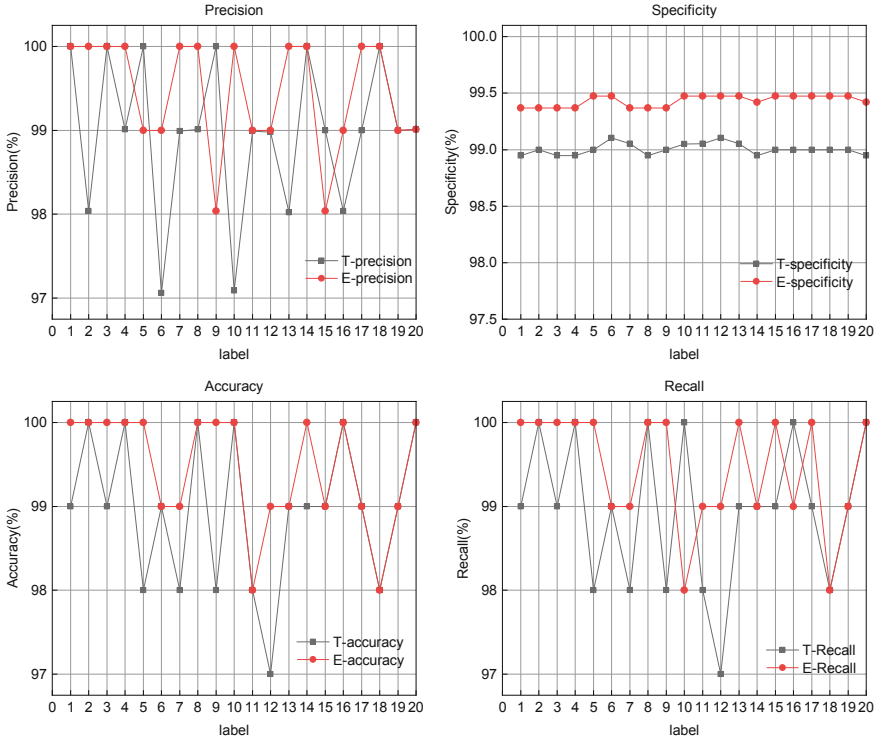


Fig. 12. The accuracy, recall, precision and specificity of USR-Net in three-point bending (T represents) and eccentric compression experiments (E represents).

4.2 Discussion

Comparison We compared the performance of USR-Net with the non-DL algorithms, but the non-DL algorithms evaluation index varies with the characteristic of them, such as the SVD utilizes correlation coefficient [42] to evaluate and CWI uses goodness of fit [43]. Specially, the correlation coefficient is a statistical measurement used to assess the strength of the linear relationship between two variables, and goodness of fit is a statistical measurement used in statistics to assess the fit degree of a statistical model. For the purpose of comparison, these evaluation indexes are all defined with the same physical meaning as classification accuracy and uniformly referred to as classification accuracy in the subsequent description. In the comparison, we use all samples to analyze, and calculate the average accuracy which is shown in Fig. 13. Comparing with traditional non-DL algorithms, the USR-Net proposed in this study has a higher average identification accuracy for its better feature extraction capability. For instance, USR-Net improve 17.5% and 15.9% of eccentric compression and three-point bending experiment than the popular non-DL algorithm CWI, which only gets 85% and 86.2% on our dataset.

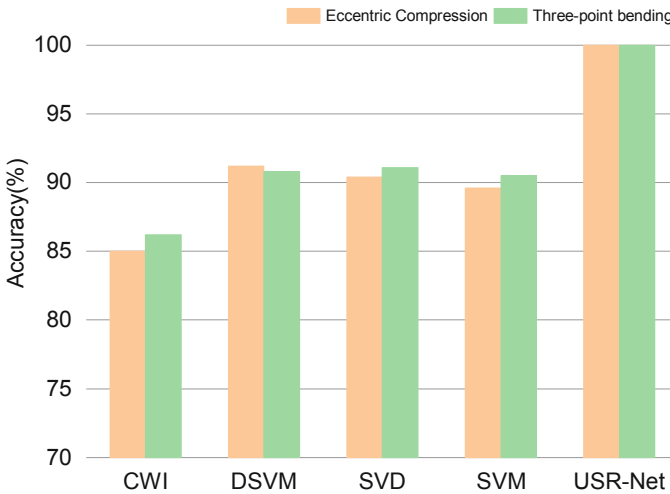


Fig. 13. Comparison with non-DL algorithms.

The verification of Kaiser effect As Fig. 10 shows, with the increase of number of iterations, the accuracy rises gradually, but when the numbers of iteration are 8 and 18, both the train accuracy and the validation accuracy suddenly decreased, we argued that is because the Kaiser effect [44] which refers to the phenomenon in which materials exhibit a characteristic pattern of stress-strain behavior during cyclic loading. Due to the existence of Kaiser effect, the concrete structure has a stress-memory effect, when a cycle of loading is complete and immediately unloading, the stress of concrete will not change and keep the state for a long time. Thus, another cycle of loading even begins, and the collected coda wave signals still present the state of last loading.

To verify the idea of Kaiser effect, we use the signals of loading level is 20 (the red label in Fig. 14) in the last loading cycle and the signals of loading level is 1 (the green

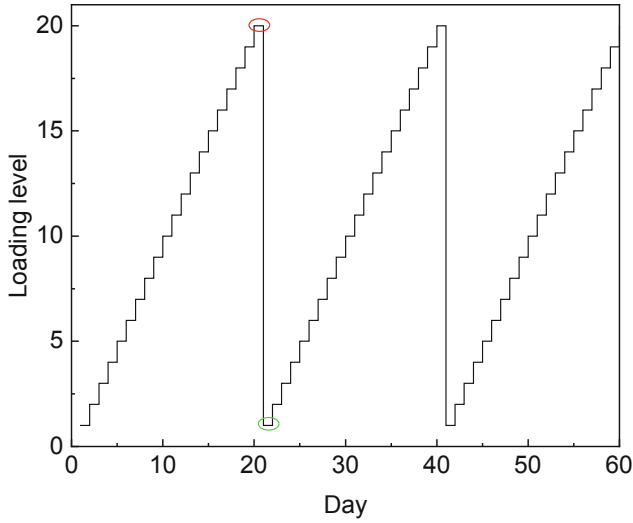


Fig. 14. Loading sequence.

label in Fig. 14) in the next loading level cycle to identify via USR-Net and non-DL algorithms. The result is shown in Table 5, we can see the results of SVD and CWI are not like the others, because the SVD and CWI mainly establish the mathematical linear model between the specific features and stress. Due to the interference from various factors, they cannot give the accuracy prediction like these labeled values. When identifying the label of 20, USR-Net, SVM, and DSVM can output the value corresponding to the label, moreover, SVD and CWI also give the relatively similar results that is relatively close to the actual stress. However, when identifying the label of 1 in the next cycle, the predict result of non-DL algorithms is similar to the label of 20 in the last cycle. Therefore, the Kaiser effect has significant influence in the coda wave signals in the coda wave measurement.

Table 5 The verify result of Kaiser effect via USR-Net and non-DL algorithms.

True label	Predicted label (convert to stress value, KPa)				
	USR-Net	SVD	CWI	SVM	DSVM
20 (630.5)	630.5	626.3	600.4	630.5	630.5
1 (144.1)	144.1	598.5	588.2	604.9	604.9

The performance of anti-noise We test the performance of anti-noise via adding random Gaussian white noise to raw signal and compare the identification accuracy of USR-Net and non-DL algorithms under different SNR. In the process of test, the TF-images generated from signals polluted by noise are input into USR-Net, and these signals are analyzed by non-DL algorithms, the accuracy of all methods is shown in Fig. 15.

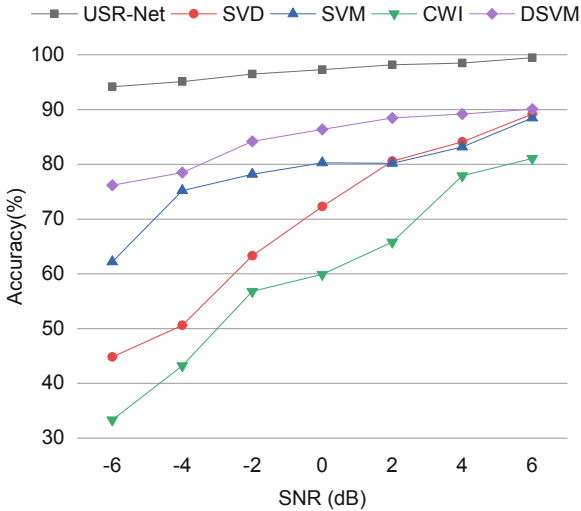


Fig. 15. The accuracy varies with SNR.

From Fig. 14, the USR-Net had the highest accuracy than other methods. Under SNR = -6 dB, the accuracy of the USR-Net is still 94.2%, and the accuracy of SVD, SVM, CWI and DSVM are 44.8%, 62.2%, 33.3% and 76.2. The filter of USR-Net removes some noise in the signal processing, and after learning step by step, the deep learning model can correctly extract valuable features from noised TF-images. Besides, USR-Net has a multi-scale respective field and multi-channel feature fusion, which can capture richer features of TF-images. Therefore, USR-Net has the best classification effect in low-SNR environments and keep the good performance on anti-noise testing. Even when the SNR is more than 2 dB, USR-Net shows a distinctive identification accuracy rate of 98.5% comparing with other coda wave measurement methods.

5 Conclusion

In response to the limitations in measuring reinforced concrete stress via ultrasonic coda wave, this study proposes the USR-Net, a method for combining the S-transform and deep learning model. USR-Net employs S-transform-based Time-frequency Filter to eliminate random noise and S-transform-based Time-frequency Analyzer to extract two-dimensional feature of the coda wave signal for generating TF-image. Then, Feature Extractor with Residual Blocks is applied to extract detailed valuable features from

the TF-images, which utilizes shortcut connection to maintain local features while feature map size reduces. Finally, Classifier takes use of linear layers to establish local and global features with actual stress labels. To evaluate the performance of USR-Net, we designed three-point bending and eccentric compression experiments and achieves the identification accuracy of 99.99% in two load patterns under continuous changing environment. Further, we explore the result of confusion matrix and use four indexes, accuracy, recall, precision, and specificity, to verify our model. Additionally, we compare USR-Net with some popular non-DL algorithms, which demonstrates USR-Net can achieves better results than other methods even under different noise condition. We also discuss about the Kaiser effect in coda wave stress measurement to explain the exceptional loss fluctuation. The good anti-noise performance and better identification accuracy of USR-Net enable it has application prospects in practical engineering. The USR-Net highlights the potential of deep learning algorithms in monitoring the health of reinforced concrete structures and contributes to the development of more precise and effective stress measurement techniques.

References

1. Jiang H, Zhang J, Jiang R, F.: Stress evaluation for rocks and structural concrete members through ultrasonic wave analysis. *Journal of Materials in Civil Engineering* 29(10), 04017172 (2017).
2. Castaneda, D. I, F.: New field-testing procedure for measuring residual stress in plain concrete pavements and structures. M.Sc. thesis, Univ. of Illinois at Urbana-Champaign, Champaign, IL (2010).
3. Herraiz, M., Espinosa, F.: Coda waves: A review. *PAGEOPH* 125, 499–577 (1987).
4. T. Planès, E. Larose, F.: A review of ultrasonic Coda Wave Interferometry in concrete. *Cement and Concrete Research* 53, 248–255 (2013).
5. Liu, Y., Guo, S., Xu, Y., & Xia, F.: Real-time monitoring of stresschanges in concrete using coda wave interferometry. *Sensors* 18(2), 626 (2018).
6. Snieder, R., Greötz, A., Douma, H., & Scales, F.: Coda wave interferometry for estimating nonlinear behavior in seismic velocity. *Science* 295(5559), 2253-2255 (2003).
7. Payan, F.: Coda wave interference for extracting the third-order nonlinear coefficient of concrete materials: functional relationship between the relative change rate of wave velocity and stress. *Journal of Nondestructive Evaluation* 28(1), 1-10 (2009).
8. Payan, C., Garnier, V., & Moysan, F.: Determination of third order elastic constants in a complex solid applying coda wave interferometry. *Applied Physics Letters* 94(1), 011904 (2009).
9. Hafiz, A., Schumacher, F.: Monitoring of Stresses in Concrete Using Ultrasonic Coda Wave Comparison Technique. *J Nondestruct Eval* 37, 73 (2018).
10. Clauß, F.; Epple, N.; Ahrens, M.A.; Niederleithinger, E.; Mark, F.: Correlation of Load-Bearing Behavior of Reinforced Concrete Members and Velocity Changes of Coda Waves. *Materials* 2022(15), 738 (2022).
11. Wang, X., Chakraborty, J. & Niederleithinger, F.: Noise Reduction for Improvement of Ultrasonic Monitoring Using Coda Wave Interferometry on a Real Bridge. *J Nondestruct Eval* 40, 14 (2021).
12. Grêt A, Snieder R, Scales J, F.: Time-lapse monitoring of rock properties with coda wave interferometry. *Journal of Geophysical Research: Solid Earth* (B3), 111 (2006).

13. Serra M, Festa G, Vassallo M, F.: Damage detection in elastic properties of masonry bridges using coda wave interferometry. *Structural Control and Health Monitoring* 24(10), e1976 (2017).
14. Onur Avci, Osama Abdeljaber, Serkan Kiranyaz, Mohammed Hussein, Moncef Gabbouj, Daniel J. Inman, F.: A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. *Mechanical Systems and Signal Processing* 147, 107077 (2021).
15. Zhang J, Han B, Xie H B, F.: Correlation between coda wave and stresses in uni-axial compression concrete. *Applied Sciences* 8(9), 1609 (2018).
16. M. Schimmel and J. Gallart, F.: The inverse S-transform in filters with time-frequency localization. *IEEE Transactions on Signal Processing* 53(11), 4417–4422 (2005).
17. N. Liu, J. Gao, B. Zhang, F. Li and Q. Wang, F.: Time–Frequency Analysis of Seismic Data Using a Three Parameters S Transform. *IEEE Geoscience and Remote Sensing Letters* 15(1), 142–146 (2018).
18. Das M K, Ari S, F.: Analysis of ECG signal denoising method based on S-transform. *Irbm* 34(6), 362–370 (2013).
19. Parolai S, F.: Denoising of seismograms using the S transform. *Bulletin of the Seismological Society of America* 99(1), 226–234 (2009).
20. S. S. Sahu, G. Panda and N. V. George, F.: An Improved S-Transform for Time-Frequency Analysis. 2009 IEEE International Advance Computing Conference, Patiala, India, 315–319 (2009).
21. N. Liu, J. Gao, B. Zhang, Q. Wang and X. Jiang, F.: Self-Adaptive Generalized S-Transform and Its Application in Seismic Time–Frequency Analysis. *IEEE Transactions on Geoscience and Remote Sensing* 57(10), 7849–7859 (2019).
22. E. Goceri, F.: Analysis of Deep Networks with Residual Blocks and Different Activation Functions: Classification of Skin Diseases. 2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA), Istanbul, Turkey, 1–6 (2019).
23. Sarwinda D, Paradisa R H, Bustamam A, F.: Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. *Procedia Computer Science* 179, 423–431 (2021).
24. Stockwell, R.G., Mansinha, L., Lowe, R.P., F.: Localization of the complex spectrum: the s transform. *IEEE Trans. Signal Process* 44(4), 998–1001 (1996).
25. Beuter C, Oleskovicz M. F.: S-transform: from main concepts to some power quality applications. *IET Signal Processing* 14(3), 115–123 (2020).
26. Stockwell, R.G., F.: A basis for efficient representation of the s-transform, *Digit. Signal Process.* 17(1), 371–393 (2007).
27. 28. Battisti, U., Riba, F.: Window-dependent bases for efficient representations of the stockwell transform. *Appl. Comput. Harmon. Anal.* 40(2), 292–320 (2016).
28. Wang, Y. F.: Efficient Stockwell transform with applications to image processing. PhD thesis, UWSpace, University of Waterloo (2011).
29. 30. Beuter, C. and Oleskovicz, F.: S-transform: from main concepts to some power quality applications. *IET Signal Process.* 14, 115–123 (2020).
30. Reddy, M.J.B., Raghupathy, R.K., Venkatesh, K.P., F.: Power quality analysis using discrete orthogonal s-transform (dost). *Digit. Signal Process.* 23(2), 616–626 (2013).
31. Wang, Y. Efficient Stockwell transform with applications to image processing. PhD thesis, UWSpace, University of Waterloo (2011).
32. B. Li and Y. He., F.: An Improved ResNet Based on the Adjustable Shortcut Connections. *IEEE Access* 6, 18967–18974 (2018).
33. Zhang T, Waqas M, Liu Z, F.: A fusing framework of shortcut convolutional neural networks. *Information Sciences* 579, 685–699 (2021).

34. 37. Wu, S., Zhong, S. & Liu, F.: Deep residual learning for image steganalysis. *Multimed Tools Appl* 77, 10437–10453 (2018).
35. Cai P, Sur S, F.: MilliPCD: Beyond Traditional Vision Indoor Point Cloud Generation via Handheld Millimeter-Wave Devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6(4): 1–24 (2023).
36. He K, Zhang X, Ren S, F.: Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778 (2016).
37. Wen, L., Li, X. & Gao, L, F.: A transfer convolutional neural network for fault diagnosis based on ResNet-50. *Neural Comput & Applic* 32, 6111–6124 (2020).
38. Shafiq, M.; Gu, Z, F.: Deep Residual Learning for Image Recognition: A Survey. *Appl. Sci.* 12, 8972 (2022).
39. A. Jafar and L. Myungho, F.: Hyperparameter Optimization for Deep Residual Learning in Image Classification. *2020 IEEE International Conference on Autonomic Computing and Self-Organizing Systems Companion (ACSOS-C)*, Washington, DC, USA, 2020, 24–29 (2020).
40. Basha S H S, Dubey S R, Pulabaigari V, F.: Impact of fully connected layers on performance of convolutional neural networks for image classification. *Neurocomputing* 378: 112–119 (2020).
41. Bian X, Jiang H, Cheng C, F.: Full-scale model testing on a ballastless high-speed railway under simulated train moving loads. *Soil Dynamics and Earthquake Engineering* 66: 368–384 (2014).
42. Uchida, Y., Kanade, T., Shiozawa, D. F.: Thermoelastic Stress Measurement Using SVD Thermo-Component Analysis. *Exp Mech* 63, 337–347 (2023).
43. Stähler S C, Sens-Schönfelder C, Niederleithinger E., F.: Monitoring stress changes in a concrete bridge with coda wave interferometry. *The Journal of the Acoustical Society of America* 129(4), 1945–1952 (2011).
44. Shokouhi, P., Zoega, A., Wigggenhauser, F.: Nondestructive investigation of stress-induced damage in concrete. *Adv. Civil Eng*, 1–9 (2010).
45. Donoho D. De-noising by soft-thresholding. *IEEE Trans Inform Theory*, 41:613–27 (1995).



Determination of the Presence or Absence of Defect for Laser Ultrasonic Visualization Testing Using Transfer Learning

Takahiro Saitoh¹(✉), Yuto Kuwabara¹, and Tsuyoshi Kato²

¹ Department of Civil and Environmental Engineering, Gunma University, 1-5-1, Tenjin, Kiryu, Gunma 376-8515, Japan

t-saitoh@gunma-u.ac.jp

² Department of Informatics, Gunma University, Aramaki, Maebashi, Gunma 371-0044, Japan

Abstract. LUVT (Laser Ultrasonic Visualization Testing) is known as an effective experimental technique for detecting surface defects. It can visualize ultrasonic wave propagation on a defect surface. Inspectors typically determine the presence or absence of defects by visually observing the propagation of ultrasonic waves and checking for the generation of scattered waves. If AI is used to automate this visual judgment made by inspectors, it could reduce the workload of the inspectorate. In this study, we attempt to automate the LUVT inspection by incorporating transfer learning into deep learning, which is the basis of AI. The effectiveness and other aspects of the proposed method are discussed by presenting several defect detection results.

Keywords: Laser ultrasonic testing · Transfer learning · Deep learning · Ultrasonic non-destructive testing

1 Introduction

LUVT (Laser Ultrasonic Visualization Testing) [1] is one of the ultrasonic non-destructive inspection methods [2, 3] to find defects in a specimen by visualizing ultrasonic wave propagation on the surface of a specimen. In general, LUVT can produce time-series ultrasonic propagation images, as shown in the next section. In LUVT, the presence or absence of defects is determined by LUVT inspectors confirming the scattered waves from defects. Therefore, if their scattered wave verification process of LUVT inspectors could be automated, it would save inspection time required for LUVT.

In recent years, the use of artificial intelligence (AI) has attracted attention as a means of automating non-destructive inspections. In fact, the concept of NDE4.0 [4] has been proposed, and attempts to apply various information technologies to non-destructive inspection have begun. Meng et al. [5] used a

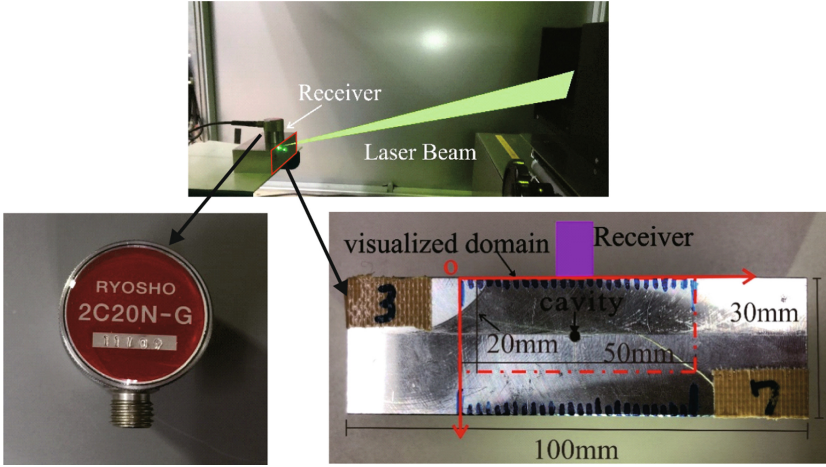


Fig. 1. LUVT experimental setup.

deep convolutional neural network to identify a defect from ultrasonic waveforms. Saitoh et al. [6] utilized deep learning to identify the presence or absence, and type of a defect in images that are equivalent to those obtained by LUVT. These images were numerically prepared using the time-domain boundary element method [7]. Ye et al. [8] also attempted to identify a defect in images such as those obtained with LUVT. Tsuzuki et al. [9] uses a convolutional neural network to determine a defect in images obtained by LUVT experiments. The key to improving the accuracy of deep neural networks is how to create an appropriate layered structure for the problem to be solved.

In this research, we aim to develop an automatic LUVT system that uses AI to determine the presence or absence of a defect in specimens, eliminating the need for inspectors. The deep learning and transfer learning [10, 11] are utilized in order to create the AI for LUVT. Transfer learning is a popular approach in deep learning that involves utilizing pre-trained weights from published convolutional neural network (CNN) models, such as VGG16 [12], to improve the performance of a new model on a related task. After training the AI model on LUVT images, we evaluate its performance by testing it on a set of previously untrained LUVT images. Finally, we check whether the AI can accurately determine the presence or absence of a defect in these unknown images, in order to confirm the validity of our proposed method. Note that the deep learning used in this study does not take into account the time axis.

2 LUVT Setup

LUVT setup is discussed in this section. The upward in Fig. 1 shows the practical LUVT situation for an isotropic specimen with or without a penetrate cavity defect. In this LUVT experiment, a 2MHz transducer with 2cm diameter positioned on the top center of the specimen, as shown on the right side of Fig. 1,

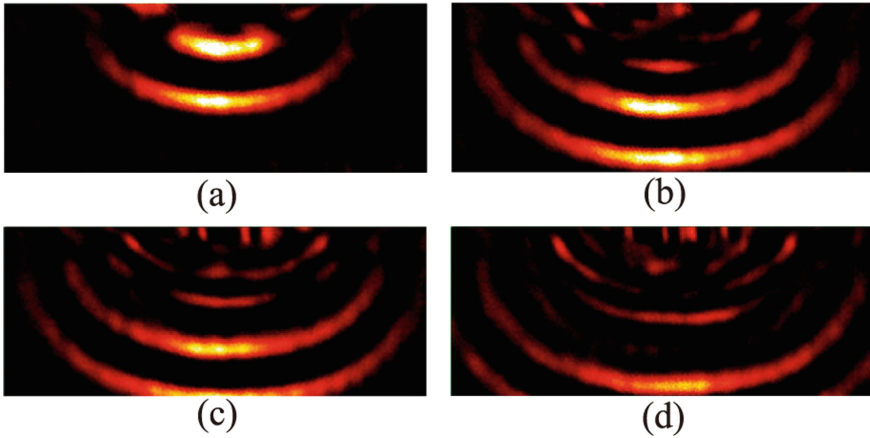


Fig. 2. Time-variations of ultrasonic wave propagation on the surface of the isotropic specimen without a cavity.

is used to transmit ultrasonic waves. The horizontal and vertical lengths of an isotropic specimen are 100mm and 30mm, respectively. The red rectangular area in Fig. 1 indicates the LUVT visualization area, which allows for the visualization of ultrasonic wave propagation in that specific region of the specimen. A 2mm diameter cavity is located within this visualization area.

Figures 2 and 3 show examples of LUVT results for this isotropic specimen without and with a defect, respectively. As shown in Fig. 2, the ultrasonic waves propagate in the aluminum surface and no scattered waves can be seen. However, for the case with a defect in Fig. 3, scattered waves can be seen in Figs. 3(c)-(d). The inspectors can judge whether the specimen has a defect or not from scattered wave generation like this. Automating this judgement process can reduce the effort required by inspectors and increase efficiency.

However, there are two problems for establishing the automation process. One is whether AI can determine the presence or absence of a defect by focusing on the generation of scattered waves as well as LUVT inspectors. The second problem concerns the preparation of training data. Creating an AI typically requires a large amount of training data. Therefore, in this case, numerous LUVT experiments need to be performed for a number of specimens, with a variety of artificial defect prepared for them. However, preparing the necessary specimens and conducting LUVT experiments can be time-consuming and require a significant amount of effort. To overcome these two difficulties, a numerical technique are transfer learning are utilized in this work.

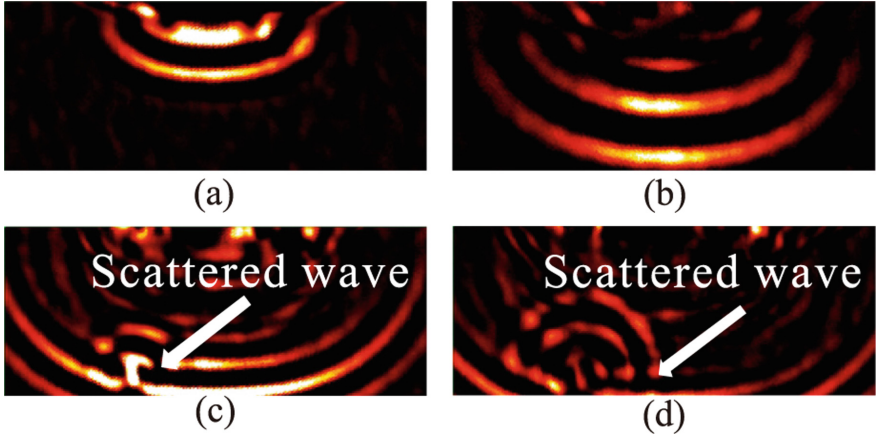


Fig. 3. Time-variations of ultrasonic wave propagation on the surface of the isotropic specimen with a penetrate cavity.

3 FDTD Simulation for Ultrasonic Wave Propagation

Numerical analysis techniques are a potent tool for investigating physical phenomena, just as experiments are. Generally, the finite element method (FEM), the finite difference time-domain method (FDTD) and the boundary element method (BEM) are used for ultrasonic wave simulation. The BEM [13–15] is known as an effective numerical technique for wave analysis. However, the BEM requires more computation time compared to the other two methods [16]. Therefore, the BEM may not always be the most efficient method for generating the large amount of training data required for AI development. On the other hand, while the FDTD method [17] may be less accurate than the FEM [18] and BEM, it offers the advantage of allowing numerical analyses to be conducted in a relatively short amount of time. Although the merits and demerits of the BEM and FDTD methods mentioned here depend on the tuning ability of the program code, in this study, the FDTD method is used to simulate ultrasonic wave propagation and scattering. The ultrasonic wave propagation simulation is to investigate whether AI can detect a defect from the generation of scattered waves. The objective of the ultrasonic wave propagation simulation is to investigate the capability of AI to detect defects based on the scattered wave generation in this work.

3.1 FDTD Formulation for 2-D Elastodynamics

Even if FDTD is used, it requires a significant amount of computational time to perform a large number of 3-D calculations. Additionally, since this study does not consider the effects of ultrasonic on the bottom and side surfaces, the influence of 3-D shapes on the wave field is relatively small. Therefore, in this

study, ultrasonic wave propagation simulations are performed using a 2-D FDTD method. In general, the ultrasonic waves exhibit the properties of elastic waves in solids. Therefore, to simulate the ultrasonic waves, it is necessary to solve the wave equation for solids [19]. The displacement $u_i(\mathbf{x}, t)$ of the ultrasonic waves at time t satisfies both the elastic wave equation and the constitutive equation as follows [20]:

$$\rho \frac{\partial^2}{\partial t^2} \mathbf{u}(\mathbf{x}, t) = \nabla \cdot \mathbf{T}(\mathbf{x}, t) + \mathbf{f}(\mathbf{x}, t) \quad (1)$$

$$\mathbf{s} \cdot \mathbf{T}(\mathbf{x}, t) = \frac{1}{2} (\nabla \mathbf{u}(\mathbf{x}, t) + \{\nabla \mathbf{u}(\mathbf{x}, t)\}^T) \quad (2)$$

where ∇ is the Auld's operator, ρ is the density, \mathbf{T} is the stress, and \mathbf{f} is the body force. The superscript symbol T shows the transposition. \mathbf{s} is the fourth-order tensor and elastic compliance, and the inverse is the elastic constant tensor \mathbf{c} . In this study, for simplicity, the body force F is set to zero and the two-dimensional in-plane wavefield of (x_1, x_3) in the Cartesian coordinate system (x_1, x_2, x_3) is considered. In this situation, Eq.(1) is shown as follows:

$$\rho \frac{\partial}{\partial t} \begin{pmatrix} \dot{u}_1 \\ \dot{u}_3 \end{pmatrix} = \begin{pmatrix} \frac{\partial T_1}{\partial x_1} + \frac{\partial T_5}{\partial x_3} \\ \frac{\partial T_3}{\partial x_3} + \frac{\partial T_5}{\partial x_1} \end{pmatrix} \quad (3)$$

where the stresses T_1 and T_3 are the vertical stresses in the x_1 and x_3 directions, respectively, and T_5 is the shear stress. The symbol $(\dot{\quad})$ indicates the partial derivative with respect to time t . In addition, Eq.(2) is rewritten as follows:

$$\frac{\partial}{\partial t} \begin{pmatrix} T_1 \\ T_3 \\ T_5 \end{pmatrix} = \begin{pmatrix} C_{11} & C_{13} & 0 \\ C_{31} & C_{33} & 0 \\ 0 & 0 & C_{55} \end{pmatrix} \begin{pmatrix} \frac{\partial u_1}{\partial x_1} \\ \frac{\partial u_3}{\partial x_3} \\ \frac{\partial u_3}{\partial x_1} + \frac{\partial u_1}{\partial x_3} \end{pmatrix} \quad (4)$$

Note that the elastic constant tensor c_{ijkl} is represented by the Voigt Notation $C_{\alpha\beta}$ ($\alpha, \beta = 1, \dots, 6$) [20] in this equation. Eqs.(3) and (4) are discretized for the time and space using well known leapfrog algorithm. The detail algorithm can be skipped in this paper.

3.2 Learning Data Created by FDTD

This section presents some of the learning data generated by FDTD. In the FDTD simulation, the corresponding area, 100mm \times 30mm, on the lower right in Fig. 1 is considered as an analysis model.

The density ρ is given by $\rho = 7850.0 \text{ kg/m}^3$. The elastic constants C_{11} , $C_{13} = C_{31}$, C_{33} , and C_{55} required for the calculation are set as $C_{11} = 282.7 \text{ GPa}$,

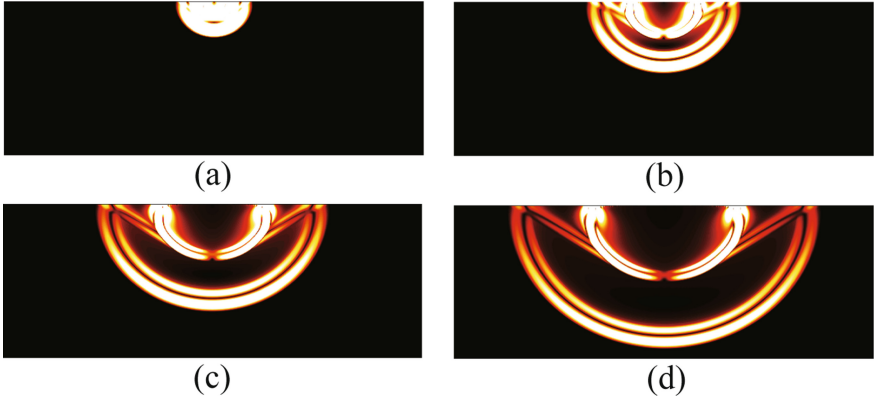


Fig. 4. Time-variations of ultrasonic wave propagation obtained by FDTD at $f^{\text{in}} = 2\text{MHz}$ for the case without a defect.

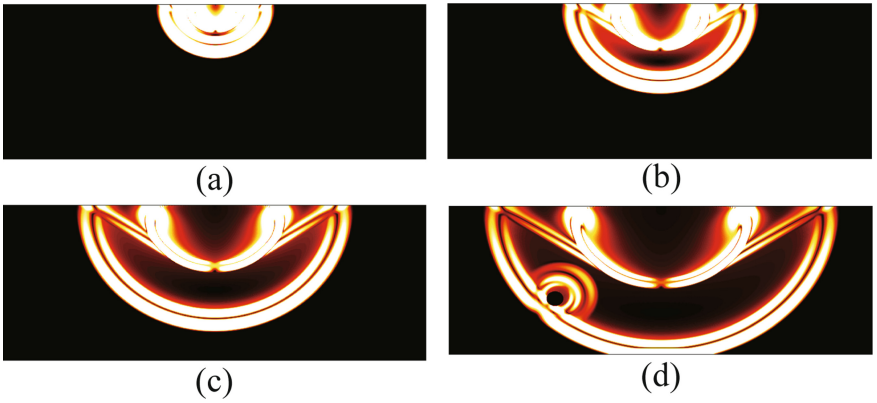


Fig. 5. Time-variations of ultrasonic wave propagation obtained by FDTD at $f^{\text{in}} = 2\text{MHz}$ for the case with a defect.

$C_{13} = C_{31} = 121.2$ GPa, $C_{33} = 282.7$ GPa, and $C_{55} = 80.8$ GPa, respectively, as isotropic parameters. As an incident source,

$$\dot{u}_3(\mathbf{x}, t) = \begin{cases} \sin\left(\frac{2\pi}{T}t\right) & \text{for } 0 \leq t \leq T \\ 0 & \text{for otherwise} \end{cases} \quad (5)$$

is given at the top center as the boundary condition of this simulation. In Eq. (5), T is the incident wave period and the central angular frequency is given by $\omega = 2\pi f = 2\pi/T$ using this T and the frequency f . The traction free boundary condition is imposed on the other boundary surface of the isotropic specimen.

Figures 4 and 5 show the time-variations of ultrasonic wave propagation obtained by FDTD for the case without and with a defect. The total wave

fields $|\mathbf{u}|$ are plotted in Figs. 4 and 5. In Fig. 5, a defect exists in the lower left of the analysis domain. As shown in Fig. 4, the ultrasonic wave is excited at the center of the top surface. The excited ultrasonic waves propagate in concentric circles pattern as shown in Fig. 4(b)-(d) as well as in Fig. 5(b) and (c) because isotropic parameters are given in this analysis. The incident ultrasonic waves are scattered by a defect in the case with a defect as shown in Fig. 5.

As mentioned in Sect. 2, the ability of the AI to recognise these scattered waves is key to the development of automatic LUVT, which is the purpose of this work. A lot of figures such as Figs. 4 and 5 are created by FDTD, changing the defect position to prepare learning data. By providing anisotropic parameters in Eq. 4, it is also possible to calculate ultrasonic waves in an anisotropic elastic solid [21]. 11353 images created by FDTD are used in the CNN training.

4 CNN and Transfer Learning

As mentioned before, CNN (convolutional neural network) has gained significant attention in recent years. In general, CNN is highly effective when the training data are images. The CNN-trained AI can detect features in images by itself. In this work, actual LUVT images, as shown in Figs. 2, 3, and images created by FDTD, as shown in Figs. 4 and 5, are used as training data for CNN. In all experiments, the ultrasonic transducer used is the one described in Sect. 2, and its position is fixed at the center of the upper surface. In deep CNN, the weights in the neural network are determined using the back propagation method [10]. In recent years, several deep neural network structures, along with corresponding training data and weights, have been made available for various object recognition tasks [12, 22]. As CNN has been thoroughly examined in many literatures, we will not go into its detail in this paper.

5 CNN Results

Some results obtained by CNN and transfer learning are shown in this section. Nvidia GPU Geforce3090RTX with 24GB memory is utilized for CNN calculations.

5.1 Scattered Wave Detection Using CNN

First, we investigate whether CNN can effectively detect scattered waves. The training data for CNN utilized in this work are generated using the FDTD, as detailed in Sect. 4. The CNN used here does not include experimental data. Typically, CNN classification problems will provide an categorized output if the input image is correct. Therefore, CNN is sometimes referred to as black boxes. However, some practical methods have been proposed to address the issue of CNN being treated as black boxes. One such method is Grad-CAM [23], which enables us to investigate how CNN arrive at their predictions. Grad-CAM

calculates the impact on each classification probability when a small modification is made to a specific part of the image. If the amount of change is large, it can be inferred that the modified part of the image has a considerable impact on the classification, and Grad-CAM can visualize the machine’s focus location in the image by displaying the magnitude of this change as a heat map on the corresponding location in the input image.

Figure 6 shows results of the GradCAM visualization for a defect detection using AI, which is trained on image data generated by the FDTD in Sect. 3. The images obtained by FDTD do not contain uncertainties such as noise. Therefore, they are convenient for the problem verification of whether AI can detect scattered waves. The heatmap in Fig. 6 shows the areas in the image that the AI has identified as defective. It should be noted that in this case, the defect is located in the bottom right-hand corner of each image. The scattered waves do not exist in Fig. 6(a) and (b), as they only show the incident waves without any scattering from a defect. Further investigation is necessary to determine why large values are observed on both sides of the incident wavefront when there is no defect as shown in Fig. 6(a) and (b).

On the other hand, the scattered waves are generated by the interaction between the incident wave and a defect in Fig. 6(c) and (d). In Fig. 6(c) and (d), the heat map values show large values at the vicinity of the generated scattered waves. Therefore, it can be observed that the created AI can capture the generated scattered waves as well as the LUVT inspectors. The computation time required for creating the training image data using the FDTD method was approximately 224 h.

Table 1. CNN architecture used in Sect. 5.2.

Layer(type)	Output shape	Param.	
Conv2D	(None,224,224,64)	640	
Conv2D-1	(None,224,224,32)	18464	
MaxPool.2D	(None,112,112,32)	0	
Conv2D-2	(None,110,110,16)	4624	Total params.: 6446642
MaxPool.2D-1	(None,56,56,16)	0	Trainable params.: 6446642
Dropout	(None,56,56,16)	0	Non-trainable params:0
Flatten	(None, 50176)	0	
Dense	(None, 128)	6422656	
Dense-1	(None, 2)	258	

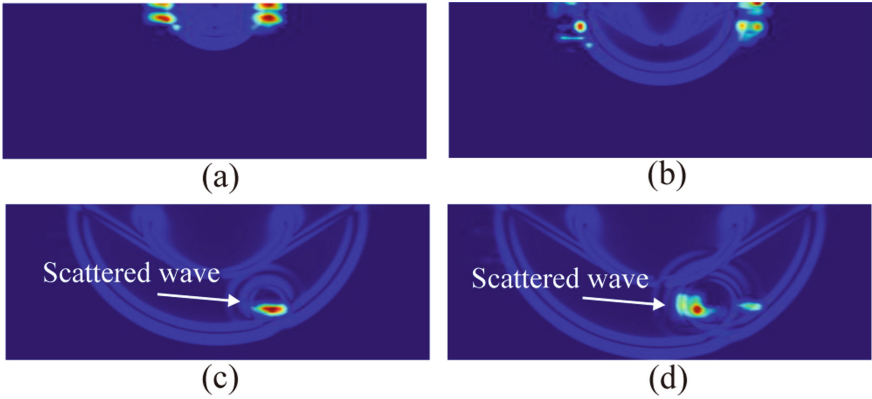


Fig. 6. Feature extraction results using Grad-CAM for a defect.

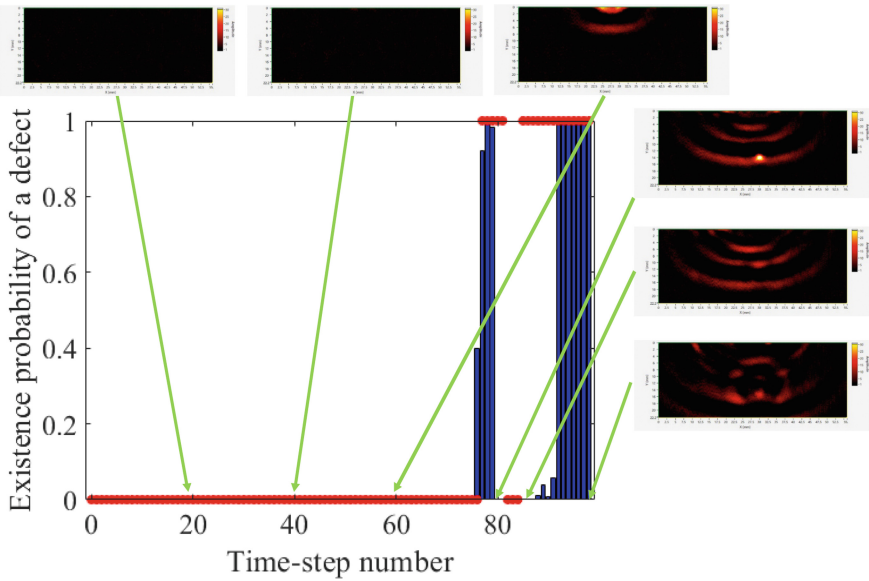


Fig. 7. Probability of a defect presence determined by the CNN architecture defined in Table 1.

5.2 Automatic Defect Detection Using CNN

Next, a CNN is applied to LUVT data to detect a defect. The CNN architecture used in this analysis can be seen in Table 1. The number of convolutional layers and hyperparameter values in the deep learning model were determined based on several parameter studies. However, since this problem is a binary classification problem that only determines the presence or absence of a defect, the structure is relatively simple, as shown in Table 1. As shown in this Table, the input image

size is 224×224 and the presence or absence of a defect is determined in the output of this CNN.

Figure 7 shows the results of applying the CNN to identify the presence or absence of a defect in a set of time history images acquired through LUVT. The CNN is trained to make these determinations for each time step captured in the images. The vertical axis shows the existence probability of a defect in each time step LUVT result. The LUVT images at representative time steps are also shown around the bars in Fig. 7. Note that this group of time-series images is an unknown group of images that have not been used as training data for creating an AI. The red circles in Fig. 7 indicate the presence or absence of a defect as determined by a human inspector. Note that, a value of 1 indicates by the red circles that a defect was presented and 0 indicates that no defect is detected.

As shown in Fig. 7, the incident wave reaches the defect in around 75 time-step and the scattered waves are generated. It can be seen that the AI's prediction results show approximately the same trend as the LUVT inspectors' predictions. Similar to LUVT inspector predictions, the probability of an AI-predicted defect's presence sharply increases from around the 75th time-step. After that, both the AI and the LUVT inspector's probability of the presence of a defect decrease. This may be because, after the scattered waves are generated, several images appear that the LUVT inspector eye cannot distinguish from the scattered waves caused by the defect due to phase interference of the scattered and the incident waves. The training time required for this CNN was approximately 13 minutes.

Table 2. CNN architecture with VGG16 used in Sect. 5.3.

Layer(type)	Output shape	Param.	
VGG16	(None,7,7,512)	14714688	Total params.: 14977857
MaxPool.2D	(None,512)	0	Trainable params.: 7342593
Dense	(None, 512)	262656	Non-trainable params: 7635264
Dropout	(None, 512)	0	
Dense-1	(None, 1)	513	

5.3 Automatic Defect Detection Using CNN with VGG16 Architecture

Finally, the CNN with VGG16 architecture is applied to LUVT data to create an AI. The CNN architecture with VGG16 used in this analysis can be seen in Fig. 8. As indicated in Fig. 8, pre-trained weights from VGG16 are utilized for defect detection in LUVT.

Figure 8 shows the results of applying the CNN with VGG16 to identify the presence or absence of a defect in a set of time history images acquired through

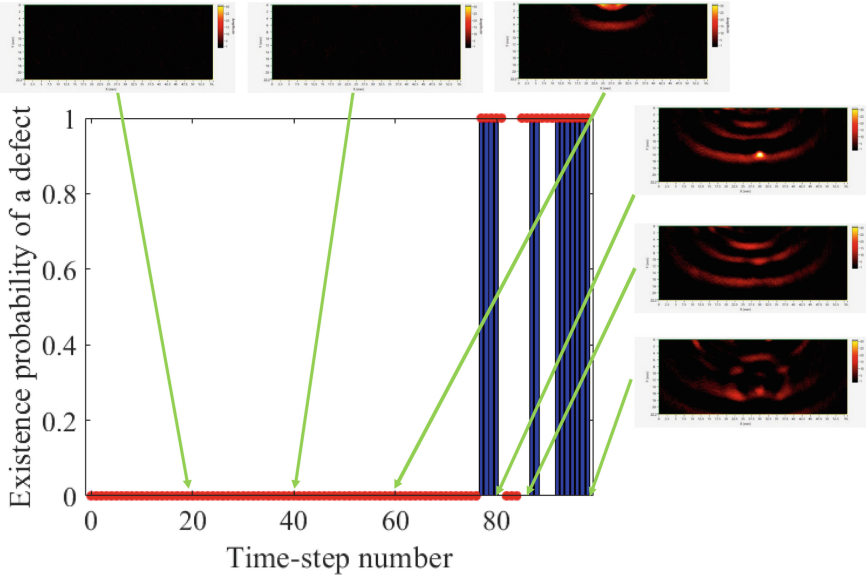


Fig. 8. Probability of a defect presence determined by the CNN architecture defined in Table 2.

LUVT. The test data used in this analysis are the same as those in Sect. 5.2. The trend of the results in Fig. 8 is almost the same as those in Fig. 7. However, a peak that is not observed in Fig. 7 is visible near the 90th time-step. This result shows that the transfer learning using VGG16 may not always yield good results for the automatic defect detection using LUVT. Therefore, it is necessary to continue to examine what kind of transfer learning is effective. For reference, the computation time for this transfer learning was approximately 19 min.

6 Conclusion

In this paper, we established automatic defect detection for LUVT experiments using CNN and transfer learning. Our results showed that the AI created by CNN judges the presence or absence of defects by focusing on the generation of scattered waves, similar to the LUVT inspector. In addition, it was shown that CNN and CNN using transfer learning could roughly determine the presence or absence of a defect. It is effective to transfer weights and network structures trained in other object recognition problems to the defect detection in LUVT, but it is necessary to carefully consider what kind of results to transfer. This study assumes the material is isotropic, but if anisotropic materials are targeted, additional training may be necessary because the scattering patterns, as shown in Figs. 3 and 5, caused by defects could change. In addition, if the defect shape is changed, such as in the case of a crack-type defect, additional training will be required to accurately detect.

In the future, we plan to detect a defect in LUVT images using GAN [24]. Furthermore, the transfer learning of image data obtained by FDTD as shown in Figs. 4 and 5 will be considered for the automatic defect detection in LUVT.

Acknowledgement. This work was supported by “Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures”, and “High Performance Computing Infrastructure” in Japan (Project ID: jh220033 and jh230036). In addition, JSPS KAKENHI(21K0423100) and SECOM Science and Technology Foundation supported this work.




References

1. Takatsubo, J., Miyauchi, M., Tsuda, H., Toyama, N., Urabe, K. and Wang, B.: Generation laser scanning method for visualizing ultrasonic waves propagating on a 3-D object, 1st International Symposium on Laser Ultrasonics: Science, Technology and Applications, (2008)
2. Rose, J. L.: Ultrasonic waves in solid media, Cambridge University Press, (2008)
3. Schmerr, L. W.: Fundamentals of ultrasonic nondestructive evaluation, Plenum Press, (1998)
4. Singh, R. and Vrana, J.: The World of NDE 4.0: Let’s think about it, Independently published, (2022)
5. Meng, M., Chua, Y. J., Wouterson, E. and Ong, C. P. K.: Ultrasonic signal classification and imaging system for composite materials via deep convolutional neural networks, *Neurocomputing*, vol. 257, pp. 128–135, (2017)
6. Saitoh, T., Kato, T. and Hirose, S.: Deep learning for scattered waves obtained by time-domain boundary element method and an attempt to classify defect types, *Journal of JSNDI*, vol.70(7), pp.272–279, (2021) (in Japanese).
7. Saitoh, T., Hirose, S., Fukui, T. and Ishida, T.: 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., EC Ltd. Eastleigh, pp. 355–360, (2007)
8. Ye, J, Ito, S. and Toyama, N.: Computerized ultrasonic imaging inspection: From shallow to deep learning, *Sensors*, vol. 18(11), p. 3820, (2018)
9. Tsuzuki, Y., Saitoh, Y., Nakahata, K., Minowa, R. and Saitoh, T.: Automatic detection of flaws from wavefield data using convolutional neural network, *Artificial Intelligence and Data Science*, vol. 1, pp. 339–348, (2020) (in Japanese).
10. Chollet, F.: Deep learning with Python, Manning Publications, (2017)
11. Pan, S. J. and Yang, Q. Y.: A survey on transfer learning, *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, pp. 1345–1359, (2009)
12. Simonyan, K. and Zisserman, A.: Very deep convolutional networks for large-scale image recognition, *ICLR 2015*, (2015)
13. Ang, W. T.: A beginner’s course in boundary element methods, Universal Publishers, Boca Raton, USA, (2007)
14. Dominguez, J.: Boundary element approach for dynamic poroelastic problems, *International Journal for Numerical Methods in Engineering*, vol. 35, pp. 307–324 (1992)
15. Beskos, D.: Boundary element methods in mechanics, North Holland, (1987)
16. Saitoh, T., Hirose, S. and Fukui, T.: Convolution quadrature boundary element method and acceleration by fast multipole method in 2-D viscoelastic wave propagation, *Theoretical and Applied Mechanics Japan*, vol. 57, pp. 385–393 (2009)

17. Kunz, K. S. and Luebbers, R. J.: Finite difference time domain method for electromagnetics, CRC Press, (1993)
18. Hughes, T. J. R.: The finite element method: linear static and dynamic finite element analysis, Dover Publications, (2012)
19. Achenbach, J. D.: Wave propagation in elastic solids, North Holland, (1987)
20. Auld, B. A.: Acoustic fields and waves in solids, vol.1, 2, R.E. Krieger, (1990)
21. Saitoh, T., Mori, A., Ooashi, K. and Nakahata, K.: Development of a new dynamic elastic constant estimation method for FRP and its validation using the FDTD method, *Insight- Non-Destructive Testing and Condition Monitoring*, vol. 61(3), pp. 162–165, (2019)
22. Tan, M. and Le, Q.: EfficientNet: Rethinking model scaling for convolutional neural networks, in: K. Chaudhuri, R. Salakhutdinov (Eds.), *Proceedings of the 36th International Conference on Machine Learning*, Vol. 97 of *Proceedings of Machine Learning Research*, PMLR, pp. 6105–6114, (2019)
23. Selvaraju, R. R., Cogswell, M., Das, R., Vedantam, R., Parikh, D. and Batra, D.: Grad-CAM: Visual explanations from deep networks via gradient-based localization, *Computer Vision and Pattern Recognition, Proceedings of IEEE International Conference on Computer Vision*, pp.618–626, (2017)
24. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y.: Generative adversarial nets, in: *Advances in neural information processing systems*, pp. 2672–2680, (2014)



Development of Nanomodified Graphene Concrete Using Machine Learning Methods

Thusitha Ginigaddara^(✉) , Thushara Jayasinghe , and Priyan Mendis 

University of Melbourne, VIC 3010, Australia
thusitha.ginigaddara@unimelb.edu.au

Abstract. Enhancing the performance of cementitious composites using nanomaterials have gained a significant research interest over the last few decades. Many studies have explored the use of various nanomaterial additives such as Carbon Nanotubes (CNTs), Carbon Nanofibers (CNFs), nano-silica and Graphene Oxide (GO) for cementitious composites. GO has shown more promising results over other nanomaterial additives as GO is hydrophilic and improves many properties of cementitious composites such as the mechanical strength, durability, resistance to fire, etc. While many studies have reported the compressive strength improvement of GO induced cementitious composites, there is a notable variance between the increments. Based on parameters such as GO dosage and water/cement ratio, the strength increments vary from as little as 5% to more than 125%. This paper primarily investigates the relationship between water/binder ratio and GO dosage in cementitious composites. The study includes a series of experiments on mechanical performance of GO induced cementitious composites, an intensive data collection from previous studies, synthesis, and characterization of GO. Moreover, for the first time, this paper presents a machine learning (ML) regression model to identify the most important factors for strength improvements and to predict the compressive strength gain of GO induced cementitious composites. The ML model was developed using more than 200 datapoints including various inputs as training sets and testing tests. The ML model was tuned to achieve a reliable accuracy and the model was validated through experimental investigations on cementitious composites. Based on the outcomes of the ML model, the most important variables for strength enhancement were identified and then applied to develop a GO induced high performance concrete which showed significant strength improvements over reference samples.

Keywords: Graphene oxide · Cementitious composites · Nanotechnology · Artificial intelligence

1 Introduction

After water, concrete is the most consumed material in the world [1]. According to 2019 records, the annual global consumption of cement was over 4 billion tons and the global demand for cement and concrete continue to increase parallel to infrastructure development of developed and developing countries [2]. Cement is the primary binding material

in concrete which sets, hardens, and adheres to aggregates to hold them together and deliver the desired properties of concrete [3]. The production and use of cement is a major contributor to CO₂ emissions, which has significant implications for global climate change. Cement is responsible for approximately 8% of global CO₂ emissions as production of 1 ton of cement emits 1 ton of CO₂ into the atmosphere [4]. This is mainly due to calcination of limestone during cement manufacturing process, which results in the release of CO₂ as a byproduct. In addition, the energy-intensive nature of cement production, including the mining of raw materials, transportation, and processing, contributes to CO₂ emissions from the cement industry [5]. To address the issue of carbon emissions of concrete, various nanomaterial additives have been explored. Carbon Nanotubes (CNT), Carbon Nanofibers (CNF) and Nano silica are some nanomaterials which have gained research interests over the last few years [6–9]. However, due to limitations in large scale production, high cost and hydrophobicity, most carbon based nanomaterials have proven to be ineffective as concrete additives.

Graphene Oxide (GO) is an advanced nanomaterial, which has shown promising results as an additive for cementitious composites. Graphene is a single layer of hexagonally arranged carbon atoms and since the Nobel prize-winning discovery of Graphene in 2004, various branches of pure science, applied science and engineering have explored its applications [10]. Graphene is extremely strong and stiff (100 times stronger than steel), yet also lightweight and flexible. It is also an excellent conductor of electricity and heat (1000 times more conductive than copper), as well as being highly resistant to chemicals and UV radiation [11]. Although graphene is hydrophobic, GO is a hydrophilic nanomaterial as it consists of few graphene layers decorated with oxygen containing functional groups [12]. The hydrophilicity allows GO to be added to concrete as an additive through a simple and cost-effective mixing process. Studies have found that adding small amounts of GO to cementitious composites can improve its strength, toughness, and durability. Ginigaddara et al. shows that GO can achieve similar performance of concrete by using 15–20% less cement. Use of GO improves the compressive strength of concrete up to 39% and GO can be used to reduce the embodied carbon content of concrete by 30% [13].

However, GO is still in its preliminary stage of research and development. For an instance, the strength increments gained by GO is associated to many variables such as the water/cement ratio [14], mix design, use of superplasticizer [15] and GO dosage. In order to identify the effect of each parameter, case-by-case experiments needs to be carried out. Therefore, a large number of research investigations including laboratory experiments are required before GO induced cementitious composites can be evaluated through large scale tests. These experiments require a great deal of resources including researchers, equipment, time, cost, materials, and training. However, machine learning regression models can be used for advanced and effective data processing through which the experimental test procedures can be expedited.

During the last decade, many studies have applied machine learning strategies to solve engineering problems [16]. Among many machine learning techniques, Artificial neural network (ANN) has been extensively employed to predict the compressive strength of cementitious composites. More recently, ensemble learning methods have been used by researchers owing to the meta-analysis capabilities of it [17]. Through

ensemble learning multiple machine learning models can be combined to obtain better outcomes. Ensemble learning methods such as random forest, gradient boosting, adaptive boosting, extreme gradient boosting, light gradient boosting, and categorical boosting are some of the most useful techniques that have been used by the researchers in the recent past.

This paper proposes a machine learning based regression model to predict strength increments of GO induced cementitious composites. The database for the machine learning model was developed from a combination of experimental test results and previously published data on GO based cementitious composites. For the experimental tests, GO was synthesized and characterized using various techniques. Moreover, mechanical properties of GO induced cementitious composites were evaluated through compressive strength tests. The developed machine learning model was used to identify the most important variables for strength gain and based on the findings a GO induced high performance concrete was developed.

2 Experimental Section

This section presents the laboratory work carried out to synthesize GO and to identify the relationship between variables such as w/c ratio and GO dosage and strength gain of cementitious composites. Under this section, the synthesis of GO, materials used and preparation of cementitious test specimens are discussed.

2.1 GO Synthesis

The GO was synthesized using an improved version of Hummer's method, which involved oxidizing natural graphite. Natural graphite was obtained from the Kahatagaha Mine in Sri Lanka. The startup graphite size used in the synthesis process was 63–90 μm and shown in Fig. 1a. The first step of the synthesis process consists of the intercalation of graphite using H_2SO_4 . For this purpose, the concentrated (98%) acid was cooled to 10 $^\circ\text{C}$ and raw graphite powder was added and continuously stirred while maintaining the temperature below 20 $^\circ\text{C}$. This step is then followed by the oxidation of intercalated graphite using KMnO_4 as the oxidizing agent. KMnO_4 was added gradually into the graphite-acid mixture and the reaction temperature was maintained at 50 $^\circ\text{C}$ for 3 h. Later, a mixture of ice and chilled distilled water was added, and the suspension was stirred for a further hour at 60 $^\circ\text{C}$. The whole process was then followed by a dropwise addition of H_2O_2 to quench the reaction and terminate any further oxidation. The final suspension was centrifuged and the supernatant was decanted away. The residual was then washed with HCl and DI multiple times to remove residual salts and un-exfoliated graphite oxide. The synthesis process was carried out by Ceylon Graphene Technologies Pvt Ltd, Sri Lanka and Fig. 1b illustrates the synthesized GO in the form of an aqueous solution of 5g/l concentration [18].

2.2 Preparation of Cementitious Composite Test Specimens

Test specimens of varying w/c ratio and GO dosage were casted for the experimental work. Grade 42.5 Ordinary Portland Cement (OPC) as the binder, natural sand as

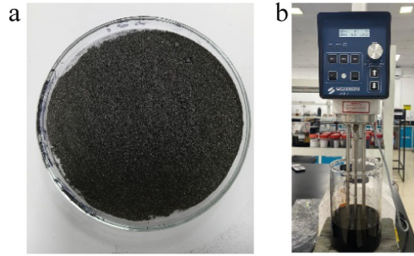


Fig. 1. Startup graphite particles (a) and synthesized GO dispersion (b)

fine aggregates and a polycarboxylate based superplasticizer (PCE) as admixture were used to cast the specimens. Tests were conducted as per ASTM C-109 - Standard Test Method for Compressive Strength of Hydraulic Cement Mortars and for each mix design, three specimens were casted for testing. Detailed mix proportions for all specimens are illustrated in Table 1.

Table 1. Mix proportions of cementitious composites

W/C ratio	Notation	Cement (g)	Fine aggregates (g)	Water (g)	GO (g)	PCE (ml)
0.3	G-0-0.30	450	1350	135	0	0.9
	G-01-0.30	450	1350	135	0.045	0.9
	G-03-0.30	450	1350	135	0.135	0.9
	G-05-0.30	450	1350	135	0.225	0.9
0.4	G-0-0.40	450	1350	180	0	0.9
	G-01-0.40	450	1350	180	0.045	0.9
	G-03-0.40	450	1350	180	0.135	0.9
	G-05-0.40	450	1350	180	0.225	0.9
0.5	G-0-0.50	450	1350	225	0	0.9
	G-01-0.50	450	1350	225	0.045	0.9
	G-03-0.50	450	1350	225	0.135	0.9
	G-05-0.50	450	1350	225	0.225	0.9
0.6	G-0-0.60	450	1350	270	0	0.9
	G-01-0.60	450	1350	270	0.045	0.9
	G-03-0.60	450	1350	270	0.135	0.9
	G-05-0.60	450	1350	270	0.225	0.9

The following mixing sequence was ensued to effectively disperse GO in the cement matrix without agglomerations. The PCE content was manually mixed in the total water content of each mixture for one minute. Then, GO was added into the PCE and Water

content and magnetically stirred for a period of 10 min. Concurrently, sand and cement contents were mixed using an automatic mortar mixer at a speed of 62.5 rpm for a period of three minutes. Then, the liquid content (PCE, Water and GO) was added to the dry content (Sand and Cement) and mixed for a period of 6 min. At the end, confirming a good mixture of all materials, the mortar was placed in steel molds and vibrated for 30 s for compaction. After keeping 24 h in the room temperature, the specimens were demolded and cured in a water bath for 28 days.

3 Results and Discussion

This section presents the results of an extensive characterization of synthesized GO which confirms the oxidation of graphite. Moreover, the test results of compressive strength tests are also discussed in this section.

3.1 X-ray Diffraction Analysis (XRD)

The XRD analysis was conducted for the synthesized GO samples using a BRUKER D8 Focus X-ray diffractometer to confirm conversion from graphite to GO based on interlayer distances (d spacing). The analysis parameters were set as, Cu $K\alpha$ radiation ($\lambda = 0.154$ nm) over a 2θ range of $5\text{--}60^\circ$ with a step size of 0.02° and a step time of 10 s. Bragg's Law was used to quantify the d spacing based on the 2θ values of the analysis [19].

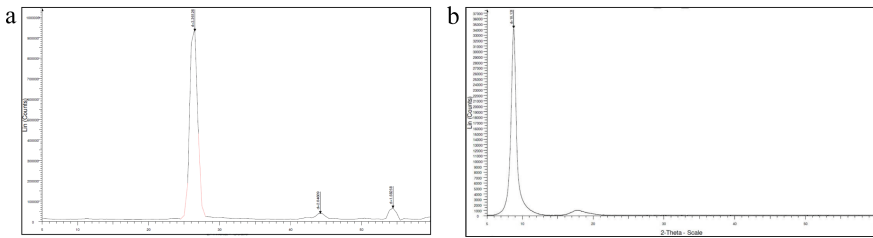


Fig. 2. XRD analysis of graphite (a) and synthesized GO (b)

Results of the graphite XRD analysis are depicted in Fig. 2a and the first sharp peak is observed at 2θ value of 36.50° which corresponds to a d spacing of 3.36\AA (0.036nm). Contrastingly, the GO analysis shown in Fig. 2b depicts a sharp peak at 2θ value of 8.85° which is associated with a d spacing of 10.138\AA (1.0138nm). The higher interlayer distance of GO is due to the availability of functional groups and water molecules in between the layers of GO. However, further characterizations were carried out for elemental analysis and quantify the oxidation level of GO.

3.2 X-ray Photoelectron Spectroscopy (XPS)

XPS was used to quantify the Carbon to Oxygen ratio (C/O) of synthesized GO. A Thermo Scientific™ ESCALAB™ Xi + X-ray Photoelectron Spectrometer was used

for analysis, The synthesized GO sample was mounted on a glass substrate and the parameters of the X-Ray source was set to Monochromatic Al K α (1486.6 eV) with a spot size of 900 μm . Survey scans and high-resolution scans were collected with pass energies of 150 and 20 eV with a step size of 1.0 and 0.05 eV. Detailed spectra processing was performed by Thermo Avantage (5.982) software.

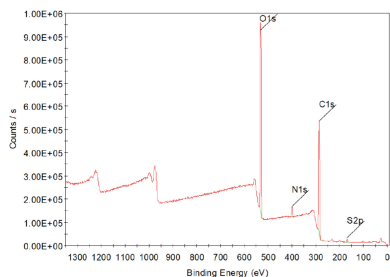


Fig. 3. XPS spectra of synthesized GO

Figure 3 and Table 2 illustrate the XPS spectra and quantification of elements respectively.

Table 2. Elemental compositions

Element	Peak binding energy (eV)	Average %
Carbon (C1s)	284.24	64.67
Oxygen (O1s)	530.80	31.29
Nitrogen (N1s)	397.92	2.77
Sulfur (S2p)	167.10	1.16

Based on the XPS analysis, the C/O ratio of synthesized GO was found to be 2.07. The availability of 31.29% Oxygen confirms that the graphite particles are substantially oxidized. Further characterization was carried out to identify the types of oxygen containing functional groups in GO.

3.3 Fourier-Transform Infrared Spectroscopy (FTIR)

FTIR was used to assess the types of functional groups available in synthesized GO sheets. A BRUKER Vertex80 FTIR microscope (Hyperion) was used for the analysis and the attenuated total reflection fourier transform infrared (ATR-FTIR) spectra of the graphene oxide were recorded in the region 800 to 4000 cm^{-1} at a resolution of 4 cm^{-1} .

As shown in the results in Fig. 4, the peaks of 3199 cm^{-1} , 1725 and 1378 cm^{-1} represents $-\text{OH}$, $-\text{COOH}$, $-\text{O}-$ functional groups, respectively. Similarly, $\text{C}-\text{O}-\text{C}$ and $\text{C}=\text{C}$ bonds are associated with the peaks observed at the peaks obtained at 1036 cm^{-1} and 1623 cm^{-1} respectively [20].

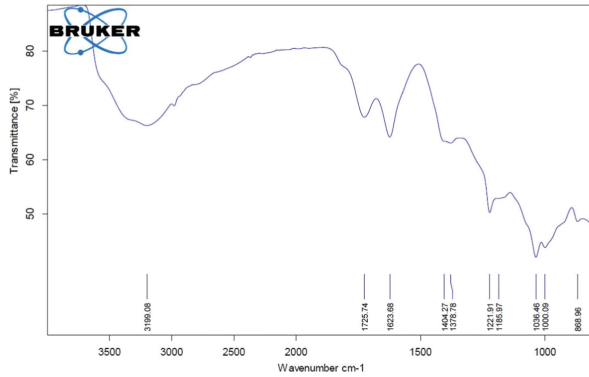


Fig. 4. FTIR spectra of synthesized GO

3.4 Raman Spectroscopy

The Raman Spectroscopy was carried out using a BRUKER SENTERRA II -Confocal Raman Microscope. The synthesized GO samples were drop casted on a substrate (SiO₂ wafers of 300 nm thickness) and three different spots were analyzed. Parameters were set to 532 nm green laser with 50X optical zooming. The results of the analysis are shown in Fig. 4 (Fig. 5).

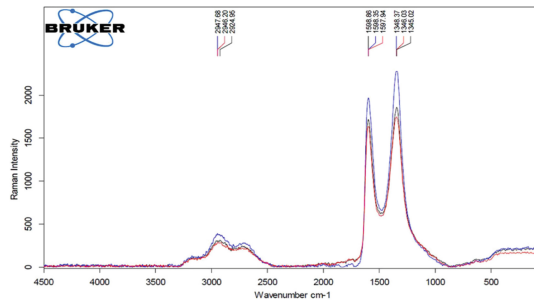


Fig. 5. Raman spectra of synthesized GO

The D band and G band of Raman spectroscopy are indications of defects and graphitic nature of the GO samples, respectively [21]. The D band confirms the availability of oxygen containing functional groups in GO sheets while the G band characterizes the C-C vibrational mode [22]. In the analysis of synthesized GO, the D band peak was observed at 1341 cm^{-1} and the G band peak was observed at 1587 cm^{-1} . Based on the intensities of G and D peaks, it was calculated that the I_D/I_G is 1.095 and sufficient defects are available in the GO sheets. Additionally, as observed between 2500 cm^{-1} and 3300 cm^{-1} a 2D band further confirms the availability of oxygen containing functional groups.

3.5 Experiment Test Results of Cementitious Composites

The mechanical properties of cementitious composites were evaluated through compressive strength tests of hardened test specimens. The compressive strength tests were conducted at 28 days curing period and the average compressive strength of three specimens for each mix design were determined. Results of the 28-day compressive strength tests are summarized in Table 3 and visualized in Fig. 6 with 95% confidence interval.

Table 3. Mechanical properties of cementitious composites

W/C	Notation	Average compressive strength	Strength gain
0.3	G-0-0.30	82.30	–
	G-01-0.30	93.64	14%
	G-03-0.30	87.5	6%
	G-05-0.30	86.12	5%
0.4	G-0-0.40	58.57	–
	G-01-0.40	69.14	18%
	G-03-0.40	66.22	13%
	G-05-0.40	61.05	4%
0.5	G-0-0.50	43.58	–
	G-01-0.50	49.25	13%
	G-03-0.50	54.26	25%
	G-05-0.50	48.63	12%
0.6	G-0-0.60	29.10	–
	G-01-0.60	31.23	7%
	G-03-0.60	33.64	16%
	G-05-0.60	29.65	2%

As presented in Table 3, for all GO induced samples, regardless of the w/c ratio, GO had a positive influence on the mechanical strength. As predicted, the strength of control samples (G-0-0.30, G-0-0.40, G-0-0.50, G-0-0.60) have reduced with the increasing w/c ratio. This is associated with the generation of capillary pores in cement matrix from excess amount of water. Moreover, as visualized in Fig. 6, results show that the w/c ratio has an impact on the optimum dosage of GO (i.e. highest strength gain). For an instance, for the samples of 0.3 and 0.4 w/c ratios, the optimum GO dosage is 0.01% by weight of cement (bwoc). Contrastingly, the optimum GO dosage for samples of 0.5 and 0.6 w/c ratios is 0.03%. This clearly indicates a relationship between free water in the cement matrix and GO. It is noted that, with more free water in the cement matrix, more GO is needed for strength improvements.

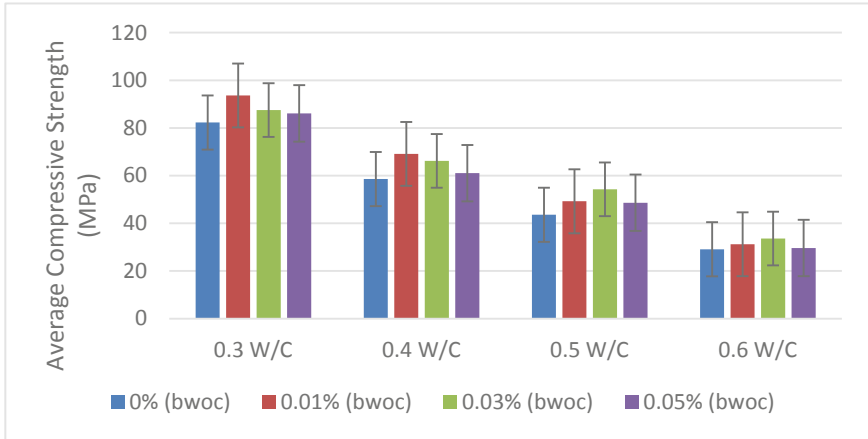


Fig. 6. Variation of average compressive strength with W/C ratio

However, for certain w/c ratios, the strength gain is more prominent when compared with other samples. For an instance, the samples of 0.5 w/c ratio reports the highest strength gain whereas the samples of 0.6 w/c ratio shows the minimum strength gain. This variation of strength gain is depicted in Fig. 7.

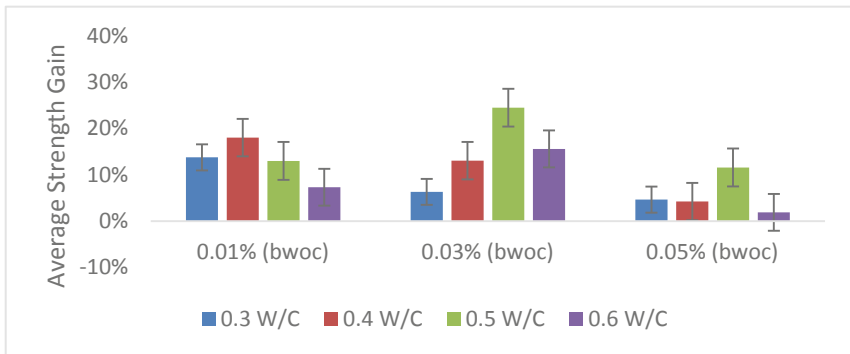


Fig. 7. Variation of average strength gain with GO dosage

Although 0.01 and 0.03% GO dosages have delivered highest strength gains in different w/c ratios, with 0.05% GO dosage a reducing pattern is observed for the strength gains. This is due to the agglomeration of GO which occurs as a result of high number of nanoparticles in the cement matrix.

4 Machine Learning Method and Graphene Concrete

4.1 Random Forest Algorithm

The machine learning framework includes an ensemble learning method which is a combination of decisions of two or more individual learning tools such as decision trees [23]. The “random forest” technique, which is a parallel ensemble method, was first adopted to process the developed dataset [24]. The random forest algorithm is an ensemble of multiple random decision trees which is less sensitive to the training data. The algorithm first performs bootstrapping and then aggregating, which is commonly known as bagging for which the framework is depicted in Fig. 8 [25].

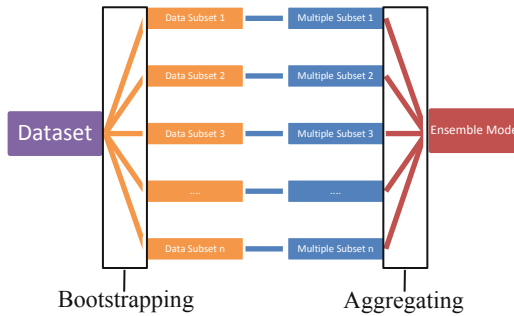


Fig. 8. Machine learning framework for random forest algorithm

The accuracy of the developed random forest machine learning model was identified through several statistical parameters. These parameters are Root Mean Square Error (RMSE), Coefficient of Determination (R^2), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) for which the applicable comparisons are shown from Eqs. 1–4 respectively.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (SG_{Predict} - SG_{Exp})^2}{n}} \quad (1)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (SG_{Exp} - SG_{Predict})^2}{\sum_{i=1}^n (SG_{Exp} - \overline{SG}_{Predict})^2} \quad (2)$$

$$MAE = \frac{\sum_{i=1}^n |SG_{Predict} - SG_{Exp}|}{n} \quad (3)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{SG_{Exp} - SG_{Predict}}{SG_{Exp}} \right| \quad (4)$$

The statistical parameters of the developed model are shown in Table 4. It is noted that the R^2 value for both training set and testing set are close to 1, which indicates a good prediction model [26]. Similarly, the lower values in RMSE, MAE and MAPE indicates good performance of the developed model. Moreover, the relationship between predicted strength gain and tested strength gain is shown in Fig. 9a.

Table 4. Statistical parameters of the developed model

Parameter	Value	
	training set	Testing set
R^2	0.921	0.817
RMSE	4.579	8.658
MAE	18.664	45.417
MAPE	0.288	0.361

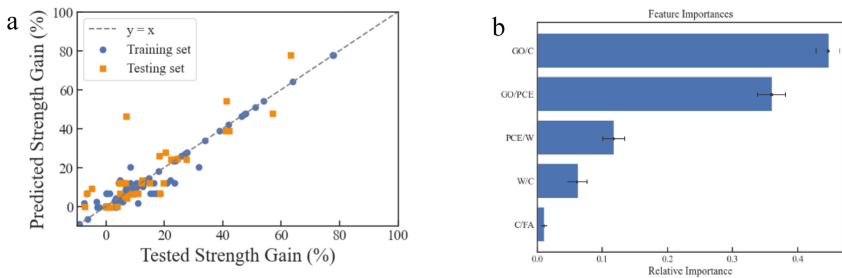


Fig. 9. Relationship between predicted strength gain and tested strength gain (a) and feature importance of parameters (b)

Based on the developed machine learning model, the most significant features which have influenced the strength gain of GO induced cementitious composites were identified. As shown in Fig. 9b, after the GO dosage (denoted as GO/C ratio), GO/PCE ratio has the highest impact on the strength gain. Similar conclusions have been drawn in experimental investigations on GO induced cementitious composites as the steric hinderance of PCE promotes the dispersion of GO in cementitious matrix without agglomerations. The PCE/Water ratio (denoted as PCE/W) also shows moderate significance and relative importance as the total PCE content affects the workability and packing density of cementitious composites.

4.2 Nanomodified Graphene Concrete

From the developed model, it was found that 0.03% GO dosage (bwoc) has delivered the highest strength increment for mix designs with w/c ratio of 0.45. Moreover, the optimum GO/PCE ratio was found to be 0.15. Based on these parameters, a mix design

was developed for a nanomodified graphene concrete. Concurrently, a reference mix design was formulated to evaluate the performance improvement of graphene concrete. The mix designs of graphene concrete and reference concrete are shown in Table 5.

Table 5. Proportion of concrete mix design (Kg/m³)

Mix ID	Water	Cement	Coarse aggregate	Fine aggregate	GO (% bwoc)	PCE
Reference samples	205	456	1175	579	0 (0%)	0.9
Graphene concrete	205	456	1175	579	0.1368 (0.03%)	0.9

Six concrete samples were prepared for each mix design as per ASTM C172, and the compressive strengths of specimens were tested at 7 days and 28 days periods as per ASTM C39 (Testing 3 specimens from each mix design at 7 days and 28 days). Figure 10a shows the hardened concrete specimens and Fig. 10b shows a concrete specimen subjected to compressive strength test. Figure 11 shows the variation of compressive strength at different curing periods (7 days and 28 days).

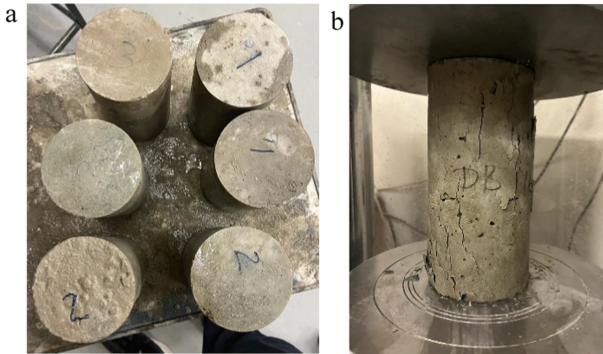


Fig. 10. Prepared concrete specimens (a) and a tested concrete sample (b) captured at 28 days curing period

Graphene concrete shows significantly better mechanical performance in terms of average compressive strength when compared with reference samples both at early age (7 days curing period) and later (28 days curing period). Comparing the average strength of graphene concrete and reference samples, it was found that at 7 days and 28 days, graphene concrete shows significant strength gain of 40% and 30% respectively.

The strength gain during the early age in graphene concrete extremely advantageous as early age strength plays a crucial role in construction processes. This enables applications in early formwork removing, shorter construction duration, more effective post tensioning, reducing the initiation of micro cracks, reducing early age shrinkage

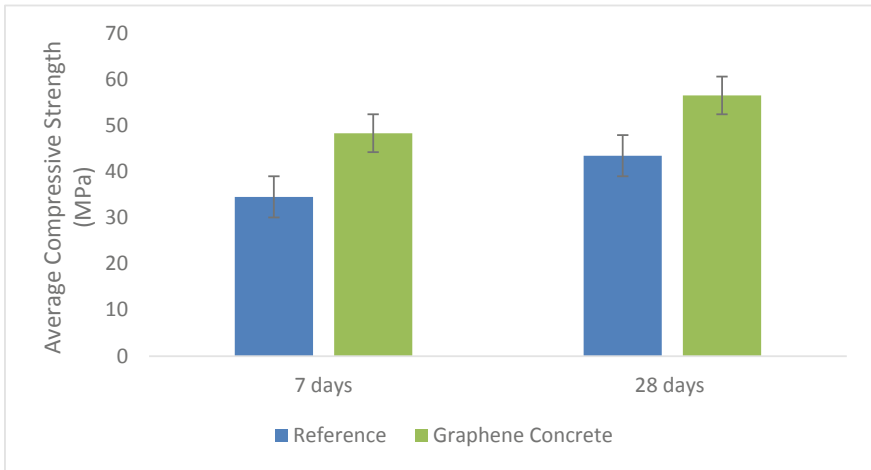


Fig. 11. Average compressive strength at different curing periods

and thermal contraction. These applications are advantageous to enhance the ongoing practices in the construction industry decreasing both time and cost.

Similarly, the significant 28 day strength gain of 30% implies that the overall cement quantity of the concrete can be reduced. Reduction of cement content will not only deliver financial benefits, but more importantly, reduce the embodied carbon of concrete. Authors, in their previous work have highlighted the opportunities in reducing the embodied carbon of concrete through graphene [4].

5 Conclusion

This study comprehensively covered extensive laboratory work as well as machine learning regression models for optimized data processing. As parts of the laboratory work, GO was synthesized using natural graphite and a full characterization using techniques such as XRD, XPS, FTIR and Raman spectroscopy was carried out. The characterization results confirmed the oxidation of graphite and conversion of graphite to GO. Through XRD, it was found that during the synthesis, interlayer distance was increased from 0.036 nm (graphite) to 1.0138 nm, successfully converting graphite to GO. XPS spectroscopy revealed availability of 31.29% oxygen in synthesized GO and a C/O ratio of 2.07. Availability of various oxygen containing functional groups such as -OH and -COOH was confirmed through FTIR spectra. The lower graphitic nature and higher order of defects in the synthesized GO sheets were investigated through Raman spectroscopy which confirmed an I_D/I_G ratio of 1.095.

Subsequently, a series of tests were carried out inducing the synthesized GO in cementitious composites to evaluate the mechanical performance of test specimens through compressive strength tests. The results showed that all samples containing GO having higher compressive strength when compared to reference samples. However, it was discovered that the strength gain attributed to GO varies with parameters such as the GO dosage and the W/C ratio.

Moreover, in this study, for the first time, machine learning methods were used to develop a model for GO induced cementitious composites. The laboratory work was combined with data from previous studies to formulate a database for a machine learning model containing more than 200 datapoints. The machine learning model was developed using random forest algorithm and the statistical parameters such as R^2 , RMSE, MAE and MPAE confirmed the development of a reliable machine learning model. From the outcomes of the model, it was identified that parameters such as GO dosage (bwoc), GO/PCE ratio and W/C ratio are significant variables which affect the strength gain attributed to GO. Finally, based on the parameters of the machine learning model, a graphene concrete was developed, which showed a significant strength improvement of 40% at 7 days following a 30% strength improvement at 28 days, when compared with reference samples.

Further investigations on graphene concrete are underway and machine learning tools are being used to optimize the mix design parameters of graphene concrete. Moreover, the characterization techniques are being used to understand the microstructural behavior between GO and cementitious compounds and findings of those work will be published in the future.

References

1. Gagg, C.R.J.E.F.A.: Cement and concrete as an engineering material: An historic appraisal and case study analysis. 40, 114–140 (2014)
2. Petroche, D.M., Ramirez, A.D.J.B.: The environmental profile of clinker, cement, and concrete: a life cycle perspective study based on ecuadorian data. 12, 311 (2022)
3. Akroyd, T.N.W.: Concrete: properties and manufacture. Elsevier (2016)
4. Ginigaddara, T., Mendis, P., Devapura, P., Mohotti, D., Fonseka, I.: Innovative low carbon graphene concrete. Concrete in Australia 48, 33–41 (2022)
5. Van Oss, H.G., Padovani, A.C.J.J.o.I.E.: Cement manufacture and the environment: part I: chemistry and technology. 6, 89–105 (2002)
6. Feng, Y., Su, Y.-F., Lu, N., Shah, S.J.E.S.: Meta concrete: Exploring novel functionality of concrete using nanotechnology. 8, 1–10 (2019)
7. Zhu, X., Gao, Y., Dai, Z., Corr, D.J., Shah, S.P.J.C., Research, C.: Effect of interfacial transition zone on the Young's modulus of carbon nanofiber reinforced cement concrete. 107, 49–63 (2018)
8. Kwalramani, M.A., Syed, Z.J.I.J.o.I.E.: Application of nanomaterials to enhance microstructure and mechanical properties of concrete. 10, (2018)
9. Vera-Agullo, J., Chozas-Ligero, V., Portillo-Rico, D., García-Casas, M., Gutiérrez-Martínez, A., Mieres-Royo, J., Grávalos-Moreno, J.: Mortar and concrete reinforced with nanomaterials. In: Nanotechnology in Construction 3: Proceedings of the NICOM3, pp. 383–388. Springer, (Year)
10. Novoselov, K.S., Geim, A.K., Morozov, S.V., Jiang, D.-e., Zhang, Y., Dubonos, S.V., Grigorieva, I.V., Firsov, A.A.J.s.: Electric field effect in atomically thin carbon films. 306, 666–669 (2004)
11. Geim, A.K.J.s.: Graphene: status and prospects. 324, 1530–1534 (2009)
12. Dimiev, A.M., Eigler, S.: Graphene oxide: fundamentals and applications. John Wiley & Sons (2016)
13. Ginigaddara, T., Ekanayake, J., Mendis, P., Devapura, P., Liyanage, A., Vaz-Serra, P.J.E.J.o.S.E.: An Introduction to High Performance Graphene Concrete. 22, 11–18 (2022)

14. Devasena, M., Karthikeyan, J.J.I.E.S.I.R.D.: Investigation on strength properties of graphene oxide concrete. 1, 307–310 (2015)
15. Lu, Z., Hanif, A., Ning, C., Shao, H., Yin, R., Li, Z.J.M., Design: Steric stabilization of graphene oxide in alkaline cementitious solutions: Mechanical enhancement of cement composite. 127, 154–161 (2017)
16. Jayasinghe, T., Gunawardena, T., Mendis, P.J.C.S.i.C.M.: Assessment of shear strength of reinforced concrete beams without shear reinforcement: A comparative study between codes of practice and artificial neural network. 16, e01102 (2022)
17. Hafez, H., Teirelbar, A., Kurda, R., Tošić, N., de la Fuente, A.J.C., Materials, B.: Pre-bcc: A novel integrated machine learning framework for predicting mechanical and durability properties of blended cement concrete. 352, 129019 (2022)
18. <https://ceylongraphene.com/graphene-in-concrete/>
19. Elton, L., Jackson, D.F.J.A.J.o.P.: X-ray diffraction and the Bragg law. 34, 1036–1038 (1966)
20. Surekha, G., Krishnaiah, K.V., Ravi, N., Suvarna, R.P.: FTIR, Raman and XRD analysis of graphene oxide films prepared by modified Hummers method. In: Journal of Physics: Conference Series, p. 012012. IOP Publishing, (Year)
21. Lee, A.Y., Yang, K., Anh, N.D., Park, C., Lee, S.M., Lee, T.G., Jeong, M.S.J.A.s.s.: Raman study of D* band in graphene oxide and its correlation with reduction. 536, 147990 (2021)
22. Xu, L., Cheng, L.J.J.o.N.: Graphite oxide under high pressure: a Raman spectroscopic study. 2013, 47–47 (2013)
23. Dietterich, T.G.J.T.h.o.b.t., networks, n.: Ensemble learning. 2, 110–125 (2002)
24. Chen, J., Li, K., Tang, Z., Bilal, K., Yu, S., Weng, C., Li, K.J.I.T.o.P., Systems, D.: A parallel random forest algorithm for big data in a spark cloud computing environment. 28, 919–933 (2016)
25. Breiman, L.J.M.I.: Bagging predictors. 24, 123–140 (1996)
26. Chicco, D., Warrens, M.J., Jurman, G.J.P.C.S.: The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. 7, e623 (2021)

Optimization Techniques



A Multi-objective Optimization Design Framework for Thin-Walled Tubular Deployable Composite Boom for Space Applications

Tian-Wei Liu^{1,2,3}, Jiang-Bo Bai^{1,2(✉)}, and Nicholas Fantuzzi³

¹ School of Transportation Science and Engineering, Beihang University, Beijing 100191, People's Republic of China

baijiangbo@buaa.edu.cn

² Jingdezhen Research Institute of Beihang University, Jiangxi Province 333000, People's Republic of China

³ DICAM Department, University of Bologna, 40136 Bologna, Italy

Abstract. The thin-walled tubular deployable composite boom (DCB), which can achieve folding and deploying functions by storing and releasing strain energy, has great potential applications in the space field, such as being the main supporting part of deployable structures like solar sails. This paper proposes a multi-objective optimization design framework for the DCB. First, the energy method was used to determine the folding moment required during the folding process of the DCB. Based on the analytical model of the folding moment, an analytical model was proposed to predict the failure index of the DCB in the folding state, using the classical laminate theory and three failure criteria. Then, a multi-objective optimization design framework for optimizing the DCB was proposed based on the popular NSGA-III algorithm. The optimization results show that a total of 458 design schemes are found on the Pareto front, all of which are superior to the data in the literature, demonstrating the effectiveness of the multi-objective optimization design framework. The research results are of great significance for the practical engineering application of the DCB.

Keywords: Deployable composite boom · Folding behaviour · Analytical model · Multi-objective optimization · NSGA-III

1 Introduction

Deployable composite structures are becoming increasingly popular in the space structures due to their ability to fold and self-deploy, which overcomes the strict volume constraints of spacecraft [1–4]. These structures can be used to create a variety of deployable space structures, such as helical antennas [5, 6], hinges [7, 8], cabin segments [9, 10], and booms [11, 12]. The thin-walled tubular deployable composite boom (DCB) is a classic example of these structures, as it can be folded and packaged into a small space, then restored to its original configuration by releasing the stored strain energy from the folding process. The tubular DCB serves as a fundamental component for constructing

deployable space structures, featuring advantages like lightweight, high stiffness, and repeatable deployment. Additionally, the tubular DCB can also be utilized as a structural backbone component in deployable parabolic mesh antennas, deployable solar sails, and space probes.

The mechanical properties of DCBs have become a research hotspot in recent years, making it necessary to optimize the design of these structures. Liu et al. [13] presented analytical models and an optimization design framework for the lenticular DCB, which are effective and show good prediction accuracy. NSGA-II is expected to provide better optimization results for the lenticular DCB. Yang et al. [14] developed a surrogate modeling-based optimal design method for a new N-shape cross-section composite ultra-thin deployable boom, integrating design of experiments, computational analyses, response surface method, and multi-objective optimization, which can be applied to different problems in structural and material design. Yang et al. [15] proposed an optimization method for deployable composite triangle rollable and collapsible booms, which can be flattened and self-deployed. The optimal design configuration features smaller concentrated stress, higher fully deploying status fundamental frequency, and appropriate wrapping peak moment. In the follow-up work, Yang et al. [16] presented an optimal design method for the coiling and deploying of a C-cross section thin-walled composite deployable boom based on response surface method and multi-objective optimization, and validated the accuracy of the optimization and response surface results using finite element models. Zhang et al. [17] presented a data-driven computational framework combining machine learning and multi-objective optimization for the design of space deployable bistable composite booms with C-cross section, and included a parametric study of geometric parameters.

As evident from the previous studies, existing optimization methods for design suffer from issues such as complexity, time consumption, and low accuracy due to their reliance on coupled surrogate models and genetic algorithms. This study proposed a novel optimization method that combines genetic algorithms with analytical models to achieve higher accuracy and computational efficiency. Compared to existing methods that use coupled surrogate models and genetic algorithms, this approach is both practical and able to accurately approximate the true optimization results. Additionally, it avoids issues such as surrogate model error and parameter tuning, providing a feasible and valuable approach for the optimization design of composite structures. To optimize the DCB, the folding moment required during the folding process was first determined using the energy method. Based on the resulting analytical model, a new model was developed to predict the failure index of the DCB in the folding state, incorporating the classical laminate theory and three failure criteria. Finally, a multi-objective optimization design framework for the DCB was proposed, utilizing the widely-used NSGA-III algorithm.

The organization of the paper is as follows: Sect. 2 establishes an analytical model for predicting the folding behavior of the DCB. Section 3 proposes a multi-objective optimization design framework. The optimization results are presented in Sect. 4. Finally, Sect. 5 provides a summary of the full paper.

2 Analytical Models

2.1 Folding Moment

The DCB can be folded using a roller, as illustrated in Fig. 1a, to occupy minimal storage space. Upon deployment, the DCB can recover to the initial configuration by utilizing the elastic strain energy stored in its own coiling deformation. The geometric configuration of the DCB in the initial state is defined by the length L , thickness t , cross-sectional radius R , and cross-sectional angle θ , as shown in Fig. 1b and 1c. To characterize the geometric model of the DCB during the folding deformation process, the thickness variation is neglected, and the overall deformation is described by the shape of the neutral surface and the variation of the curvature radius, without any stretching of the neutral surface.

The moment of inertia about the y -axis of the cross-section of the DCB can be expressed as [10]

$$I = \frac{\theta + \sin \theta}{2} \left(tR^3 + \frac{t^3 R}{4} \right) + t\theta R \left[\frac{2}{\theta} \sin \frac{\theta}{2} \left(R + \frac{t^2}{12R} \right) \right]^2 - t \sin \frac{\theta}{2} \left(4R^2 + \frac{t^2}{3} \right) \left[\frac{2}{\theta} \sin \frac{\theta}{2} \left(R + \frac{t^2}{12R} \right) \right] \quad (1)$$

The thickness of the deployable composite boom can be expressed as

$$t = nt_0 \quad (2)$$

where n is the number of plies and t_0 is the thickness of a single ply.

The energy method is employed to determine the folding moment required for the DCB during the folding process. To obtain the strain energy of the DCB, the folding process is divided into two stages, as illustrated in Fig. 1d. In the first stage, a bending moment M_y is applied at the edge to change the transverse curvature of the DCB from $1/R$ to 0, and in the second stage, a bending moment M_x is applied at the edge to change the longitudinal curvature of the DCB from 0 to $1/r$. r is the radius of the roller. The external work is converted into bending strain energy in both stages. This approach enables the determination of the folding moment required for the DCB during the folding process and provides a means to evaluate the strain energy generated during the process.

By using the ABD matrix used in classical laminate theory, the elastic deformation of the DCB during the folding process can be written as

$$\begin{bmatrix} \mathbf{N} \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon}_0 \\ \Delta \boldsymbol{\kappa} \end{bmatrix} \quad (3)$$

The DCB is typically prepared as a symmetric laminate (i.e., $\mathbf{B} = 0$). Moreover, due to the constraints imposed by the folding mechanism, the torsional deformation κ_{xy} can be neglected. Therefore, Eq. (3) can be simplified as

$$\begin{bmatrix} M_x \\ M_y \\ 0 \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} & D_{16} \\ D_{12} & D_{22} & D_{26} \\ D_{16} & D_{26} & D_{66} \end{bmatrix} \begin{bmatrix} \Delta \kappa_x \\ \Delta \kappa_y \\ 0 \end{bmatrix} \quad (4)$$

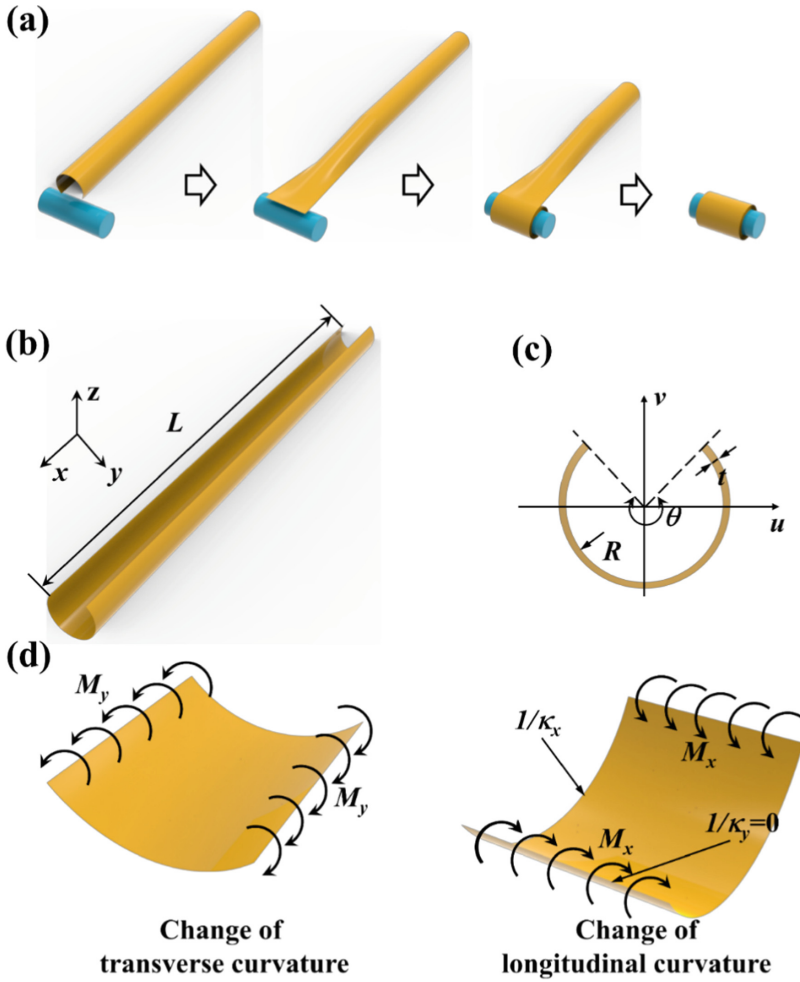


Fig. 1. Schematic diagram of the DCB: (a) Folding deformation process (b) Deploying state (c) Cross-section of the deploying state (d) Force analysis during the folding process.

The strain energy per unit area generated when the DCB changes from the deploying state to the folding state can be expressed as [18]

$$\begin{aligned}
 u &= \frac{R\theta}{2} [\Delta\kappa_x \ \Delta\kappa_y \ 0] \begin{bmatrix} D_{11} & D_{12} & D_{16} \\ D_{12} & D_{22} & D_{26} \\ D_{16} & D_{26} & D_{66} \end{bmatrix} \begin{bmatrix} \Delta\kappa_x \\ \Delta\kappa_y \\ 0 \end{bmatrix} \\
 &= \frac{R\theta}{2} (D_{11}\Delta\kappa_x^2 + 2D_{12}\Delta\kappa_x\Delta\kappa_y + D_{22}\Delta\kappa_y^2)
 \end{aligned}
 \tag{5}$$

where R and θ represent the radius and central angle of the cross-section of the DCB, respectively. The values of $\Delta\kappa_x$ and $\Delta\kappa_y$ can be derived as

$$\begin{cases} \Delta\kappa_x = \frac{1}{r} \\ \Delta\kappa_y = -\frac{1}{R} \end{cases} \quad (6)$$

Therefore, the strain energy generated during the folding process of the DCB can be expressed as

$$U(\alpha) = \frac{R\theta r\alpha}{2} \left(\frac{D_{11}}{r^2} - \frac{2D_{12}}{rR} + \frac{D_{22}}{R^2} \right) \quad (7)$$

where α represents the coiling angle of the DCB in any state.

Therefore, the folding moment required for the DCB during the folding process can be deduced as

$$T = \frac{dU(\alpha)}{d\alpha} = \frac{R\theta r}{2} \left(\frac{D_{11}}{r^2} - \frac{2D_{12}}{rR} + \frac{D_{22}}{R^2} \right) \quad (8)$$

2.2 Failure Index

Figure 2 is a schematic diagram of a laminate. The stress-strain relationship for the k th ply in the laminate can be expressed as

$$\begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix}_k = \begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} & \bar{Q}_{16} \\ \bar{Q}_{12} & \bar{Q}_{22} & \bar{Q}_{26} \\ \bar{Q}_{16} & \bar{Q}_{26} & \bar{Q}_{66} \end{bmatrix}_k z \Delta\kappa \quad (9)$$

Based on the coordinate transformation equation for the stress components of the k th ply in the laminate, the maximum principal stress in the k th ply in the laminate can be expressed as

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{bmatrix}_k = \begin{bmatrix} \cos^2 \beta & \sin^2 \beta & 2 \sin \beta \cos \beta \\ \sin^2 \beta & \cos^2 \beta & -2 \sin \beta \cos \beta \\ -\sin \beta \cos \beta & \sin \beta \cos \beta & \cos^2 \beta - \sin^2 \beta \end{bmatrix}_k \begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix}_k \quad (10)$$

By substituting Eqs. (4) and (7) into Eq. (8), the maximum principal stresses for the k th ply in the laminate can be obtained as σ_1^k , σ_2^k , and τ_{12}^k . By substituting the obtained maximum principal stresses for the k th ply into the expressions for three failure criteria (i.e., Tsai-Hill, Tsai-Wu and maximum stress criteria [4–6, 19]), the three corresponding failure indices for the DCB in the folding state can be obtained, namely $I_{f,1}$, $I_{f,2}$, and $I_{f,3}$.

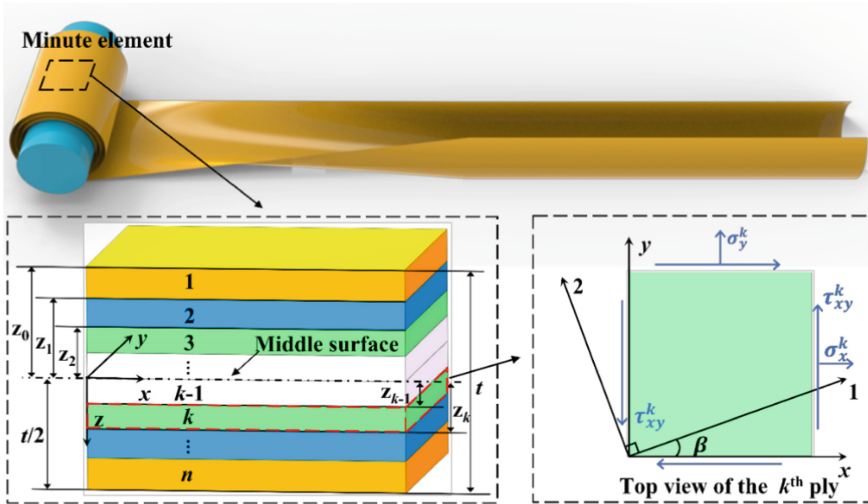


Fig. 2. Schematic diagram of the laminate ply.

3 Optimal Design Schemes

3.1 Problem Description

The optimal design of the DCB is a typical multi-objective optimization problem. The general multi-objective optimization formulation can be described as

$$\begin{aligned} & \text{Minimize } \phi(X) = [\phi_1(X), \phi_2(X) \dots \phi_i(X)] \\ & \text{subject to: } \begin{cases} X = (x_1, x_2 \dots x_m)^T \\ x_g \leq x_g \leq \bar{x}_g (1 \leq g \leq m) \end{cases} \end{aligned} \tag{11}$$

where ϕ represents the objective function, $X=(x_1, x_2 \dots x_j)^T$ are the design variables, i and m are respectively the number of objective functions and design variables, and x_g and \bar{x}_g are the upper and lower limits of the variable space, respectively.

In practical engineering applications of the DCB, the focus is on its axial load capacity, natural frequency, and folding behavior. To facilitate the folding of the DCB and ensure its reliability, the folding moment should be minimized as much as possible. The mass of the deployable composite boom should also be minimized to reduce launch costs. Additionally, the axial load capacity and natural frequency mainly depend on the size of the cross-sectional moment of inertia, so it is necessary to have a sufficiently large moment of inertia to achieve higher axial load capacity and natural frequency. Therefore, the objective functions can be described as

- Objective 1: Minimize the folding moment, as defined in Eq. (6), which can be expressed as

$$\phi_1(X) = T \tag{12}$$

- Objective 2: Maximize the cross-sectional moment of inertia about the u-axis, as defined in Eq. (10), which can be expressed as

$$\phi_2(X) = 1/I_u \quad (13)$$

According to the analytical model established in Sect. 2, the geometric parameters of the DCB hinge have a significant impact on the folding moment and the cross-sectional moment of inertia. Therefore, in this study, the radius of the cross-section, central angle, and thickness of the ply are chosen as the three design variables, which can be expressed as

$$X = (R\theta t_0)^T \quad (14)$$

The optimization design of the DCB needs to satisfy that all three failure indices are less than 1 in the folding state. Therefore, the constraints can be described as

$$\max\{I_{f,1}, I_{f,2}, I_{f,3}\} < 1 \quad (15)$$

Penalty techniques are a commonly used method for handling optimization constraints. However, penalty techniques often require users to define problem-specific parameters, which can affect the performance of the algorithm. In this paper, an adaptive penalty technique is used to handle the constraints of the optimization problem. A penalty factor is set in the fitness function. For any feasible solution, the penalty factor is 1; for any infeasible solution, the penalty factor is greater than 1. In addition, the penalty factor increases as the degree of constraint violation increases.

3.2 NSGA-III Algorithm

Multi-objective optimization problems provide a set of points called the Pareto optimal set, which represents trade-off solutions between conflicting objectives, unlike single-objective optimization problems that provide only a single optimal solution. Among the many-objective evolutionary algorithms used to obtain Pareto fronts, NSGA-III [20, 21] is considered one of the most powerful genetic algorithms. This algorithm is an improved version of NSGA-II and has been widely used in various engineering design optimization problems over the past decade. In this study, NSGA-III is used to solve the optimization problem. The hyperparameters are defined in Table 1.

3.3 Multi-Objective Optimization Design Framework

The multi-objective optimization design framework for the DCB is presented in Fig. 3, and the optimization process is implemented using a high-performance Python-coded genetic algorithm toolbox called Geatpy. The material properties and specifications of the DCB, except for the three design variables, are shown in Table 2. To achieve a comprehensive Pareto front, all design variables are set as continuous variables. The variable space and constraint settings for the optimization in this paper are shown in Table 3.

Table 1. NSGA-III algorithm hyperparameter definition.

Initialization	Initialization type	Random
Reproduction	Encoding type	Real value coded
	Number of populations	1500
	Number of generations	200
	Selection type	Well-spread reference points
	Crossover type	Real variable SBX
	Crossover rate	0.7
	Mutation type	Polynomial mutation
	Mutation rate	0.08
	Elitism operator	Active
Termination	Stop criterion	Reach the maximum generation

4 Optimization Results

Using the optimization framework established in Sect. 3, Fig. 4 shows the results of the Pareto front, which consists of 1500 design points. According to Reference [11], the folding moment and inertia moment about the u -axis of the experimental sample are 87.30 N·mm and 6958.5 mm⁴, respectively. Two points were manually selected from the Pareto front: Point A, with an inertia moment about the u -axis of 6958.5 mm⁴, and Point B, with a folding moment of 87.30 N·mm. These two points represent two different types of designs. Point A and the experimental sample have the same inertia moment about the u -axis, but Point A has a 67.3% reduction in folding moment compared to the experimental sample. Point B and the experimental sample have the same folding moment, but Point B has a 119.4% increase in the inertia moment about the u -axis compared to the experimental sample. There are 458 points on the Pareto front between these two designs, providing a range of designs exhibiting different secondary properties. Therefore, the proposed multi-objective optimization design framework can effectively optimize the DCB.

To further analyze the relationships between the variables and the objective functions, three points with special positions were manually selected on the Pareto front, namely Point C at the left front end ($T = 15.4$ N·mm and $I_u = 558.0$ mm⁴), Point D at the inflection point ($T = 32.6$ N·mm and $I_u = 10978.8$ mm⁴), and Point E at the right rear end ($T = 261.2$ N·mm and $I_u = 21992.3$ mm⁴), as shown in Fig. 4. The cross-sectional radius, cross-sectional angle, and ply thickness for Point C were all set to the minimum values, resulting in the minimum folding moment and inertia moment about the u -axis among all points on the Pareto front. For Point D, the cross-sectional radius and angle were set to the maximum values, while the ply thickness was set to the minimum value. For Point E, the cross-sectional radius, cross-sectional angle, and ply thickness were all set to the maximum values, resulting in the maximum folding moment and inertia moment about the u -axis among all points on the Pareto front.

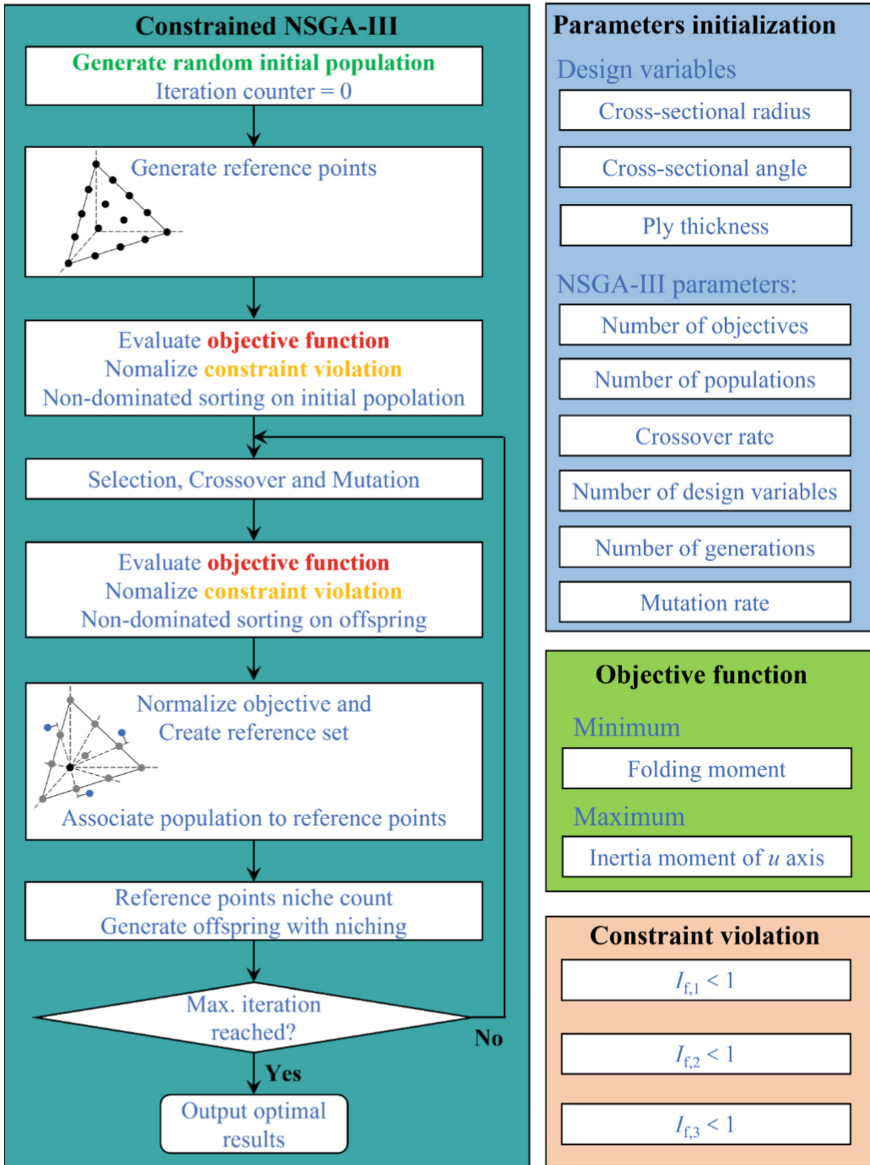


Fig. 3. Multi-objective optimization design framework of the DCB with NSGA-III.

5 Conclusions

Due to its low mass-to-volume ratio, high packaging efficiency, and excellent mechanical properties, the DCB is increasingly used in the aerospace field. The focus of this paper is to propose a multi-objective optimization design framework for optimizing the DCB. The following five important results were obtained in this study:

Table 2. Specifications and properties of T700/epoxy composite ply and DCB [11].

Specifications and properties	Values
Longitudinal elastic modulus E_1 (GPa)	128.61
Transverse elastic modulus E_2 (GPa)	7.52
In-plane shear modulus G_{12} (GPa)	4.82
Inter-laminar shear modulus G_{13} (GPa)	4.50
Inter-laminar shear modulus G_{23} (GPa)	4.50
Poisson's ratio ν_{12}	0.314
Longitudinal tensile strength X_t (MPa)	2103.44
Transverse tensile strength Y_t (MPa)	75.97
Longitudinal compressive strength X_c (MPa)	1233.65
Transverse compressive strength Y_c (MPa)	181.46
In-plane shear strength S_{12} (MPa)	216.36
Density ρ (g/mm ³)	1.6×10^{-3}
Length L (mm)	800
Number of plies n	6
Stacking configuration	[0/45/-45/-45/45/0]

Table 3. Parameter setting of design variables and constraints.

Parameter		Lower bound	Upper bound
Design variables	R (mm)	20	50
	q (deg)	180	270
	t_0 (mm)	0.02	0.04
Constraints	$I_{f,1}$	–	1
	$I_{f,2}$	–	1
	$I_{f,3}$	–	1

- (1) The folding moment required during the folding process of the DCB was determined based on the energy principle, and the failure indices in the folding state were determined using the Tsai-Hill, Tsai-Wu, and maximum stress criteria.
- (2) A multi-objective optimization design framework for optimizing the DCB was proposed based on the classical NSGA-III algorithm.
- (3) The optimization results showed that a total of 458 design solutions with both minimum folding moment and maximum inertia moment about the u -axis were found among 1500 Pareto front points, which were superior to the experimental sample. This helps designers to determine the desired design solutions.

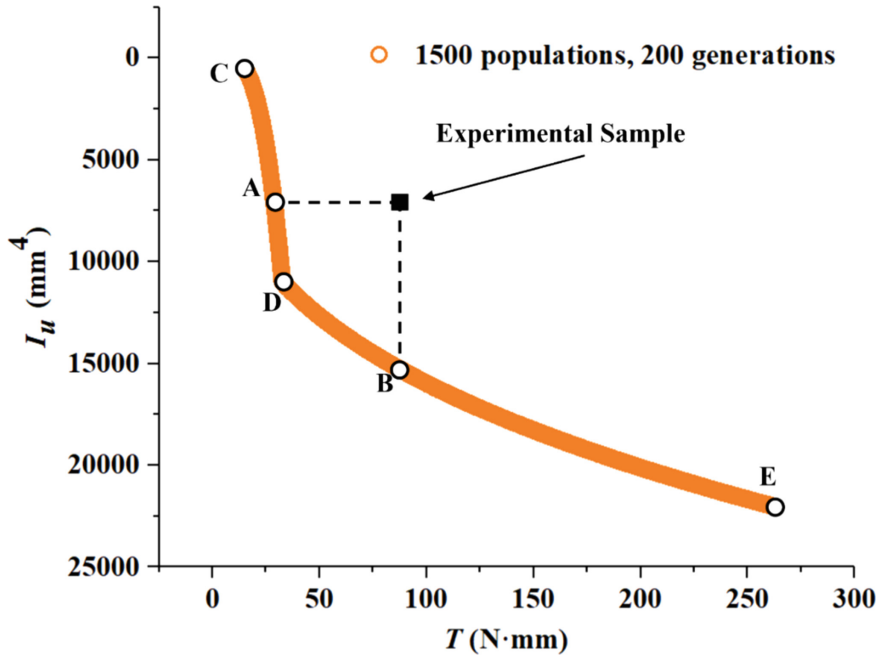


Fig. 4. Optimal design points of the DCB.

- (4) Point A and the experimental sample have the same inertia moment about the u -axis, but Point A has a 67.3% reduction in folding moment compared to the experimental sample. Point B and the experimental sample have the same folding moment, but Point B has a 119.4% increase in the inertia moment about the u -axis compared to the experimental sample.
- (5) The proposed multi-objective optimization design framework can effectively optimize the deployable composite boom.

Overall, this study provides valuable insights and a useful design framework for optimizing the DCB, which has important practical implications for the aerospace industry.

Acknowledgements. This project was supported by the National Natural Science Foundation of China (Grant No. 52275231 and Grant No. 51875026) and the National Defense Basic Research Program of China (Grant No. JCKY2019205C002). The first author acknowledges the China Scholarship Council (Grant No. 202106020152).

Credit Authorship Contribution Statement. Tian-Wei Liu: Investigation, Methodology, Software, Formal analysis, Validation, Writing-original draft. Jiang-Bo Bai: Conceptualization, Supervision, Resources, Funding acquisition, Project administration, Methodology, Writing-review and editing. Nicholas Fantuzzi: Supervision, Writing-review and editing.

References

1. An N, Jia Q, Jin H, et al. Multiscale modeling of viscoelastic behavior of unidirectional composite laminates and deployable structures. *Materials & Design*, 2022, 219: 110754.
2. Bai JB, Liu TW, Yang GH, et al. A variable camber wing concept based on corrugated flexible composite skin. *Aerospace Science and Technology*, 2023, 13: 108318.
3. Bai JB, Xiong JJ, Gao JP, et al. Fabrication and validation of collapsible composite lenticular tubes. *Acta Aeronautica et Astronautica Sinica*, 2011, 32(7):1217–1223.
4. Liu TW, Bai JB, Li SL, et al. Large deformation and failure analysis of the corrugated flexible composite skin for morphing wing. *Engineering Structures*, 2023, 278: 115463.
5. Liu TW, Bai JB, Lin QH, et al. An analytical model for predicting compressive behaviour of composite helical Structures: Considering geometric nonlinearity effect. *Composites Structures*, 2021, 255:112908.
6. Bai JB, Liu TW, Wang ZZ, et al. Determining the best practice – Optimal designs of composite helical structures using Genetic Algorithms. *Composites Structures*, 2021, 268:113982.
7. Liu TW, Bai JB, Fantuzzi N, et al. Multi-objective optimisation designs for thin-walled deployable composite hinges using surrogate models and Genetic Algorithms. *Composites Structures*, 2022, 280:114757.
8. Sakovsky M, Pellegrino S. Closed cross-section dual-matrix composite hinge for deployable structures. *Composite Structures*, 2019, 208:784–795.
9. Liu TW, Bai JB. Folding behaviour of a deployable composite cabin for space habitats-part 1: Experimental and numerical investigation. *Composite Structures*, 2022, 302: 116244.
10. Liu TW, Bai JB. Folding behaviour of a deployable composite cabin for space habitats-Part 2: Analytical investigation. *Composite Structures*, 2022, 297: 115929.
11. Liu TW, Bai JB, Fantuzzi N. Analytical models for predicting folding behaviour of thin-walled tubular deployable composite boom for space applications. *Acta Astronautica*, 2023, 208:167–178.
12. Yang H, Guo HW, Wang Y, et al. Analytical solution of the peak bending moment of an M boom for membrane deployable structures. *International Journal of Solids and Structures*, 2020, 206:236–246.
13. Liu TW, Bai JB, Fantuzzi N. Folding behavior of the thin-walled lenticular deployable composite boom: Analytical analysis and many-objective optimization. *Mechanics of Advanced Materials and Structures*, 2022: 1–19.
14. Yang H, Lu F, Guo H, et al. Design of a new N-shape composite ultra-thin deployable boom in the post-buckling range using response surface method and optimization. *IEEE Access*, 2019, 7: 129659–129665.
15. Yang H, Liu L, Guo H, et al. Wrapping dynamic analysis and optimization of deployable composite triangular rollable and collapsible booms. *Structural and Multidisciplinary Optimization*, 2019, 59: 1371–1383.
16. Yang H, Guo H, Liu R, et al. Coiling and deploying dynamic optimization of a C-cross section thin-walled composite deployable boom. *Structural and Multidisciplinary Optimization*, 2020, 61: 1731–1738.
17. Zhang Z, Zhou H, Ma J, et al. Space deployable bistable composite structures with C-cross section based on machine learning and multi-objective optimization. *Composite Structures*, 2022, 297: 115983.
18. Iqbal K, Pellegrino S, Daton-Lovett A. Bistable composite slit tubes. *IUTAM-IASS Symposium on Deployable Structures*. Cambridge, UK, 1998.
19. Guo S, Li X, Liu T, et al. Parametric study on low-velocity impact (LVI) damage and compression after impact (CAI) strength of composite laminates. *Polymers*, 2022, 14(23): 5200.

20. Deb K, Jain H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: solving problems with box constraints. *IEEE Transactions on Evolutionary Computation*, 2014, 18(4):577–601.
21. Jain H, Deb K. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part II: handling constraints and extending to an adaptive approach. *IEEE Transactions on Evolutionary Computation*, 2014, 18(4):602–622.



Highly Efficient Multi-resolution Topology Optimization Based on the Kriging-Interpolation Network

Wenliang Qian^{1,2} and Hui Li^{1,2}(✉)

¹ School of Civil Engineering, Harbin Institute of Technology, Harbin, China
lihui@hit.edu.cn

² Key Lab of Smart Prevention and Mitigation of Civil Engineering Disaster of the Ministry of Industry and Information Technology, Beijing, China

Abstract. Topology optimization aims to search for the optimal material distribution with a prescribed volume fraction. Recently, to reduce the computational cost of finite element analysis, multi-resolution topology optimization (MTO) has been proposed to decouple the finite elements and density elements. However, MTO introduces new problems with checkerboard patterns and numerous design variables, which hinder its popularity. To overcome these problems of MTO and significantly improve the computational efficiency, an efficient multi-resolution topology optimization method based on the Kriging-Interpolation network (MTO-KIN) is proposed in this paper. In the proposed MTO-KIN, a customized single-layer neural network (Kriging-Interpolation network) is designed to express the topology description function of the design domain, avoiding the checkerboard patterns without filtering techniques or higher-order elements and reducing the design space. Several two-dimensional and three-dimensional numerical examples of topology optimization problems with compliance minimization are studied to demonstrate the effectiveness of the proposed method. The results show that the proposed MTO-KIN can obtain better optimization results than MTO. Meanwhile, compared with MTO, an acceleration of about 11 to 18 times can be achieved.

Keywords: Multi-resolution topology optimization · Kriging-Interpolation network · Knot mesh · Topology description function

1 Introduction

To obtain the optimum structural performance, researchers have shown an increased interest in topology optimization since the pioneering work [1]. The goal of topology optimization is to find the optimal material distribution with the prescribed volume fraction to minimize the objective [2–4]. However, topology optimization is still not popular in the engineering field, especially for large-scale problems, due to the need for high computational resources.

To overcome this computational burden, a series of works have been done to reduce computational cost or improve computational efficiency by saving the computational costs of updating the design variables or calculating the governing equations. In terms of design variables, Guest et al. [5] proposed a decoupling strategy to separate the design variable and the analysis mesh by the Heaviside projection method. Kim et al. proposed the reducible design variable method (RDVM) to save the computational costs of updating design variables by not considering the design variables that have converged [6]. In terms of solving the governing equations, iterative solvers, for example, a multigrid preconditioned conjugate gradients (MGCG) solver, are adopted to fast solve large-scale 3-D structural topology optimization problems on a standard PC [7, 8]. Another approach is to use a multiresolution topology optimization (MTOP) scheme to obtain high resolution designs by using a coarser discretization for finite elements and finer discretization for both density elements and design variables [9], which is a universal topology optimization method. According to this approach, finite element analysis will be performed in a relatively coarse mesh, which will significantly save time in solving the governing equations. However, MTOP has two disadvantages: the QR-patterns and numerous design variables. The QR-patterns means that the optimization results by this method consisted of artificially stiff regions named the QR-patterns as shown in Fig. 1, which seriously hurts the application of MTOP. The QR-patterns arise mainly due to the discontinuity of the density distribution in the element leading to the inability of the low-order shape functions to accurately model its displacement. To conquer this problem, several methods have been proposed, including the use of filtering (density projection) and the use of higher polynomial shape functions. When using the filtering in MTOP, the density field will be smoothed equivalent to imposing a restriction on minimum feature size. However, the filter radius is not easy to determine an appropriate value, where the large filter radius will restrict the design field from expressing a high order material distribution [10]. Meanwhile, the additional computational costs of computing filter weights and performing convolution operations are also an obstacle, especially on large scale problems. When using the higher polynomial elements in MTOP, the QR-patterns can be avoided in the simplest way by accurately modeling the displacement field [10–12]. But, the use of high order elements will increase the degrees of freedom so as to take more time to solve the governing equation, which contradicts the original purpose of MTOP. Another drawback, i.e. numerous design variables, arises from the introduction of the density mesh in MTOP. To overcome this problem, an adaptive isosurface variable grouping (aIVG) is proposed to group design variables of similar grouping criteria into a single grouped design variable [13], which will significantly reduce computation time in optimization.

In recent years, deep learning has made successive breakthroughs in several fields due to its powerful non-linear representation and efficient optimization algorithm [14–16]. Therefore, for the sake of completeness, it is also necessary to mention that several machine learning-based studies for topology optimization have been proposed. For example, a deep learning-based method is proposed to predict an optimized structure by designing convolutional neural networks to model the relationship between the optimization results and some parameters of the optimal topologies [17]. A novel two-phase methodology based on deep learning also is proposed to find the relationship between

the final density values of each finite element and the results of the first few iterations of the SIMP method [18]. Some reparameterization methods based on deep learning also be studied, where the design variables are the parameters of the network [19–21].

Since the above studies based deep learning either face generalizability problems, or interpretability problems, or computational efficiency problems, a general and efficient topology optimization method is worth investigating. As mentioned above, MTOP is a universal topology optimization method, and significantly save time in solving the governing equations. However, the QR-patterns and numerous design variables will lose the computational advantage of MTOP, preventing the further popularity of MTOP. So, once these problems can be overcome, MTOP will be the most promising approach to help topology optimization become popular in the engineering field.

Motivated by the aforementioned problems, this study aims to provide a novel method to avoid the QR-patterns and reduce the number of design variables without losing the computational advantage of MTOP. As is well-known, level set methods can avoid the checkerboard pattern, similar to the QR-patterns, without the filter technology. Topology description functions (TDF) based on the Kriging interpolation model [22, 23] are a class of explicit level-set methods to solve topology optimization, which can not only avoid the checkerboard phenomenon but also significantly reduce the number of design variables compared with element-wise density-based topology optimization. However, the methods based on the Kriging interpolation model are non-gradient methods and thus require high computational resources, since the gradient calculation is very difficult due to the fact that the material distribution is obtained by cutting the TDF with a certain threshold [23, 24]. Therefore, when the gradients can be easily calculated, the Kriging interpolation model will be a potential way to avoid the QR-patterns while reducing the number of design variables.

The remainder of this article is organized as follows. Section 2 describes the proposed multi-resolution topology optimization based on the Kriging-Interpolation network (MTOP-KIN) in detail. Section 3 conducts a series of 2D and 3D experiments to validate the proposed method and discusses the results. Section 4 presents a discussion about the effect of the Multiresolution Model on MTOP-KIN and the influence of the Knot mesh and Density mesh on MTOP-KIN. Section 5 summarizes and concludes the paper.

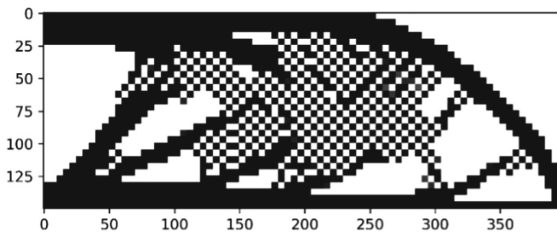


Fig. 1. The QR-patterns in optimization design obtained by MTOP

2 Method

A novel multi-resolution topology optimization using the Kriging-Interpolation network is proposed in this study, which is referred to as MTOP-KIN. The proposed MTOP-KIN can avoid QR patterns without filtering techniques or high-order elements and reduce the design space, thereby improving the computational efficiency for the topology optimization problems. It can divide into three parts: Multiresolution model, Physical model solver, and Model optimization, as shown in Fig. 2. The three parts of the proposed MTOP-KIN are described in detail in Sects. 2.1, 2.2, and 2.3 below, respectively.

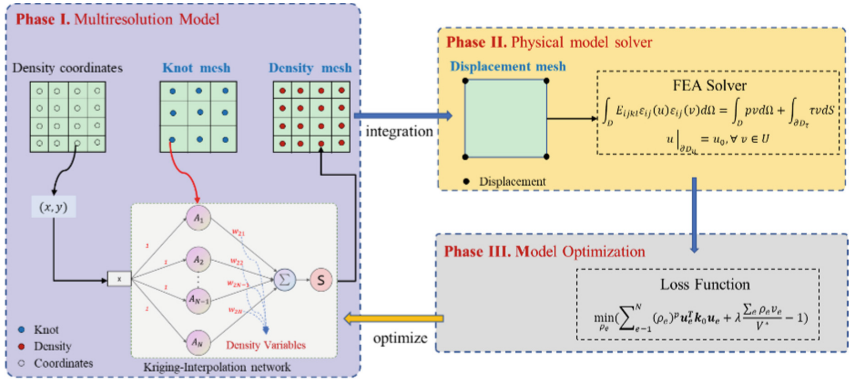


Fig. 2. The framework for the proposed multi-resolution topology optimization

2.1 Multiresolution Model

The goal of the Multiresolution model is to establish the relationship between the Knot mesh, Density mesh, and Displacement mesh, as shown in Fig. 3. By using these meshes, the Knot mesh to establish a topology description function, the Density mesh to determine density variables in the Displacement mesh, and Displacement mesh to perform finite element analysis can be decoupled from each other.

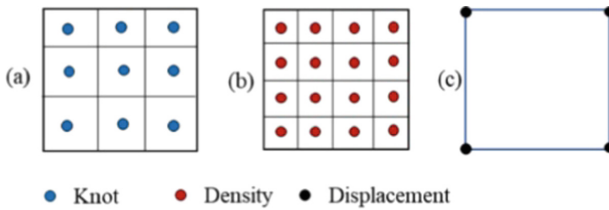


Fig. 3. The three different meshes for the multiresolution model: (a) Knot mesh, (b) Density mesh, (c) Displacement mesh

In the conventional MTOP, the design variable is equal to the number of density elements in Density mesh. But, in the MTOP-KIN, the number of design variables

is equal to the number of elements in the Knot mesh. Since the number of elements in the Knot mesh is usually smaller than the number of density variables in Density mesh, e.g. 9%, the Knot mesh can reduce the number of design variables than directly using the Density mesh. Also, as in conventional MTOP, the Displacement mesh can significantly reduce the time of solving the governing equation as the finite element analysis is performed in a relatively coarse mesh. For example, in Fig. 3, the number of elements in the Displacement mesh is only 6.25% of those in the Density mesh.

To avoid QR-patterns without losing the computational advantage of MTOP and reduce the design space, the Knot mesh adopts a Kriging-Interpolation method based a single-layer neural network to obtain density variables. Using this Kriging-Interpolation network, the topology description function (TDF) $f(x)$ is a combination of nonlinear activation function $A_i(x)$ in the hidden layer weighted by weights w_{ij} in the hidden layer. Then the $f(x)$ is acquired by the forward propagation of the network as follows

$$f(x) = \sum_i w_{2i} A_i(w_{i0}x) = \sum_i w_{2i} e^{-\alpha^2 \|w_{i0}x - x_i\|^2} \quad (1)$$

Here, $A_i(w_{i0}x) = e^{-\alpha^2 \|w_{i0}x - x_i\|^2}$ is the customized nonlinear activation function, $w_{i0} = 1$ and w_{2i} denote the corresponding weights of the single-layer neural network, x_i and α denote the center coordinates of elements in the Knot mesh and tuning constant, respectively. The α denotes the sharing degree between knots by

$$\alpha = \frac{1}{\tilde{r}} \sqrt{\ln(1/c)} \quad (2)$$

where \tilde{r} is the minimum Euclidean distance between knots and c is a constant setting as 0.5 in this study. The learnable weights w_{2i} , i.e. design variables, will be updated by a backpropagation algorithm, where the number is equal to the numbers of elements in the Knot mesh. So, the number of design variables can be greatly reduced, which will improve the computational efficiency of MTOP. Meanwhile, since the only learnable weights w_{2i} represents the weights of different basis functions, this network is interpretable. The architecture of the proposed Kriging-Interpolation network is presented in Fig. 4, where the input is the center coordinates in the Density mesh and the output are the corresponding density values.

Here, the sigmoid activation function is adopted to convert the value $f(x)$ to a continuous density value $\rho(x) \in [0, 1]$ as:

$$\rho(x) = \frac{1}{1 + e^{-f(x)}} \quad (3)$$

This activation function will overcome the difficulty of calculating the gradient of the objective function with respect to the TDF by the cutting operation in the classical level-set method, where the aim of this cutting operation is to determine the material properties in the design domain by the relationship between the value $f(x)$ and the threshold value T_S as follows.

$$\begin{aligned} f(x) &\geq T_S, & \rho(x) &= 1 \\ f(x) &< T_S, & \rho(x) &= 0 \end{aligned} \quad (4)$$

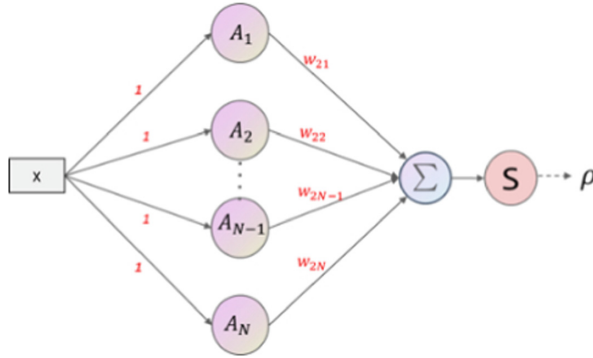


Fig. 4. The architecture of the proposed Kriging-Interpolation network

By using the proposed Kriging-Interpolation network, the density distribution in the density mesh will change smoothly, which can be observed in Fig. 5. The smaller the number of knots in the knot mesh, the smoother the change of density distribution in the Density mesh. Meanwhile, the values of density function $\rho(x)$ is a continuous function within the range of $[0, 1]$ with convenience in the calculation of gradient, which also improves computational efficiency of MTOP by using gradient-based optimization methods, such as stochastic gradient descent (SGD), AdaGrad [25], and Adam [26].

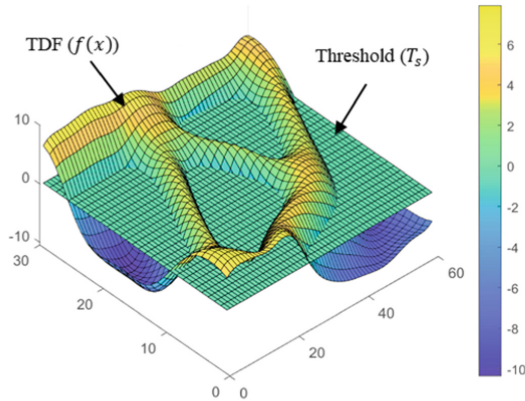


Fig. 5. The topology description function by the proposed Kriging-Interpolation network

2.2 Physical Model Solver

The goal of the Physical model solver is to solve the state field, involving the integration of the element stiffness matrix in the displacement mesh, material penalization model, and finite element analysis.

The integration of the element stiffness matrix is to compute the stiffness by considering the corresponding contribution of the density element over the displacement

element domain. Figure 6 presents the superposed meshes of the Density mesh and Displacement mesh, where a_i^e is the area or volume of the i th density element in the e th displacement element and ρ_i^e is the density value of the i th density element in the e th displacement element. Using the Gauss-Legendre quadrature, a modified displacement element stiffness can be expressed as

$$\mathbf{k}^e = \int_{\Omega^e} \mathbf{B}^{eT} \mathbf{C}^e \mathbf{B}^e \cong \sum_{i=1}^{N_n^e} \left(\mathbf{B}^{eT} \mathbf{C}^e \mathbf{B}^e \right)_i a_i^e \quad (5)$$

where \mathbf{k}^e , Ω^e , \mathbf{B}^e , \mathbf{C}^e , and N_n^e is the element stiffness matrix, design domain, strain-displacement matrix, constitutive matrix, and number of density element of the e th displacement element, respectively.

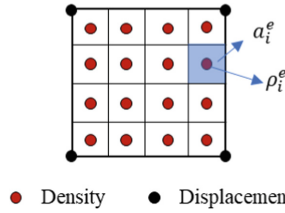


Fig. 6. The superposed meshes of the density mesh and displacement mesh

The material penalization model is to convert the integer optimization problem to a continuous optimization problem. In this study, the popular model named solid isotropic material with penalization (SIMP) is adopted [2]. And its formulation is $E_i(\rho_i) = \rho_i^p E_0$, where E_0 denotes the Young's modulus and p is the penalization factor, set to 3 in this study. Because the constitutive matrix depends on the modified Young's modulus of the material, the modified displacement element can be reexpressed by one value ρ_i^e as follows:

$$\mathbf{k}^e \cong \sum_{i=1}^{N_n^e} (\rho_i^e)^p \left(\mathbf{B}^{eT} \mathbf{C}^e \mathbf{B}^e \right)_i a_i^e \quad (6)$$

The finite element analysis is to solve the state field, where the governing equation of the state field can be expressed by the energy bilinear form and the load linear form, and its formulation as

$$\int_D C_{ijkl} \varepsilon_{ij}(\mathbf{u}) \varepsilon_{ij}(\mathbf{v}) d\Omega = \int_D \mathbf{p} \mathbf{v} d\Omega + \int_{\partial D_\tau} \boldsymbol{\tau} \mathbf{v} dS \quad (7)$$

where C_{ijkl} is the fourth-order constitutive tensor, ε_{ij} is the linearized strains, \mathbf{p} is the body forces, and $\boldsymbol{\tau}$ is the traction forces. By discretizing the design domain with displacement elements, i.e. the Displacement mesh, the governing equation can be formulated by the following discrete equations:

$$\mathbf{K} \mathbf{U} = \mathbf{F} \quad (8)$$

where \mathbf{U} , \mathbf{K} and \mathbf{F} are the global displacement, stiffness matrix and force vectors, respectively. The \mathbf{K} is obtained by assembling the element stiffness matrix \mathbf{k}^e , i.e. $\mathbf{K} = \sum_e \mathbf{k}^e$.

2.3 Model Optimization

The goal of the Model optimization is to search the optimal parameters of MTOP-KIN by optimizing the loss function, which aims at converting the constrained optimization problem of topology optimization into a network optimization by the Lagrange multiplier method. For the specific compliance minimization problems, the loss function can be written as follows:

$$L = \mathbf{U}^T \mathbf{K} \mathbf{U} + \lambda \left(\frac{V(\rho^e)}{V_0} - V_f^c \right)^2 \quad (9)$$

where λ , V_f^c , $V(\rho^e)$ and V_0 are Lagrangian penalty factor, the prescribed volume fraction, the material volume and volume of design domain, respectively. To control the order of magnitude of the different terms, the normalization is operated in Eq. (9), which is similar to that TOuNN in [19].

$$L = \frac{\mathbf{U}^T \mathbf{K} \mathbf{U}}{J_0} + \lambda \left(\frac{V(\rho^e)}{V_0 \times V_f^c} - 1 \right)^2 \quad (10)$$

where J_0 is the objective function value of the randomly initialized MTOP-KIN before network training. In order to update the network parameters w_{2i} , the gradient can be calculated using the chain derivative rule as follows:

$$\frac{\partial L}{\partial \theta_{2i}} = \sum_e \frac{\partial L}{\partial \rho^e} \frac{\partial \rho^e}{\partial f} \frac{\partial f}{\partial \theta_{2i}} \quad (11)$$

$$\frac{\partial L}{\partial \rho^e} = -\frac{1}{J_0} (\rho_j^e)^{p-1} \mathbf{u}^{eT} \left(\mathbf{B}^{eT} \mathbf{C}^e \mathbf{B}^e \right) |_{j a_j^e} \mathbf{u}^e + \frac{2\lambda}{V_f^c} \left(\frac{V(\rho^e)}{V_0 \times V_f^c} - 1 \right) \quad (12)$$

$$\frac{\partial \rho^e}{\partial f} = \frac{e^{-f(x)}}{(1 + e^{-f(x)})^2} \quad (13)$$

$$\frac{\partial f}{\partial w_{2i}} = e^{-\alpha^2 \|\mathbf{x} - x_i\|^2} \quad (14)$$

where Eqs. (13) and (14) can be obtained directly by the backpropagation algorithm. Adam [26] is one of the most popular optimization algorithms in deep learning due to its good performance. So, Adam is adopted to optimize the MTOP-KIN. In optimization, checking the convergence is also important, which can affect the computational efficiency of topology optimization. The change of the loss function in deep learning is often used to determine the convergence, so a natural idea is to use the change of the loss function as the convergence criterion for the proposed MTOP-KIN.

Here, the standard deviation of objective function values of successive iterations is adopted to determine the convergence status of MTOP-KIN as follows:

$$\bar{L}_i = \frac{L_{i-(NC-1)} + L_{i-(NC-2)} + \cdots + L_{i-1} + L_i}{NC} \quad (15)$$

$$\sigma_i = \sqrt{\frac{1}{NC-1} \sum_{j=i-(NC-1)}^i (L_j - \bar{L}_i)^2} \quad (16)$$

where NC denotes the number of successive iterations for checking the convergence status and $i \geq NC$. When the standard deviation σ_i is less than a convergence tolerance value $\tilde{\sigma}$, the convergence is considered to be reached.

3 Numerical Experiments

To investigate the performance of the proposed MTOP-KIN four structures are employed as examples, including three 2D structures and one 3D structure. In all examples, the volume fraction, the density penalization power, Young's modulus, and Poisson's ratio are set as 0.5, 3, 1.0, and 0.3. For the convenience of the subsequent discussion, the ratio of the number of the Knot mesh elements to the number of the Density mesh elements and the ratio of the number of the Density mesh elements to the number of the Displacement mesh elements are defined as follows

$$R_1 = N_{knot}/N_{density} \quad (17)$$

$$R_2 = N_{density}/N_{displacement} \quad (18)$$

where N_{knot} , $N_{density}$, and $N_{displacement}$ denote the number of elements in the Knot mesh, Density mesh, and Displacement mesh, respectively. In all four different examples below, the R_1 and R_2 are set as 0.3 and 5, respectively. The learning rate of Adam is set as 0.3 and coefficients used for computing running averages of gradient and its square are set as $\beta_1 = 0.5$ and $\beta_2 = 0.999$, respectively. The number of successive iterations and convergence tolerance value for convergence condition is set as 15 and 0.0005, respectively.

3.1 2D Examples

The design domains of all 2D examples are assumed to be discretized by using a mesh of 80×30 square finite elements (Q4 elements) in the Displacement mesh of MTOP-KIN and MTOP, where the Density mesh is 400×150 and the Knot mesh is 120×45 . The design domains and boundary conditions are shown in Fig. 7. To verify the performance of the proposed MTOP-KIN method, MTOP is selected as a baseline [9] and the filter radius is 5. Figure 8 gives the convergence curves of MTOP-KIN, from which it can be seen that the convergence is achieved in about 80 iterations.

These curves also illustrate that MTOP-KIN can achieve convergence performance under the same parameters for different examples. The optimization results generated by the proposed MTOP-KIN and MTOP are shown in Fig. 9. As presented, MTOP-KIN allows for clearer results without the use of filter techniques, i.e. gray regions are much fewer to compare with MTOP. This phenomenon indicates the feasibility of MTOP-KIN. Meanwhile, the QR-patterns also did not appear in the optimization results, which illustrates that the proposed method can avoid this problem without filtering techniques or high-order elements.

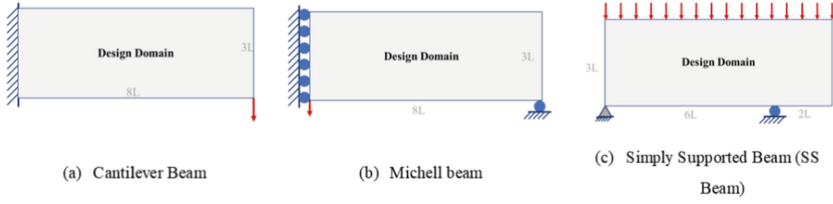


Fig. 7. The design domains and boundary conditions of the cantilever beam, Michell beam and SS beam

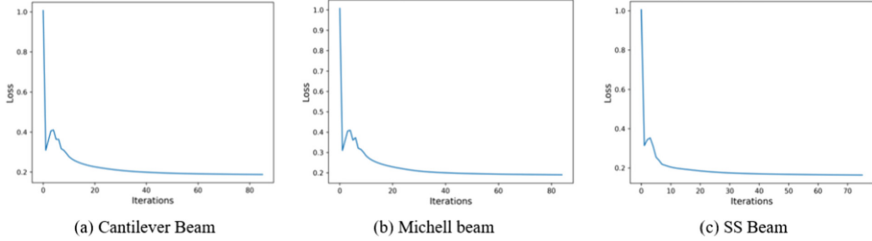


Fig. 8. The convergence curves of loss function in the training process on the cantilever beam, Michell beam and SS beam

Table 1 presents a summary of the proposed MTOP-KIN and MTOP, including the number of design variables, objective function values, and calculation time, on all test examples. These results show that the number of design variables of MTOP-KIN is much smaller than those of MTOP because the Knot mesh and nonlinear activation function $A_i(x)$ are adopted in MTOP-KIN. From these results, it can also be found that the compliance of the proposed method is smaller than that of MTOP, and a speedup of approximately 15 to 18 times is achieved compared with the MTOP method. The reason for the faster speed is mainly due to the reduction of design variables and the absence of filtering techniques.

3.2 3D Cantilever Beam

In this subsection, a 3D compliance minimization problem is also performed to verify that the proposed method can be applied not only to 2D but also to 3D. Figure 10 gives the 3D Cantilever Beam subjected to the unit loads. The design domain is discretized by using a mesh of $42 \times 12 \times 4$ cube finite elements in the Displacement mesh of MTOP-KIN and MTOP. Here, the Density mesh is $210 \times 60 \times 20$ and the Knot mesh is $63 \times 18 \times 6$. Using the MTOP-KIN, the convergence is achieved in about 70 iterations, as shown in Fig. 11. Figure 12 presents the optimization results by the MTOP-KIN and MTOP, indicating the proposed method is also suitable for 3D structures. From the optimization results, more thin members exist in the optimization result obtained by MTOP-KIN which means that better structures may be found by the proposed method. This can be verified by the compliance value, as listed in Table 2. The difference in the compliance value indicates that the stiffness of structure obtained by MTOP-KIN is much greater than the stiffness of that obtained by MTOP. Table 2 also shows that

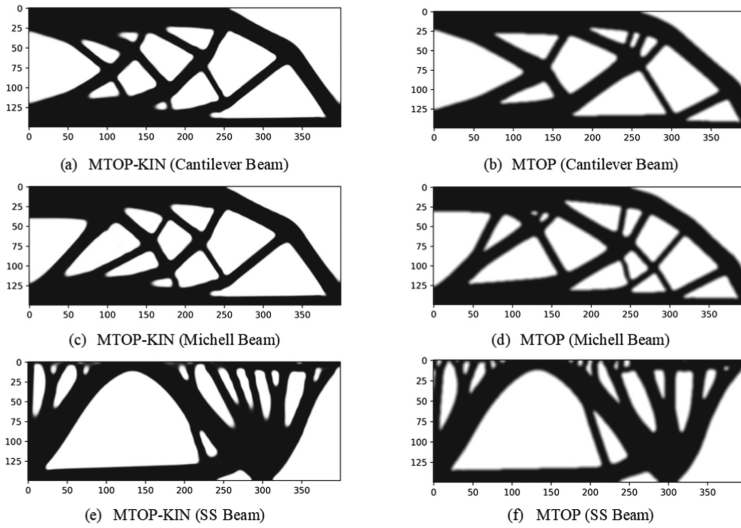


Fig. 9. Optimization results of the cantilever beam, Michell beam and simply supported (SS) beam

Table 1. Summary of the tested examples on the cantilever beam, Michell beam and simply supported (SS) beam

Case	Methods	No. of design variables	Obj. func. value	Time (s)
Cantilever beam	MTOP-KIN	5400	134.88	5.08
	MTOP	60000	140.51	80.99
Michell beam	MTOP-KIN	5400	142.10	5.18
	MTOP	60000	150.60	81.34
SS beam	MTOP-KIN	5400	22605.42	4.20
	MTOP	60000	23155.92	78.28

the proposed MTOP-KIN has a speedup of approximately 11 times compared with the MTOP method.

4 Discussions

The results of numerical experiments had illustrated the proposed MTOP-KIN can obtain better optimization results with less optimization time compared with MTOP, as outlined in Sect. 3. However, the key model and parameters still need to be investigated to explore their effect, for example, the effect of the multiresolution model and the influence of the number of knot elements and density elements on MTOP-KIN. So, first, the MTOP-KIN without the Multiresolution Model, named TOP-KIN, is investigated on the Cantilever Beam, Michell beam, and SS Beam. Then, the influence of the Knot mesh and Density

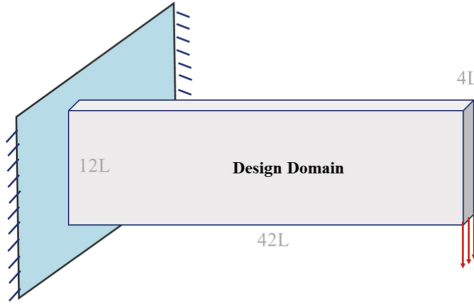


Fig. 10. The design domains and boundary conditions of the 3D cantilever beam

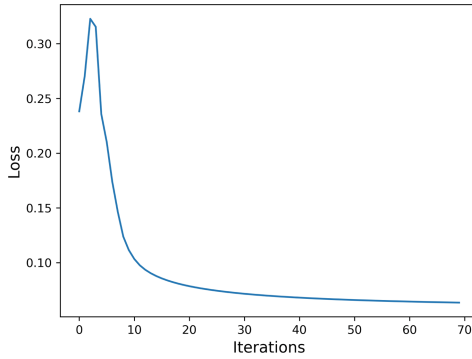


Fig. 11. The convergence curves of loss function in the training process on the 3D cantilever beam

Table 2. Summary of the tested examples on the 3D cantilever beam

Case	Methods	No. of design variables	Obj. func. value	Time (s)
3D cantilever beam	MTOP-KIN	6804	2703.69	35.59
	MTOP	252000	3742.08	396.75

mesh on MTOP-KIN is investigated on the SS Beam. Because the optimization result of the SS Beam has more members, the change of the optimized structure can be easily observed.

4.1 Effect of the Multiresolution Model on MTOP-KIN

The proposed MTOP-KIN comprises three parts: Multiresolution Model, Physical model solver, and Model optimization, as shown in Fig. 1. The most important part is Multiresolution Model to model the relationship between the Knot mesh, Density mesh, and Displacement mesh, where the Density mesh is the bridge that connects Knot mesh and

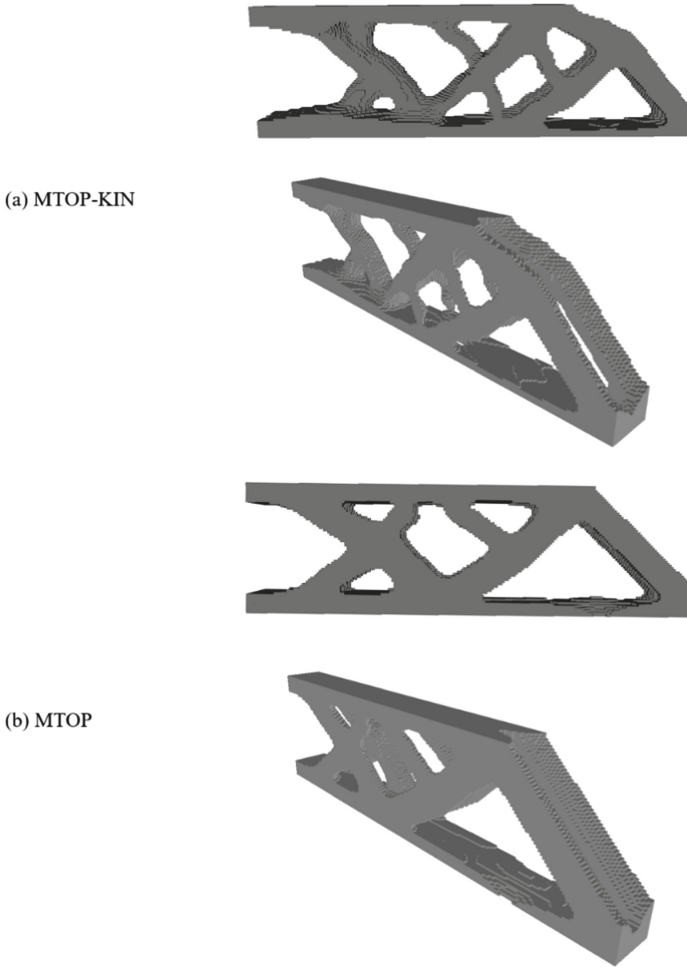


Fig. 12. Optimization results of the cantilever beam, Michell beam and simply supported (SS) beam

Displacement mesh. So, only using the Knot mesh and Displacement mesh in MTOP-KIN will be able to explain the role of the Multiresolution Model, this modified topology optimization method is called TOP-KIN.

Here, the number of elements in the Knot mesh and Displacement mesh of TOP-KIN is 120×45 and 80×30 , which equal to those of MTOP-KIN, as described in Sect. 3.1. Figure 13 and Table 3 give the optimization results and the performance comparisons, respectively. From the results of Fig. 13, it can be found that MTOP-KIN can obtain a clearer optimization result with a higher resolution which means more thin members appear in the result. Table 3 shows that the optimization results have a much lower compliance value and the training time is only about twice that of the TOP-KIN.

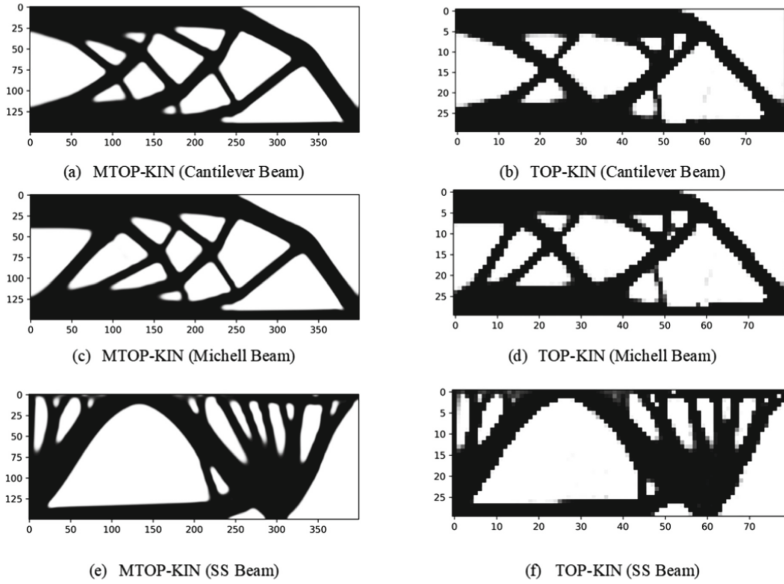


Fig. 13. Optimization results of the cantilever beam, Michell beam, and SS beam by MTOP-KIN and TOP-KIN

Table 3. Summary of the tested examples with MTOP-KIN and TOP-KIN

Case	Methods	No. of design variables	Obj. func. value	Time (s)
Cantilever beam	MTOP-KIN	5400	134.88	5.08
	TOP-KIN	5400	142.37	2.39
Michell beam	MTOP-KIN	5400	142.10	5.18
	TOP-KIN	5400	148.98	2.51
SS beam	MTOP-KIN	5400	22605.42	4.20
	TOP-KIN	5400	23122.41	2.20

Although the training time of MTOP-KIN increases compared with TOP-KIN, the TOP-KIN will consume more time to achieve the same resolution compared with the proposed MTOP-KIN. To illustrate this phenomenon, the Knot mesh and Displacement mesh in TOP-KIN are adjusted to 120×45 and 400×150 , where the aim here is to make the number of elements in the Displacement mesh of TOP-KIN equal to the number of elements in the Density mesh of MTOP-KIN for obtaining the same resolution. For ease of description, the TOP-KIN of this configuration is noted as TOP-KIN-5. Figure 14 illustrates the MTOP-KIN can get almost the same optimization results with less Displacement mesh, meaning the optimization time can be significantly reduced. Meanwhile, Fig. 14(e) and (f) also clearly illustrate that MTOP-KIN gives better optimization results. These conclusions can be verified by the results in Table 4, which shows that

MTOP-KIN can not only get a much lower compliance value but also achieve a speedup of approximately 11 times.

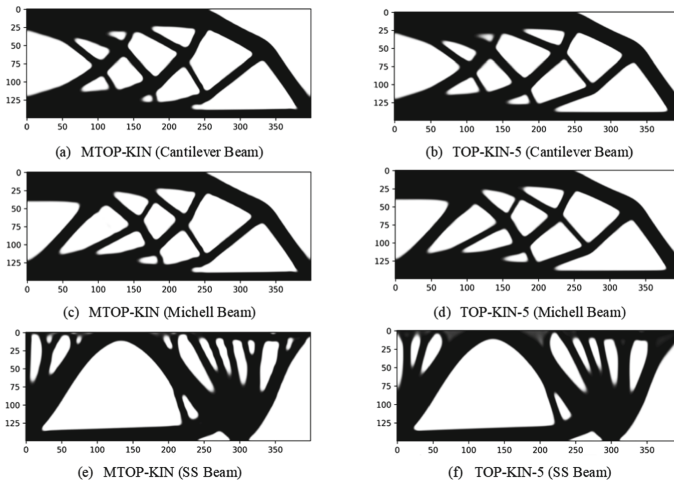


Fig. 14. Optimization results of the cantilever beam, Michell beam and SS beam by MTOP-KIN and TOP-KIN-5

Table 4. Summary of the tested examples with MTOP-KIN and TOP-KIN-5

Case	Methods	No. of design variables	Obj. func. value	Time (s)
Cantilever beam	MTOP-KIN	5400	134.88	5.08
	TOP-KIN-5	5400	138.93	60.72
Michell beam	MTOP-KIN	5400	142.10	5.18
	TOP-KIN-5	5400	148.41	59.05
SS beam	MTOP-KIN	5400	22605.42	4.20
	TOP-KIN-5	5400	22788.34	51.90

4.2 Effect of the Knot Mesh and Density Mesh on MTOP-KIN

The Knot mesh can control the degree of smoothing of the change of density distribution in the Density mesh, and the Density mesh controls the stiffness of the element in the Displacement mesh. Therefore, it is necessary to investigate the effect of the Knot mesh and Density mesh on MTOP-KIN.

The influence of the Knot mesh is first investigated. In order to exclude the effect of the Displacement and Density meshes, the number of elements remains the same, and only the number of the Knot mesh is changed. Eight different Knot mesh configurations were investigated with the ratio of the number of elements of the Knot mesh and Density

mesh are 0.20, 0.25, 0.30, 0.35, 0.37, 0.39, 0.40, and 0.41, respectively. Figure 15 and Table 5 give the optimization results and performance comparisons. These results show that the resolution of the structure becomes higher and objective function values come to be smaller, as the number of elements in the Knot mesh increases. But, the degree of smoothing of the optimization results obtained by MTOP-KIN with $R_1 = 0.40$ and $R_1 = 0.41$ are worse than the results of the other configurations.

Then the influence of the Density mesh is investigated. Similarly, only the ratio of the number of the density mesh elements to the number of the displacement mesh elements is changed with the same Displacement mesh and $R_1 = 0.30$. Six different ratios R_2 were investigated are 2, 3, 4, 5, 6 and 7. From the results of Fig. 16 and Table 5, it can be observed that structures with isolated islands will occur when unsuitable R_2 is adopted. Because the more the number of knots in the Knot mesh means that $\alpha = \frac{1}{r} \sqrt{\ln(1/c)}$ is smaller, the smoothing effect on the Density mesh will decrease, leading to drastic changes or even discontinuity of density in the Density mesh. Fortunately, through the extensive experiments, including 2D and 3D experiments, it can be found that a suitable ratio of elements of the Knot mesh and Displacement mesh should satisfy $R_1 \times R_2 \in [1.3, 1.8]$, which can help the proposed MTOP-KIN generate clear and smooth optimization results (Table 6).

Table 5. Summary of the SS beam by MTOP-KIN with different knot meshes

R_1	Obj. func. value	R_1	Obj. func. value
0.20	23004.08	0.37	22436.90
0.25	22787.96	0.39	22396.11
0.30	22605.42	0.40	22391.78
0.35	22502.78	0.41	22338.00

Table 6. Summary of the SS beam by MTOP-KIN with different density meshes

R_2	Obj. func. value	R_2	Obj. func. value
2	23588.38	5	22605.42
3	23113.61	6	22449.35
4	22812.90	7	22324.90

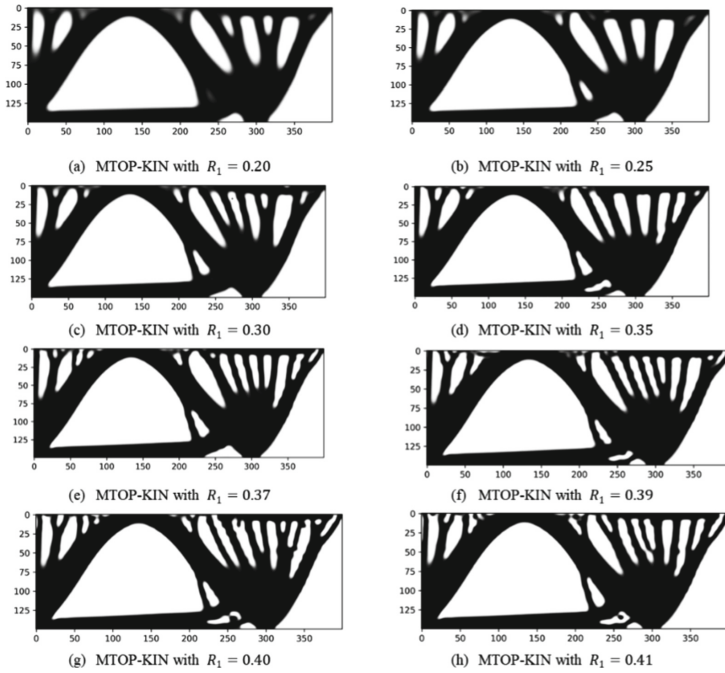


Fig. 15. Optimization results of the SS beam by MTOP-KIN with different knot meshes

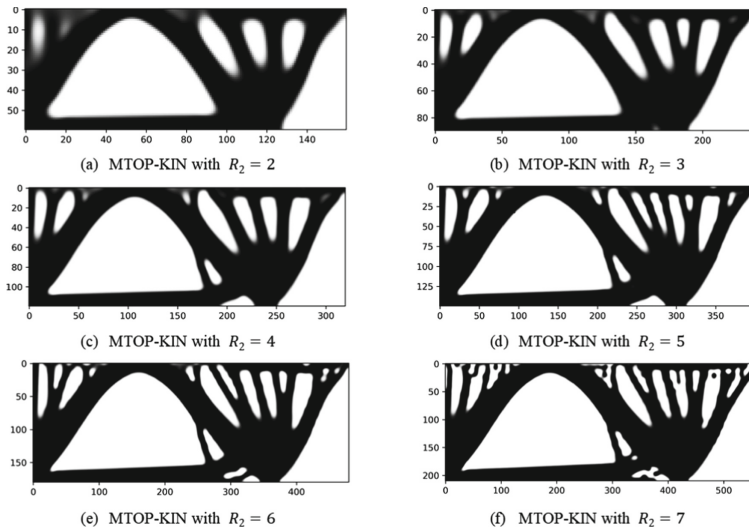


Fig. 16. Optimization results of the SS beam by MTOP-KIN with different density meshes

5 Conclusion

This paper proposes a novel multi-resolution topology optimization based on the Kriging-Interpolation network using a customized single-layer neural network to express the topology description function of a design domain. The proposed can obtain better optimization results using less time compared with MTOP.

The main conclusions are summarized as follows:

- The MTOP-KIN can avoid QR patterns without filtering techniques or high-order elements by the proposed Kriging-Interpolation network to model the relationship between the Knot elements and Density elements, which maintains the computational advantage of multi-resolution topology optimization.
- The proposed Kriging-Interpolation network in MTOP-KIN is interpretable since all parameters in this network are straightforward linked with those in topology description function and have a clear physical sense. Meanwhile, with this network, the design variables of MTOP-KIN can be updated by gradient-based optimization algorithms and GPU.
- The proposed MTOP-KIN can reduce the design space by a novel topology description function and obtain a better optimization result with a speedup of approximately 11 to 18 times compared with the MTOP method for 2D and 3D examples.

Acknowledgement. This study is financially supported by National Natural Science Foundation of China No. 51921006.

References

1. Bendsøe, Martin P. “Optimal shape design as a material distribution problem.” *Structural optimization* 1 (1989): 193–202.
2. Bendsøe, Martin P., and Ole Sigmund. “Material interpolation schemes in topology optimization.” *Archive of applied mechanics* 69 (1999): 635–654.
3. Xie, Yi Min, and Grant P. Steven. “A simple evolutionary procedure for structural optimization.” *Computers & structures* 49.5 (1993): 885–896.
4. Yulin, Mei, and Wang Xiaoming. “A level set method for structural topology optimization and its applications.” *Advances in Engineering software* 35.7 (2004): 415–441.
5. Guest, James K., and Lindsey C. Smith Genut. “Reducing dimensionality in topology optimization using adaptive design variable fields.” *International journal for numerical methods in engineering* 81.8 (2010): 1019–1045.
6. Kim, Sun Yong, Il Yong Kim, and Chris K. Mechefske. “A new efficient convergence criterion for reducing computational expense in topology optimization: reducible design variable method.” *International journal for numerical methods in engineering* 90.6 (2012): 752–783.
7. Amir, Oded, Niels Aage, and Boyan S. Lazarov. “On multigrid-CG for efficient topology optimization.” *Structural and Multidisciplinary Optimization* 49 (2014): 815–829.
8. Peetz, Darin, and Ahmed Elbanna. “On the use of multigrid preconditioners for topology optimization.” *Structural and Multidisciplinary Optimization* 63 (2021): 835–853.
9. Nguyen, Tam H., et al. “A computational paradigm for multiresolution topology optimization (MTOP).” *Structural and Multidisciplinary Optimization* 41 (2010): 525–539.

10. Gupta, Deepak K., Matthijs Langelaar, and Fred van Keulen. "QR-patterns: artefacts in multi-resolution topology optimization." *Structural and Multidisciplinary Optimization* 58 (2018): 1335–1350.
11. Groen, Jeroen P., et al. "Higher-order multi-resolution topology optimization using the finite cell method." *International Journal for Numerical Methods in Engineering* 110.10 (2017): 903–920.
12. Nguyen, Tam H., Chau H. Le, and Jerome F. Hajjar. "Topology optimization using the p-version of the finite element method." *Structural and Multidisciplinary Optimization* 56 (2017): 571–586.
13. Yoo, Jaeun, In Gwun Jang, and Ikjin Lee. "Multi-resolution topology optimization using adaptive isosurface variable grouping (MTOP-aIVG) for enhanced computational efficiency." *Structural and Multidisciplinary Optimization* 63 (2021): 1743–1766.
14. Davies, Alex, et al. "Advancing mathematics by guiding human intuition with AI." *Nature* 600.7887 (2021): 70–74.
15. Silver, David, et al. "Mastering the game of go without human knowledge." *nature* 550.7676 (2017): 354–359.
16. Tunyasuvunakool, Kathryn, et al. "Highly accurate protein structure prediction for the human proteome." *Nature* 596.7873 (2021): 590–596.
17. Yu, Yonggyun, et al. "Deep learning for determining a near-optimal topological design without any iteration." *Structural and Multidisciplinary Optimization* 59.3 (2019): 787–799.
18. Kallioras, Nikos Ath, Georgios Kazakis, and Nikos D. Lagaros. "Accelerated topology optimization by means of deep learning." *Structural and Multidisciplinary Optimization* 62.3 (2020): 1185–1212.
19. Chandrasekhar, Aaditya, and Krishnan Suresh. "TOuNN: Topology optimization using neural networks." *Structural and Multidisciplinary Optimization* 63 (2021): 1135–1149.
20. Deng, Hao, and Albert C. To. "A parametric level set method for topology optimization based on deep neural network." *Journal of Mechanical Design* 143.9 (2021): 091702
21. Zhang, Zeyu, et al. "TONR: An exploration for a novel way combining neural network with topology optimization." *Computer Methods in Applied Mechanics and Engineering* 386 (2021): 114083.
22. Hamza, Karim, Mohamed Aly, and Hesham Hegazi. "A kriging-interpolated level-set approach for structural topology optimization." *Journal of Mechanical Design* 136.1 (2014): 011008.
23. Guirguis, David, William W. Melek, and Mohamed F. Aly. "High-resolution non-gradient topology optimization." *Journal of Computational Physics* 372 (2018): 107–125.
24. Guirguis, David, and Mohamed F. Aly. "A derivative-free level-set method for topology optimization." *Finite Elements in Analysis and Design* 120 (2016): 41–56.
25. Duchi, John, Elad Hazan, and Yoram Singer. "Adaptive subgradient methods for online learning and stochastic optimization." *Journal of machine learning research* 12.7 (2011).
26. Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980* (2014).



Study on How 3D Printer Process Parameters Can Be Optimized for Use in Orthopedics Fracture Classification and Surgical Planning

Yosef W. Adugna¹ , Heinrich Brüggemann², Erik Hafnor², Ane Djuv^{2,3}, Aksel Paulsen^{2,4}, Hanne R. Hagland^{5,6}, and Hirpa G. Lemu¹ 

- ¹ FaDepartment of Mechanical and Structural Engineering and Materials Science, University of Stavanger UiS, Stavanger, Norway
yosef.w.adugna@uis.no
- ² Stavanger University Hospital, Stavanger, Norway
- ³ Department of Clinical Medicine, University of Bergen, Bergen, Norway
- ⁴ Department of Public Health, Faculty of Health Science, University of Stavanger UiS, Stavanger, Norway
- ⁵ Stavanger University Hospital SUS, Gastrointestinal Translational Research Unit, Stavanger, Norway
- ⁶ Department of Chemistry, Bioscience and Environmental Engineering, University of Stavanger UiS, Stavanger, Norway

Abstract. The revolutionary advance of 3D printing technology is currently being tested in many medical applications due to its potential for unique patient-specific treatments. In orthopedics, 3D printing of fracture models can be used to improve the accuracy of fracture identification and classification, as well as for preoperative planning and training. For 3D models to be useful in traumatology, it is necessary that they be available in the limited preoperative time from patient admission to surgery. Our study is based on the proximal ‘hip fractures’ from the Stavanger University Hospital case study. Treatment of this type of fracture occurs within 24 h. Various print parameters in the post-processing software, including layer height, print time, and material consumption, are examined, and evaluated to ensure the optimal balance between 3D print time and print quality. Wrapping and removal of unaffected parts and the reduction of the 3D model size greatly minimize the printing time. Reducing excess information and connecting cylinders decreases print time significantly while still giving enough detail without losing fracture morphology. Other parameters including whether the use of printers with multiple extruders leads to a reduction in printing time, since the distance and size of support structures significantly affect the printing time, needs to be investigated in further studies.

Keywords: Orthopedic · Fracture classification · 3D printing parameter optimization

1 Introduction

The healthcare industry is only one of several industries that are being transformed by the development of 3D printing technology [1]. 3D printing technology has been increasingly used in the medical field in recent years, particularly for preoperative surgical planning [2], prosthetics [3], and organ replacement [4]. One of the main benefits of 3D printing in the medical area is the ability to produce individualized and customized medical devices patient specific implant (PSI) and implants [5, 6]. In orthopedics, 3D Printing of fracture models may improve the accuracy of fracture identification and classification compared to X-ray images for the surgeons, which could further assist in pre-surgical planning and teaching. Furthermore, it may reduce interobserver variability when using a 3D model to classify fractures compared to regular computerized tomography images. For 3D models to be of use in surgical planning in traumatology the time from the patient enters the emergency department to an accurate 3D model printed is limited depending on the fracture location. Here, we use proximal “hip” fractures at Stavanger University Hospital as “case studies”, (REK 2019/470). Thus, among the processes involved from patients entering the emergency department to surgery, (i.e., Digital Imaging and Communications in Medicine (DICOM) and corresponding segmentation processing), 3D Printing time occupies the highest time percentage. Even though there are many parameters to be optimized for general medical treatment when introducing 3D Printing, this work focuses on optimizing the parts to be printed through the removal of the section along with 3D Printer process parameters to reduce the printing time.

Hip fractures (Fig. 1) are common injuries with exponentially increased incidents in older adults. Hip fractures can cause serious disability, worse quality of life, and increased mortality (25% 1-year mortality) [7]. For effective and proper treatment planning, accurate classification of hip fractures is crucial, and 3D printing technology has become a valuable tool in this process. 3D printing can also help identify any unusual fracture patterns or complex fractures that may be difficult to classify using traditional methods. Based on imaging data, such as computerized tomography (CT) scans, 3D printing can produce precise models of hip fractures [8–10]. These models help surgeons understand the fracture pattern better depending on the fracture type, the best and most effective surgical technique can be chosen [10]. Moreover, 3D printed models can be used to inform patients and their families about the fracture and anticipated surgical process, which can reduce anxiety and enhance patient outcomes [11]. Additionally, surgical guides can be produced using 3D printing to help ensure the precise placement of implants and other hardware during surgery [12–14]. Being able to produce accurate fracture models using 3D printing, which are made available both within a time constraint and economically feasible, is important. A good model case to test the feasibility of 3D printing as an extra resource for surgical planning in an emergency setting is using hip fractures, as patients ideally should be on the operating table within 24 h of arriving at the emergency room [15, 16]. Thus, the 3D printed model of the patient’s hip fracture needs to be made readily available for the surgeons whilst they are planning the corrective procedure. Therefore, having a fast printing of the fracture while maintaining the anatomical accuracy of the part is necessary.

To do these different ways of optimizing techniques on the parameters of the parts as well as on the printers can be done. Among them, is the optimization of the fracture

itself, where the focus will be made only on the fracture area. This can be done during the segmentation process where unnecessary parts which are not involved in the fracture or not necessary for understanding the anatomy such as the femur head will be cut out and removed, segmentation of the bone shaft close to the fracture area as much as possible and wrapping the bone shaft to reduce surface roughness (fracture-free) areas. Removing unnecessary parts of a hip bone model can help reduce printing time in cases of hip fracture by simplifying the 3D model and reducing its overall complexity. This can be done by segmenting the fracture region and isolating it from the surrounding unaffected bone. The segmentation process involves using specialized software to identify and separate the fractured region from the rest of the bone based on information from medical imaging scans such as CT or magnetic resonance imaging (MRI).

The second technique will focus on 3D printer settings and parameters optimization, including identification of proper printing layer height for faster printing, selection of nozzle diameter, and material consumption. In this study, fused deposition modeling (FDM) printing technology is used Fig. 2. This technology used is a type of additive manufacturing technology that involves the layering of a three-dimensional object using thermoplastic material extruded through a nozzle [17]. A print bed, a nozzle, and a filament feed mechanism are typical components of an FDM 3D printer. The filament feed mechanism feeds the thermoplastic filament into the printer's nozzle, which melts the filament and extrudes it onto the print bed in the desired shape. The printer's print bed moves up or down to build up the object layer by layer, and the extruded material solidifies as it cools, creating a solid object. FDM based 3D printers are commonly used for quick prototyping, model production, and the fabrication of functional parts. They are also commonly used in the medical field of orthopedics to create patient-specific implants, prostheses, and surgical guides [18]. FDM-based 3D printers are popular because of their versatility, low cost, and ease of use [19].



Fig. 1. Hip fracture

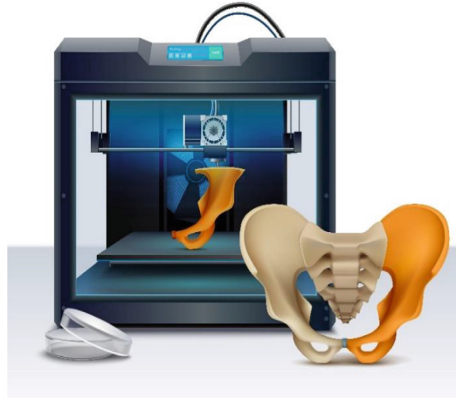


Fig. 2. FDM 3D printer

This article will assess the parameters of the FDM printer affecting the efficiency of the printed part based on the patient's hip fracture also considering 3 different cases categorized in the segmentation process (part optimization). The highest severe (Complex) case is considered for this work.

2 Materials and Methods

Digital Imaging and Communications in Medicine (DICOM) image of the patient from the CT scan is imported to Materialise Mimics v24 and the fracture is segmented. Thus, imported to 3-Matic software to convert it to a 3D Model. The Severe case (9) based on the Hospital's AO classification system of fracture is considered. Severe cases are expected to have longer 3D printing time than other fracture conditions as more fragments of the bone require more connecting rods. This condition directly affects the duration of 3D printing time.

In the first part, 3 different segmentation methods were considered on the 3D model to reduce 3D printing time:

- i. Normal condition (with head and not wrapped)
- ii. Head cut only (not wrapped)
- iii. Head cut and wrapping.

In the second part, different printer settings and parameters were assessed.

2.1 Fracture 3D Model Part Optimization

Three different 3D models (*.stl files*) according to the description of cases were prepared and 3D printed. To keep consistency, all models in the three cases were printed using the same 3D printing parameters. The printing parameter used was the 0.2-layer thickness, 10% infill density, and triangle infill pattern, whereas the printing speed was 70 mm/s with printers using 0.4 mm nozzle diameter. Cases are described below:

Case I - Normal condition (with head)

Figure 3 illustrates the normal segmentation condition considered as case 1. This case implies the full fracture site including the head of the femur where there is no fracture as well as no removal parts. One of the difficulties in orthopedic surgery is dealing with fracture segments that have detached from the bone. When these pieces are not appropriately stabilized and maintained together, they can lead to poor healing and possibly re-fracture. One way to fix this is to use segment cylinders, which are custom-designed devices that are inserted into the bone to hold the split bone segments together. The main location of the fracture, the fracture segments, and the highly porous bone areas found in the femur are shown in the below figure.

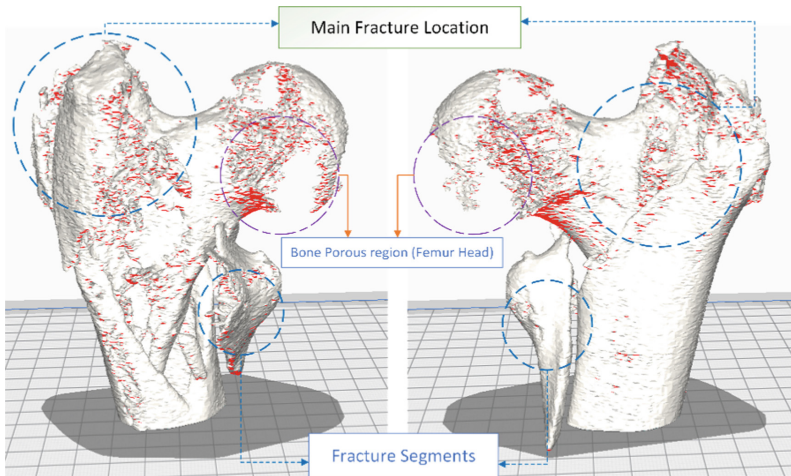


Fig. 3. Case I – Normal condition

Case II - Head cut (Removal)

The second option that was suggested is to cut the head of the femur as seen below in Fig. 4. Typically, the femoral neck or intertrochanteric portion of the femur is the location of the primary fracture in a hip fracture [18]. The femoral neck is the portion of the femur bone that joins the bone's ball-shaped head to the main shaft, while the intertrochanteric region is located right below the femoral neck. These regions are essential for the stability and movement of the hip joint, and a fracture in either of them can have a profound effect on a patient's quality of life.

Case III - Wrapping of no fracture area (shaft and calcar) and femur head removal

The final case, as shown in Fig. 5 is when the wrapping of the bone shaft, was introduced in addition to the head cut. Wrapping the fractured bone involves isolating the fracture region and removing the unaffected parts of the bone from the 3D model. This can help to simplify the 3D model and reduce its complexity, which can potentially reduce the printing time. However, this depends on the specific case and the size and location of the

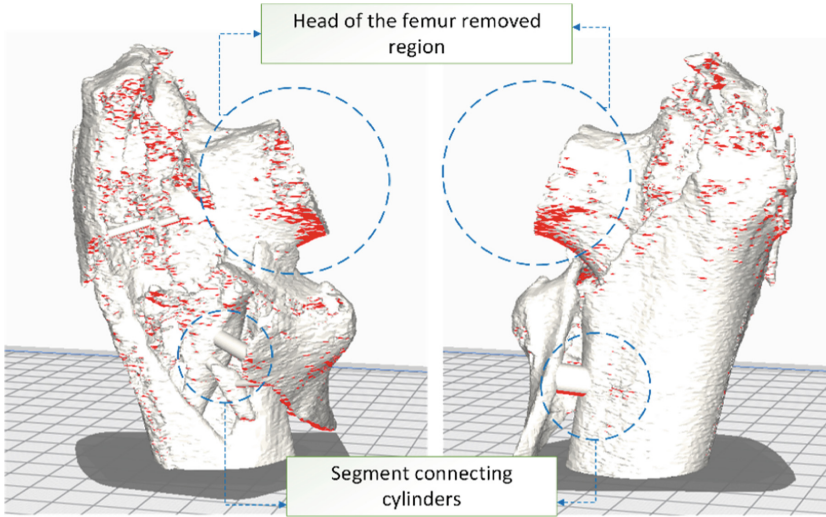


Fig. 4. Hip fracture after the removal of the femur head

fracture. Wrapping is done in Materialise Mimics segmentation software. The nature of the bone surface is rough which makes 3D printers stay longer to print the edges, it is assumed by smoothing the surface the retraction time of the 3D printer movement will decrease and that makes overall 3D printing timeless.

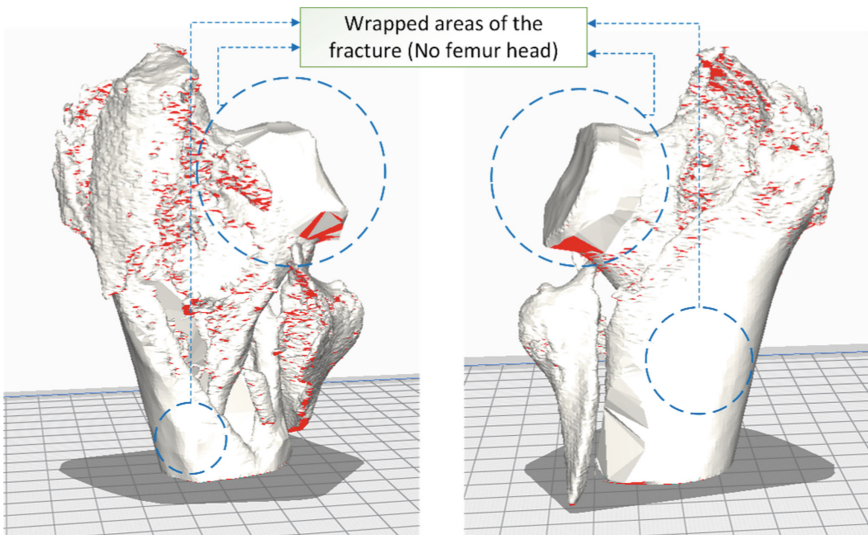


Fig. 5. Wrapped areas of the femur fracture-free zones

2.2 Study of the FDM 3D Printer Parameters

The experimental 3D printer used in this article is the Ultimaker S5. This is considered for the efficiency of the printer as well as the low cost of the 3D printer in general. The material selected for model printing is PLA, which is affordable and eco-friendly, and PVA for the support structure. Ultimaker's Cura v5 slicer was used to check the parameters. Tested parameters are:

- i. Print time
- ii. Layer height
- iii. Material consumption
- iv. Nozzle diameter (0.4 mm and 0.8 mm).

The fixed printing parameters are given in Table 1.

Table 1. Printing parameters for both 0.4 mm diameter nozzle

Print parameter	Description
Printer and nozzle	Ultimaker 5, 0.4 mm
Printing temperature	205
Print speed	70 mm/s
Infill pattern	Triangles
Infill density	10%
Wall thickness	mm

3 Results and Discussion

3.1 Fracture 3D Model Part Optimization

By using the same printing parameters for a 0.4 mm nozzle and 0.2 mm layer height, the printing time for the 3D model after part optimization for the 3 cases is recorded from slicer software. Among the 3 cases, it is found that removing the head as well as wrapping reduces the time. This is directly related to the reduction of the file size (Table 2).

Removal of unaffected parts of the bone during fracture like cutting the femur head in this case, can greatly help to reduce the printing time. This is because the printing process will go more quickly because there will be less material to build and fewer details to print. The 3D model created when a fractured bone is segmented from a medical image has rough edges or surface imperfections due to the nature of the bone that needs to be smoothed out before printing. Also, this involves additional steps in the 3D modeling preparation for models' geometry refinement. Wrapping helps simplify the 3D model and reduce complexity, which reduces the printing time. However, this depends on the specific case and the size and exact location of the fracture.

Table 2. Time recorded for the 3 different cases

	Description	Time
Case 1	Normal (with femur head)	11.7 h
Case 2	Head cut (no wrapping)	8 h
Case 3	Head cutted and wrapped	7.6 h

3.2 Study of the FDM 3D Printer Parameters

The layer height can be configured to be no more than 75% of the nozzle diameter for any specific print core or nozzle combination. Considering this, different layer height is given and the printing time for each case is recorded. The parameters are discussed below.

i. *Layer height vs print time (0.4 mm nozzle diameter).*

Recorded data of the relation between layer height and print time for 0.4 mm is shown in Table 3.

Table 3. The relation between layer height and print time

Layer height (mm)	0.04	0.06	0.08	0.1	0.15	0.2	0.25	0.3	0.35	0.4
Print time (hrs)	34	24.5	19	13.3	9.83	7.6	6.5	5.82	5.2	4.5

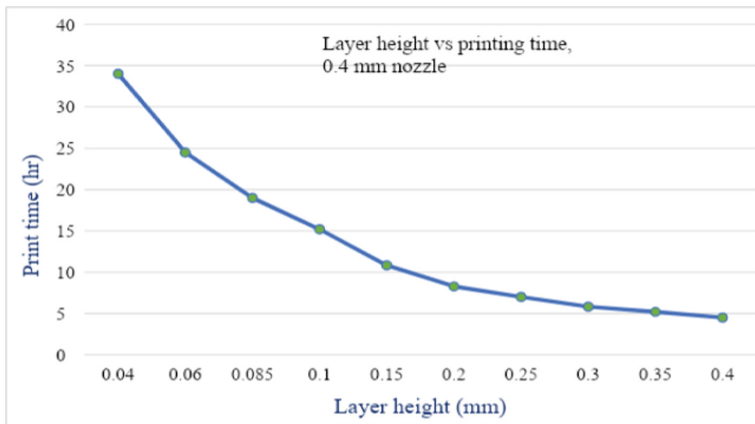


Fig. 6. Layer height (mm) and printing time (hrs)

The above experimental data (Fig. 6) result shows that there exists an increment in the height of the layer corresponding to a decrement in the printing time. Although the printing time and speed decrease with increasing layer thickness, the resulting model is rougher. This is because when the layer height drops, the printer head must make more passes to complete each layer. This indicates that printing the same thing takes longer than printing it with a thicker layer height. Furthermore, printing with thinner layers necessitates a higher level of precision and accuracy, which can extend printing time. The rate of change is increasing gradually.

ii. *Layer height vs material consumption.*

The relation between the layer height and the material is shown in Fig. 7. Table 4 also illustrates the relation between the print time and layer height recorded.

Table 4. The relation between layer height and print time

Layer height (mm)	0.04	0.06	0.085	0.1	0.15	0.2	0.25	0.3	0.35	0.4
Material consumption (m)	9.6	9.66	9.93	9.78	10.05	10.1	10.42	10.35	10.7	10.9

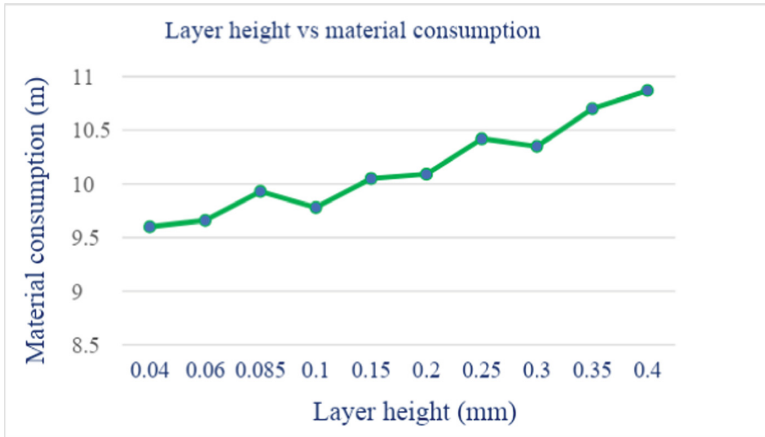


Fig. 7. Layer height (mm) and material consumption (m)

Figure 8 shows the experimental data of print layer height, printing time, and consumables, and the relationship between layer height and print time, and consumables. The figure illustrates the evolution of the relationship between layer height, printing time, consumables, and precision over time. According to the figure presented, the results of the experiments indicate that the printing time is reduced to its minimum while still maintaining a high level of print quality when the layer height is set between 0.1 and

0.15 mm. The 3D Printing time between 0.1 mm and 0.15 mm for the complex (severe) case selected is between 13 h and 9.8 h, whereas the consumable material is almost similar for the cases.

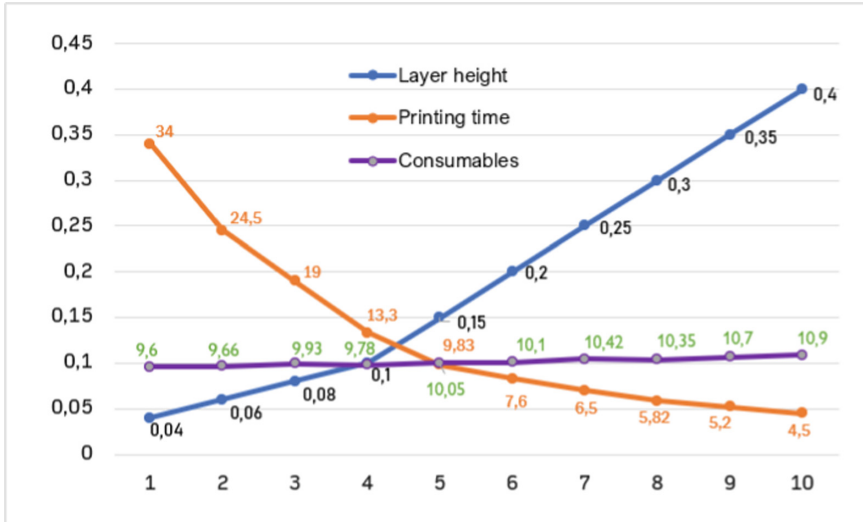


Fig. 8. Layer height, printing time, and consumables comparative illustration

3.3 Connecting Cylinders

Fracture fragments connecting cylinders (Fig. 9) can lengthen the printing process for a hip fracture model. The addition of these connecting cylinders to the fracture fragments in the 3D model ensures that the fracture fragments are kept in the proper places and simulate the true bone structure. However, including these connecting cylinders may complicate the 3D model, lengthening the print process. This is because a model with connected cylinders can require a longer print time than a simpler model because connecting cylinders require the printer to print more precise details, such as the shape and location of the cylinders. The number and size of fracture fragments in the model can also have an impact on print time. The model will demand more material to print and take longer to manufacture the larger the fracture fragments. Similarly, to this, the model will be more complex and take longer to print more fracture fragments it has.

Using fracture fragments to connect cylinders can make the printed model stronger and more stable, but it's important to consider how this will affect the 3D print time and cost. Depending on the needs of the hip fracture treatment, it may be better to make a simpler model without the cylinders in some situations.

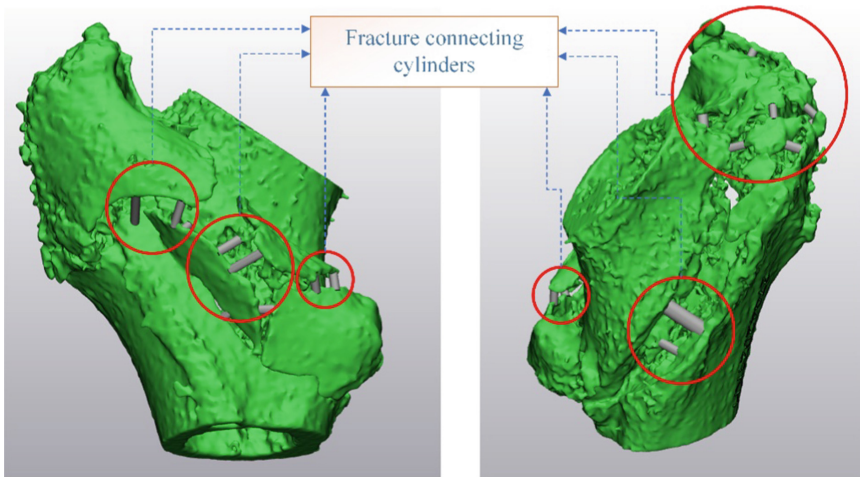


Fig. 9. Fracture connecting cylinders

4 Conclusion

It is seen that optimizing part itself plays an important role in a better 3D printing time, especially in a condition where there is a complex geometry like a hip fracture. It's important to note that reducing print time may also affect the quality and accuracy of the final print. Thus, finding the right balance between the print time and print quality to ensure the best possible outcomes for patients.

Wrapping and removal of unaffected parts greatly reduced the printing time. Reducing the size of the 3D model by removing the unaffected areas of the bone can help to speed up the printing process. This is because there will be less overall information to print and less specific information to print. Wrapping the fractured bone helps to reduce the complexity of the 3D model and makes it easier to print. However, wrapping the bone can also reduce the accuracy of the model and may not be suitable in cases where the unaffected parts of the bone are crucial for treatment planning.

One method for reducing the 3D print time for models of hip fractures is to simplify the model by lowering the number and size of fracture fragments as well as the quantity of connecting cylinders. To accomplish this, it demands only adding connecting cylinders to the most important fracture fragments and leaving the less important fragments without connecting cylinders.

References

1. Schubert, C., M.C. Van Langeveld, and L.A. Donoso, Innovations in 3D printing: a 3D overview from optics to organs. *British Journal of Ophthalmology* 98(2), 159–161 (2014).
2. Tejo-Otero, A., I. Buj-Corral, and F. Fenollosa-Artés, 3D printing in medicine for preoperative surgical planning: a review. *Annals of biomedical engineering* 48(2), 536–555 (2020).
3. Ventola, C.L., Medical applications for 3D printing: current and projected uses. *Pharmacy and Therapeutics* 39(10), 704 (2014).

4. Arslan-Yildiz, A., et al., Towards artificial tissue models: past, present, and future of 3D bioprinting. *Biofabrication* 8(1), 014103 (2016).
5. Wixted, C.M., et al., Three-dimensional printing in orthopaedic surgery: current applications and future developments. *JAAOS Global Research & Reviews* 5(4), (2021).
6. Han, Q., et al., Novel exploration of 3D printed wrist arthroplasty to solve the severe and complicated bone defect of wrist. *Rapid Prototyping Journal*, (2017).
7. Zanker, J. and G. Duque, Osteoporosis in older persons: old and new players. *Journal of the American Geriatrics Society* 67(4), 831–840 (2019).
8. Maryada, V.R., et al., pre-operative planning and templating with 3-D printed models for complex primary and revision total hip arthroplasty. *Journal of Orthopaedics* 34, 240–245 (2022).
9. Sallent, A., et al., Feasibility of 3D-printed models of the proximal femur to real bone: a cadaveric study. *HIP International* 29(4), 452–455 (2019).
10. Lim, P.K., et al., Use of 3D printed models in resident education for the classification of acetabulum fractures. *Journal of surgical education* 75(6), 1679–1684 (2018).
11. Bizzotto, N., et al., Three-dimensional printing of bone fractures: a new tangible realistic way for preoperative planning and education. *Surgical Innovation* 22(5), 548–551 (2015).
12. Kim, T., et al., Accuracy of a simplified 3D-printed implant surgical guide. *The Journal of prosthetic dentistry* 124(2), 195–201 (2020).
13. Henckel, J., et al., 3D-printed patient-specific guides for hip arthroplasty. *JAAOS-Journal of the American Academy of Orthopaedic Surgeons* 26(16), 342–348 (2018).
14. Bizzotto, N., et al., 3D printed models of distal radius fractures. *Injury* 47(4): 976–978 (2016).
15. Kjaervik, C., et al., Waiting time for hip fracture surgery: hospital variation, causes, and effects on postoperative mortality: data on 37,708 operations reported to the Norwegian Hip fracture Register from 2014 to 2018. *Bone & joint open* 2(9), 710–720 (2021).
16. Pincus, D., et al., Association between wait time and 30-day mortality in adults undergoing hip fracture surgery. *Jama*, 2017. 318(20): p. 1994–2003 (2017).
17. Wu, J. Study on optimization of 3D printing parameters. in *IOP conference series: materials science and engineering*. IOP Publishing (2018).
18. Sontakke, U. and S. Jaju, The Role of 3D Printing in the Biomedical Application: A Review. *Smart Technologies for Energy, Environment and Sustainable Development, Vol 2: Select Proceedings of ICSTEESD*: 371–381 (2020).
19. Hoy, M.B., 3D printing: making things at the library. *Medical reference services quarterly* 32(1), 93–99 (2013).
20. Woo, S.-H., et al., Three-dimensional-printing technology in hip and pelvic surgery: current landscape. *Hip & pelvis* 32(1),1–10 (2020).



Metaheuristic Optimization of Compressive Strength and Microstructures for Fused Deposition Modelling Using PLA Material

Kedir Hussen¹✉, Hirpa G. Lemu², and Amanuel Diriba Tura¹

¹ Jimma Institute of Technology, Jimma University, Jimma, Ethiopia
kedir40hussen@gmail.com

² Faculty of Science and Technology, University of Stavanger, Stavanger, Norway

Abstract. As a result of the rapidly advancing additive manufacturing technologies, parallel growth of interest in additively fabricated complex parts is observed. Fabricating parts using additive technology opens opportunities to produce parts with complex geometries and avoids manufacturing constraints. This provides full design freedom that can lead to design optimization involving several design or performance constraints. Nowadays, there are different optimization tools used in optimization of parts and materials with different capabilities. In this paper, metaheuristic optimization of fused deposition modelling process parameters was conducted to improve the compressive strength and microstructures of polylactic acid (PLA) materials. The genetic algorithm optimization tool was implemented, and analysis of variance (ANOVA) was used for the analysis of mechanical properties. In the compression strength, the ANOVA results indicated that the platform temperature is the most significant parameter for compression strength for fabrication of the material followed by print speed, number of shells, infill density, extrusion temperature and layer thickness. It has been also observed that the microstructure is highly influenced by layer thickness followed by print speed, number of shells, extrusion temperature, platform temperature and infill density.

Keywords: Additive manufacturing · Design optimization · Fused deposition modelling · Genetic algorithm · Metaheuristic optimization · Polylactic acid materials

1 Introduction

Additive manufacturing (AM) is the deposition of material layer by layer to create items [1]. 3D printing (3DP), rapid prototyping (RP), direct digital manufacturing (DDM), rapid manufacturing (RM), and solid freeform fabrication (SFF) are all acronyms used to describe AM technologies [2]. FDM is an extrusion-based AM method in which material is melted and dispensed selectively through a fine nozzle according to a slice model [3].

There are different types of 3D printing technologies, from those material extrusion, material jetting, powder bed fusion, vat photopolymerization, sheet lamination, binder

jetting, and directed energy deposition (DED) are the most common. This technology has also many applications area. For instance, construction, biomedical, academia, fashion industry, food processing, aerospace, etc. are the most common application areas of 3D printing technology.

Some researchers have done regarding to the additive manufacturing technology. Their researches are used to add some values on the specimen of different materials, even if there are many gaps.

From these researches for instance Wang et al. [4] investigated the impact of several printing parameters on the mechanical characteristics, microstructure, and surface quality of printed PEEK parts, including printing temperature, printing speed, and printing layer thickness. Naveed [5] researched the properties of five different raster angles used to build 3D components utilizing thermoplastic material–polylactic acid (PLA), in order to determine the ideal raster position for fabricating the strongest good microstructure 3D printing item. The study also recommended for further study by adjusting printing conditions. Beside this Mohanty et al. [6] have also declared the effect of process parameters on dimension precision and microstructures of FDM printed part where DOE is designed by using Taguchi-MACROS-nature inspired heuristic optimization technique. They use the specimen fabricated by using ABS M30 (Acrylonitrile Butadiene Styrene). On the other hand Anghel et al. [7] by taking process parameters of layer thickness, infill density, and imposed clearance between the shaft and the hole. The Genetic Algorithm Artificial Neural Network (GA-ANN) is utilized for output optimization of part quality and functionality of the part. The process parameters were set with three distinct levels (low, medium and high). Finally, the study got as the layer thickness has a great impact on the part quality and recommended on other parameters for further study.

Some researchers have done also on FDM process parameters in order to enhance mechanical properties of the specimen. Here are some literatures and their recommendation which deal about adding value on the properties by optimizing the process parameters. From these literatures Huu et al. [8] optimized the process parameters for the improvement of compressive strength for the products fabricated using PLA-copper. The research focused on the process parameters of raster angle, build direction, and layer thickness. The paper has used central composite design in the design of experiments and ANOVA for the data analysis and recommended on other parameters for further study. Whereas Dey et al. [9] have did a study on how process parameters are optimized in fused deposition modeling with the particle swarm optimization algorithm method. Here PLA is used as the filament material since it is non-toxic and biodegradable. In the study, four process parameters layer thickness, build orientation, infill density, and extrusion temperature is optimized to get higher compressive strength and lower build time. On the other hand optimization of the fused deposition modeling process parameters through PLA implant is done by using adaptive neuro-fuzzy inference (ANFIS) model and whale optimization algorithm by Sai et al. [10] which deals about the effect and control of FDM process parameters (layer thickness, raster angle, infill density, and internal structure) on the output of build time, compressive strength, and surface roughness of the specimen. And finally the study recommended on other optimization methods for further study When engaging in post-processing activities for 3DP components, Impens and Urbanic [11] conducted a comprehensive study to understand the green component mechanical

characteristics (compressive strength) and the various elements that can influence the mechanical properties. To do an analysis of variance (ANOVA) on the results, a full factorial design of experiments (DOE) approach was used. On the other hand Yadav et al. [12] investigated the impact of different infill design patterns on the compressive strength of FDM-processed PLA components. They compared compressive strengths for six distinct infill designs recommended on other process parameters and output parameters for further study.

Xia et al. [13] investigated the effects of binder saturation level on linear dimensional accuracy and compressive strength of 3D printed geopolymer. According to the findings, increasing the binder saturation level significantly boosted the compressive strengths of both green and post-processed samples. Oudah et al. [14] looked studied how infill pattern (Tri-Hexagon, Zig-Zag, and Gyroid), infill density (25, 50, and 75) percent, and layer thickness (0.1, 0.2, and 0.3) mm affected the compressive strength of ABS materials. According to the result the infill density is the most sensitive characteristic among the three parameters evaluated, according to the research findings and analysis of variance results. Patibandla and Mian [15] Mian investigated the effect of process parameters such as extrusion temperature and print speed on sample mechanical properties. Mechanical parameters such as failure strength, yield strength, and compressive modulus were evaluated by compression testing. The print speed has a minor impact on mechanical qualities, although greater fabrication temperatures improve mechanical properties.

When engaging in post-processing activities for 3DP components, Impens and Urbanic [11] conducted a comprehensive study to understand the green component mechanical characteristics (compressive strength) and the various elements that can influence the mechanical properties. To do an analysis of variance (ANOVA) on the results, a full factorial design of experiments (DOE) approach is used. Hernandez et al. [16] published the results of compression (ASTM standard D695), 4-point flexure (ASTM D790), and tensile (ASTM D 638 Type I) testing on ABS plastic specimens specified to ASTM standards and manufactured in various orientations using the uPrint SE Plus 3D printer. The impact of printing 3D items in various orientations on the mechanical characteristics of ABS P430 plastic was also investigated. When compared to other orientations, printing with 0° in the x-y plane boosted the material's compressive and flexure strengths dramatically. Patibandla and Mian [15] investigated the effect of process parameters such as extrusion temperature and print speed on sample mechanical properties. Meta-structure specimens were created on a MakerBot 2X Replicator 3D printer using an acrylonitrile butadiene styrene polymer. Mechanical parameters such as failure strength, yield strength, and compressive modulus were evaluated by compression testing. The print speed has a minor impact on mechanical qualities, although greater fabrication temperatures improve mechanical properties.

2 Materials and Methodology

2.1 PLA Material and 3D Printer

In the FDM process, PLA, ABS, PETG, and some other plastic filament materials can be used as a building material. In this thesis, PLA filament plastic material is used as a building material due to its behavior of in-toxicity, availability on the market, and property of low melting point. The building material properties are revealed in Table 1.

Table 1. Properties of PLA filament.

Property	Value
Name	Polylactic acid (PLA)
Chemical formula	C ₃ H ₄ O ₂
Melting temperature	157–170 °C
Typical temperature of molding	178–240 °C
Density	1.24 g/ [cm] ^3

The additive manufacturing technology is very advanced technology to fabricate a product which has complex shape. There are different types of machines under this technology based on its use of range (specification limit) and working principle. For instance, FDM machine, LOM machine, SLA machine, SLS machine, etc. In this work an FDM machine called Flashforge Guider II is used since it is available here. This machine has basically model material spool, build platform, liquefier nozzle, extruder head, etc. The build platform is free for moving up and down, whereas the extruder can move towards anywhere on the foam base horizontally. The filament (PLA) is fed through the extruder head automatically and it is melted then deposited through the nozzle to fabricate the body layer by layer. The Flashforge guider II machine is shown in Fig. 1 and its specification are presented in Table 2.

2.2 Process Parameters

A research work can't be undertaken without the use of sampling, because studying the whole population is not possible due to some practical limitations like time, cost and other factors. In the process of 3D printing, there are a population of process parameters which affect the output parameters even though their degree of influence on the output is different. In this work, stratified sampling method based on the initial trials of exhaustive literature study are used to select the process parameters. These process parameters are stratified in to groups regarding outputs and take one process parameter at a time based on their degree of influence on the outputs.

The level selection of each process parameter is according to the permissible standard slice profile setting recommendation of the machine which is recommended by the machine specifier, experience and real application in industries. By taking the standard slice profile setting of the machine specification, the sample levels are taken by going up and down proportionately then the level is leveled as five levels as shown in Table 3.

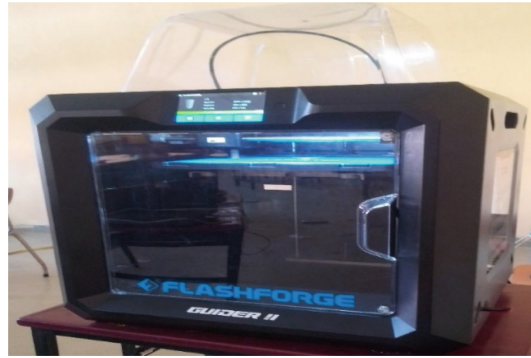


Fig. 1. Flashforge guider II 3D printer machine

Table 2. Specification of flashforge guider II 3D printer machine.

Name	Guider II
Number of extruders	1
Print technology	FDM
Build volume (mm ³)	280*250*30
Layer resolution (mm)	0.01–0.4
Build accuracy (mm)	±0.2
Position accuracy (mm)	0.0025 in z axis; 0.011 in X and Y axis
Filament diameter (mm)	1.75 ± 0.07
Nozzle diameter (mm)	0.4
Print speed (mm/s)	10–200
Software used	FlashPrint
Support format	STL/3MF/JPG
Device size (mm)	490*550*560
Net weight (Kg)	30
AC Input	Input; 100V-240 VAC, 47–63 Hz and power 500 W
Connectivity	Wifi, Ethernet, USB cable and USB stick

2.3 Design of Experiment

Design of experiment (DOE) is used to regulate and allow to observe the relation or interaction between the factors and the results. It is more important to examine the relationship between various input variables (aka factors) and important output variables in a systematic and efficient manner (aka responses). In this work DOE is also more important to set the interaction between all process parameters (Layer thickness, print speed, Infill density, Extrusion temperature, Number of shells, Platform temperature) by

Table 3. Process parameters and their levels.

S. no	Parameters	Level 1	Level 2	Level 3	Level 4	Level 5
1	Layer thickness (mm)	0.14	0.16	0.18	0.20	0.22
2	Print speed (mm/sec)	40	50	60	70	80
3	Infill density (%)	5	10	15	20	25
4	Extrusion temperature (°C)	200	205	210	215	220
5	Number of shells	1	2	3	4	5
6	Platform temperature (°C)	20	25	30	35	40

taking in to consideration their levels and the response parameters (compressive strength and microstructure). In this paper, Taguchi L25(5⁶) orthogonal array design of experiment in MINITAB software was used to analyze the interaction between factors and the responses as given in Table 4. Then finally parts are produced for each experimental run ratio after slicing with the help of a software called FlashPrint 5.

2.4 Production of Specimen

The CAD models of the workpiece are modeled with the help of CATIA V5 then it is directly converted to an STL file. Then the parts are produced after slicing by using Flash-Print 5 software. In the process of CAD model design, each part is designed according to the design standards for the experiment testing interchangeability. For the production of the specimens, the process parameter variables are fed according to the experiment plan which is already processed in the design of experiment. For the whole fabrication of parts, the PLA filament is loaded to the FDM machine nozzle head.

The compressive strength test specimen is usually fabricated in the form of a cylinder, cube, prism, and other form standards. In this work, the test specimen is fabricated in the standard of ASTM D695 [17] in cylindrical shape for saving of input material and the design standard is shown in Fig. 2. Whereas micro-structure examination was conducted on the workpiece prepared for tensile test with the standard of ASTM D638 which is the most common testing standard for the determination of the microstructure properties of materials. The part design is as shown in Fig. 2.

Table 4. Experimental runs and process parameters relation.

Experimental run	Layer thickness (mm)	Print speed (mm/sec)	Infill density (%)	Extrusion temperature (°C)	Number of shells	Platform temperature (°C)
1	0.14	40	5	200	1	20
2	0.14	50	10	205	2	25
3	0.14	60	15	210	3	30
4	0.14	70	20	215	4	35
5	0.14	80	25	220	5	40
6	0.16	40	10	210	4	40
7	0.16	50	15	215	5	20
8	0.16	60	20	220	1	25
9	0.16	70	25	200	2	30
10	0.16	80	5	205	3	35
11	0.18	40	15	220	2	35
12	0.18	50	20	200	3	40
13	0.18	60	25	205	4	20
14	0.18	70	5	210	5	25
15	0.18	80	10	215	1	30
16	0.20	40	20	205	5	30
17	0.20	50	25	210	1	35
18	0.20	60	5	215	2	40
19	0.20	70	10	220	3	20
20	0.20	80	15	200	4	25
21	0.22	40	25	215	3	25
22	0.22	50	5	220	4	30
23	0.22	60	10	200	5	35
24	0.22	70	15	205	1	40
25	0.22	80	20	210	2	20

2.5 Procedure of the Experiment

In this compression strength test the specimen was designed in the standard of ASTM D695. It was conducted using a testing machine as shown in Fig. 3. During the test the specimen is placed at the center of the two compressor discs and then the load is applied uniformly with a speed of 5 mm/min for the whole specimens. The load is increased up to the maximum stress is reached (when the part is collapsed). The result of the test consists the maximum load by which the material can withstand (kN) and maximum compression strength (MPa). Since the dimensions of the part was entered before the

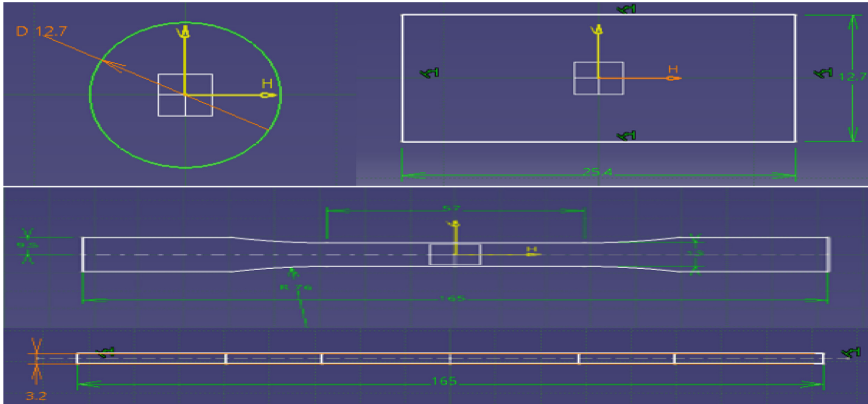


Fig. 2. Specifications of compression test and microstructure test specimens respectively.

compressive load is applied the compressive strength is automatically calculated and displayed which is as shown in Table 5.

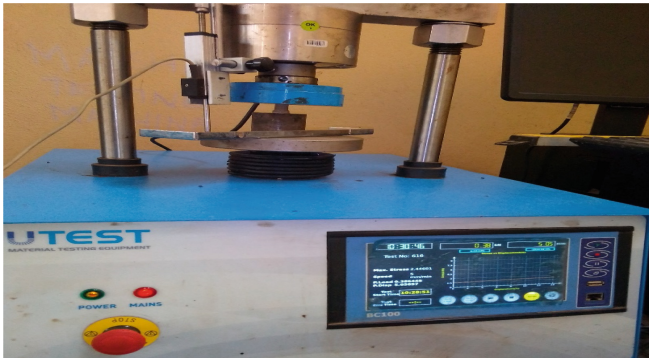


Fig. 3. Compression test machine.

Micro-structure examination is generally conducted using optical microscope. It is used for the purpose of gaining information on how the PLA material is made, the quality of the resulting PLA material and magnifying the micro-level of the detailed structure of the materials. The microstructure of the material is expressed using the size of the grain inside it, then the characteristic of the material is depended up on the grain size. The smaller the grain size the best the microstructure is. The test is conducted using optical microscope with different magnification to decide the effect of process parameters on the microstructure (grain size) as shown in Fig. 4 then the grain length is measured using ImageJ software and the result is recorded as Table 5.

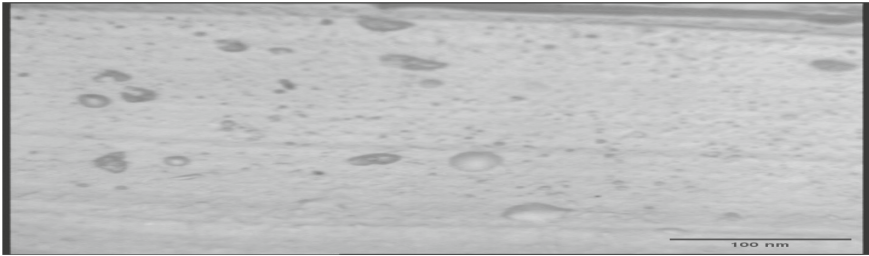


Fig. 4. Image of microstructure grains.

In this work, analysis of variance demonstrates the contribution and significance of all factors for the response. Table 6 shows the ANOVA for all parameters as which are insignificant even though their degree of insignificance is increase with layer thickness followed by extrusion temperature, number of shells, infill density, platform temperature and print speed for compression strength for the fabrication of FDM.

From the graph of analysis in Fig. 5 between layer thickness and compression strength it is observed that a part with a layer thickness of 0.2 mm has got the maximum compression strength value of 2.653 MPa and a part with a layer thickness of 0.18 mm has got the minimum compression strength value of 2.515 MPa. Beside this from the graph plotted between print speed and compression strength, it is observed a specimen with a print speed of 70 mm/sec has got the highest compression strength value of 2.760 MPa and a part with a print speed of 40 mm/sec has got the lowest compression strength value of 2.463 MPa. Whereas the graph plotted between infill density and compression strength shows that the highest compression strength of 2.716 MPa is scored from the part printed with an infill density of 5% and the lowest compression strength of 2.459 MPa is scored from the part printed with an infill density of 15%. The graph plotted between extrusion temperature and compression strength also shows the highest compression strength of 2.657 MPa is scored from the part printed with an extrusion temperature of 205 °C and the lowest compression strength of 2.522 MPa is scored from the part printed with an extrusion temperature of 210 °C. From the plotted graph between number of shells and compression strength it is observed that the part with a number of shells of 3 has got the highest compression strength of 2.706 MPa and a part with a number of shells of 1 has got the lowest compression strength of 2.465 MPa. And finally, from the plotted graph between platform temperature and compression strength it is observed that a part with a platform temperature of 25 °C has got the highest compression strength of 2.781 MPa and a part with a platform temperature of 20 °C has got the lowest compression strength of 2.426 MPa.

In this paper, analysis of variance also explains the contribution and significance of all factors for the response of microstructure. Table 7 declares the ANOVA for all process parameters are insignificant in response of microstructures for the fabrication of FDM since the P-value is above 0.05.

From the graph of analysis in Fig. 6 between layer thickness and mean microstructure observed that a sample with a layer thickness of 0.20 mm has got the best microstructure grain length of 3.723 nm and a sample with layer thickness of 0.18 mm has got the

Table 5. Experimental tests of all tests.

Experimental run	Layer thickness (mm)	Print speed (mm/sec)	Infill density (%)	Extrusion temperature (°C)	Number of shells	Platform temperature (°C)	Compression strength (MPa)	Microstructure (nm)
1	0.14	40	5	200	1	20	2.324	4.325
2	0.14	50	10	205	2	25	2.857	5.288
3	0.14	60	15	210	3	30	2.371	4.754
4	0.14	70	20	215	4	35	2.970	4.404
5	0.14	80	25	220	5	40	2.509	4.550
6	0.16	40	10	210	4	40	2.550	5.169
7	0.16	50	15	215	5	20	2.267	3.593
8	0.16	60	20	220	1	25	2.583	7.117
9	0.16	70	25	200	2	30	2.601	4.014
10	0.16	80	5	205	3	35	2.869	3.804
11	0.18	40	15	220	2	35	2.189	5.259
12	0.18	50	20	200	3	40	2.775	4.705
13	0.18	60	25	205	4	20	2.449	6.682
14	0.18	70	5	210	5	25	2.857	5.016
15	0.18	80	10	215	1	30	2.304	4.226

(continued)

Table 5. (continued)

Experimental run	Layer thickness (mm)	Print speed (mm/sec)	Infill density (%)	Extrusion temperature (°C)	Number of shells	Platform temperature (°C)	Compression strength (MPa)	Microstructure (nm)
16	0.20	40	20	205	5	30	2.451	3.068
17	0.20	50	25	210	1	35	2.454	4.773
18	0.20	60	5	215	2	40	2.836	4.097
19	0.20	70	10	220	3	20	2.714	3.127
20	0.20	80	15	200	4	25	2.810	3.552
21	0.22	40	25	215	3	25	2.799	4.017
22	0.22	50	5	220	4	30	2.693	5.241
23	0.22	60	10	200	5	35	2.423	3.735
24	0.22	70	15	205	1	40	2.658	3.985
25	0.22	80	20	210	2	20	2.376	3.669

Table 6. Analysis of variance table for compression test.

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Layer thickness	1	0.00107	0.001067	0.02	0.892
Print speed	1	0.06949	0.069490	1.22	0.283
Infill density	1	0.03011	0.030111	0.53	0.476
Extrusion temperature	1	0.00715	0.007152	0.13	0.727
Number of shells	1	0.01925	0.019247	0.34	0.568
Platform temperature	1	0.03892	0.038921	0.69	0.418
Error	18	1.02147	0.056748		
Total	24	1.18746			

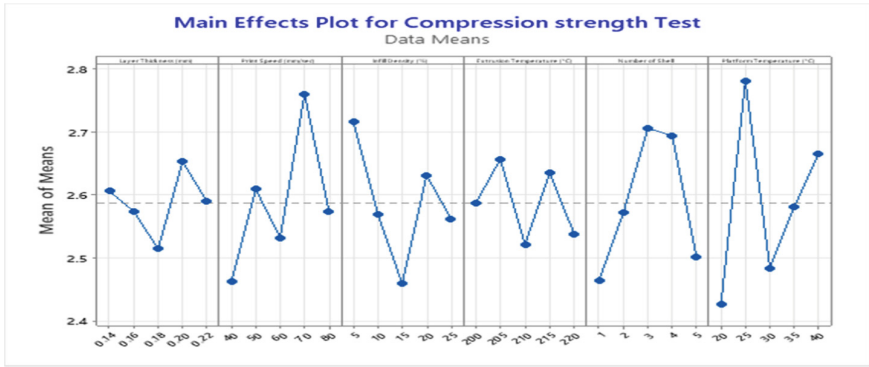


Fig. 5. Main effects plot for compression strength test.

Table 7. Analysis of variance table for microstructure.

Source	DF	Adj SS	Adj MS	F-value	P-value
Layer thickness	1	2.1749	2.17486	2.28	0.149
Print speed	1	1.0162	1.01617	1.06	0.316
Infill density	1	0.4093	0.40933	0.43	0.521
Extrusion temperature	1	1.1059	1.10588	1.16	0.296
Number of shells	1	0.7705	0.77054	0.81	0.381
Platform temperature	1	0.0126	0.01264	0.01	0.910
Error	18	17.1889	0.95494		
Total	24	22.6783			

worst microstructure grain length of 5.178 nm. And also, from the graph of analysis between print speed and mean microstructure it is obtained that a part with a print speed of 80 mm/sec has got a best microstructure grain length of 3.960 nm and a part with a print speed of 60 mm/sec has got the worst microstructure grain length of 5.277 nm. At the same time from the graph of analysis between infill density and mean microstructure it is found that a sample with an infill density of 15% has got the best microstructure grain length of 4.229 nm and a sample with an infill density of 25% has got the worst microstructure grain length of 4.593 nm. From the graph of analysis between extrusion temperature and mean microstructure it is observed that a part with an extrusion temperature of 200 °C has got the best microstructure grain length of 4.066 nm and a part with an extrusion temperature of 220 °C has got the worst microstructure grain length of 5.059 nm. And from the graph of analysis between number of shells and mean microstructure it is found that a sample with a number of shells of 5 has got the best microstructure grain length of 3.992 nm and a sample with a number of shells of 4 has got the worst microstructure grain length of 5.010 nm. And finally, from the graph of analysis between platform temperature and mean microstructure it is observed that a part with a platform temperature of 30 °C has got the best microstructure grain length of 4.261 nm and a part with a platform temperature of 25 °C has the worst microstructure grain length of 4.998 nm.

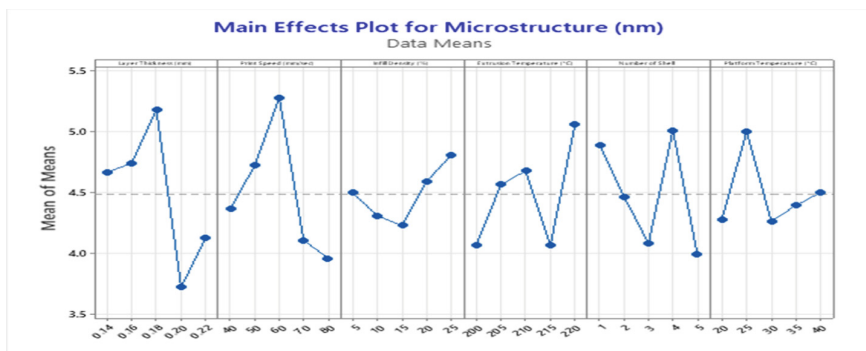


Fig. 6. Main effects plot for microstructure.

2.6 Genetic Algorithm Optimization

University of Michigan professor John Holland first up this idea for the genetic algorithm (GA) [18]. In this algorithm, each individual process parameters i.e., layer thickness, print speed, infill density, number of shells, extrusion temperature, and platform temperatures are coded mathematically. The objective of optimization is to provide an output with a gene composition that will maximize their performance. Initialization, Evaluation, Selection, Crossover and Mutation, and Recombination are the five main GA phases.

Initialization is the first step in the procedure. The initialized population $P^{[1]}$ is produced during this stage. The performance response parameters of each string are

evaluated. The individual is assigned a fitness score and a chance of survival based on their performance. Then, selection is carried out depending on each outputs chance of surviving. After the selection process, some of the population can join the mating pool. The individuals in this mating pool then exchange traits with one another, which results in crossover, which creates offspring P' with a mixture of features from the mating pool. The following generation $P^{[k+1]}$ after the k th generation is made up of this group of kids along with additional freshly generated random outputs. The loop repeats in this trend. To complete the optimization a stopping requirement should met, by the maximum number of iterations is reached and the convergence of the results. The GA optimization is done for all outputs of compression strength and microstructure in the following. This the whole operation is executed by creating a MATLAB programing code and optimization solver.

3 Optimization Results and Discussion

3.1 Compressive Strength

The genetic algorithm optimization has a regression function used as an objective function for compression strength which was defined as a function of process parameters and formed by the regression model as function $CS = \text{compression strength}(x)$;

$$CS = -(2.67 + 0.23 * x(1) + 0.00373 * x(2) - 0.00491 * x(3) - 0.00239 * x(4) + 0.0196 * x(5) + 0.00558 * x(6))$$

where; $x(1)$ = Layer Thickness (mm); $x(2)$ = Print Speed (mm/sec); $x(3)$ = Infill Density (%); $x(4)$ = Extrusion Temperature ($^{\circ}\text{C}$); $x(5)$ = Number of Shell and $x(6)$ = Platform Temperature ($^{\circ}\text{C}$).

The GA optimization for compression strength is done with a stopping criterion of exceeding the maximum number of 100 iterations (has no effects beyond this). The optimized process parameters are obtained as 0.22 mm of layer thickness, 79.988 mm/sec of print speed, 5.009% of infill density, 200.848 $^{\circ}\text{C}$ of extrusion temperature, 4.999 of number of shells and 39.998 $^{\circ}\text{C}$ of platform temperature to get a compressive strength of 2.835 MPa. The negative sign in the result indicates that the optimization algorithm is maximization of the compressive strength.

3.2 Microstructure

The genetic algorithm optimization has a regression function used as an objective function for microstructure grain length which was defined as a function of process parameters and formed by the regression model as *function* $MC = \text{microstructure}(x)$

$$MC = 1.17 - 10.43 * x(1) - 0.0143 * x(2) + 0.0181 * x(3) + 0.0297 * x(4) - 0.124 * x(5) - 0.0032 * x(6)$$

where; $x(1)$ = Layer Thickness (mm); $x(2)$ = Print Speed (mm/sec); $x(3)$ = Infill Density (%); $x(4)$ = Extrusion Temperature ($^{\circ}\text{C}$); $x(5)$ = Number of Shell and $x(6)$ = Platform Temperature ($^{\circ}\text{C}$)

The GA optimization for microstructure grain size is done with a stopping criterion of exceeding the maximum number of 100 iterations. The optimized process parameters are obtained as 0.22 mm of layer thickness, 79.998 mm/sec of print speed, 5.005% of infill density, 200.002 °C of extrusion temperature, 5 of number of shells and 27.759 °C of platform temperature to get a microstructure grain length of 3.0533 nm. Since the objective of the optimization of the algorithm is minimization, the function is positive in default.

4 Conclusions and Recommendations

4.1 Conclusions

In this paper, the optimization of FDM process parameters to improve Compressive strength and Micro-structures of PLA materials has been implemented using GA. The ANOVA was used for the analysis of mechanical properties. The major out puts of the research are presented as follow:

For the compression strength the ANOVA shows that the platform temperature is the most significant parameter for compression strength for fabrication of FDM followed by print speed, number of shells, infill density, extrusion temperature and layer thickness. The optimized process parameters are obtained as 0.22 mm of layer thickness, 79.988 mm/sec of print speed, 5.009% of infill density, 200.848 °C of extrusion temperature, 4.999 of number of shells and 39.998 °C of platform temperature to get a compressive strength of 2.835 MPa.

For the microstructure it is observed from the ANOVA the microstructure is highly influenced by layer thickness followed by print speed, number of shells, extrusion temperature, platform temperature and infill density. The optimized process parameters are obtained as 0.22 mm of layer thickness, 79.998 mm/sec of print speed, 5.005% of infill density, 200.002 °C of extrusion temperature, 5 of number of shells and 27.759 °C of platform temperature to get a microstructure grain size length of 3.0533 nm.

4.2 Recommendations

This paper specifies the effects of FDM process parameters such as layer thickness, print speed, number of shells, infill density, extrusion temperature and plate form temperature on the outputs of compression strength and microstructure using PLA filament material and the optimization was done using GA. In the future the following further study can be done.

- Conducting similar parameter studies on other additive manufacturing technology such as selective laser sintering, electron beam melting etc.
- Considering other FDM process parameters like raster angle, raster width, nozzle diameter, cooling rate with further parameter level settings.
- Extending this work to other materials like ABS, PTEEG, etc.

References

1. J. Jiang, X. Xu, and J. Stringer. Support structures for additive manufacturing: A review. *J. Manuf. Mater. Process.*, vol. 2, no. 4, 2018, <https://doi.org/10.3390/jmmp2040064>.
2. O. Abdulhameed, A. Al-Ahmari, W. Ameen, and S. H. Mian, "Additive manufacturing: Challenges, trends, and applications," *Adv. Mech. Eng.*, vol. 11, no. 2, pp. 1–27, 2019, <https://doi.org/10.1177/1687814018822880>.
3. E. Brancewicz-Steinmetz, J. Sawicki, and P. Byczkowska. The influence of 3d printing parameters on adhesion between polylactic acid (PlA) and thermoplastic polyurethane (tpu). *Materials (Basel)*, vol. 14, no. 21, 2021, <https://doi.org/10.3390/ma14216464>.
4. P. Wang, B. Zou, H. Xiao, S. Ding, and C. Huang. Effects of printing parameters of fused deposition modeling on mechanical properties, surface quality, and microstructure of PEEK. *J. Mater. Process. Technol.*, vol. 271, no. September 2018, pp. 62–74, 2019, <https://doi.org/10.1016/j.jmatprotec.2019.03.016>.
5. N. Naveed, "Investigate the effects of process parameters on material properties and microstructural changes of 3D-printed specimens using fused deposition modelling (FDM)," *Mater. Technol.*, vol. 36, no. 5, pp. 317–330, 2021, <https://doi.org/10.1080/10667857.2020.1758475>.
6. A. Mohanty *et al.* Parametric optimization of parameters affecting dimension precision of FDM printed part using hybrid Taguchi-MARCOS-nature inspired heuristic optimization technique. *Mater. Today Proc.*, no. xxxx, 2021, <https://doi.org/10.1016/j.matpr.2021.06.216>.
7. D. C. Anghel, D. M. Iordache, A. D. Rizea, and N. D. Stanescu. A new approach to optimize the relative clearance for cylindrical joints manufactured by fdm 3d printing using a hybrid genetic algorithm artificial neural network and rational function. *Processes*, vol. 9, no. 6, 2021, <https://doi.org/10.3390/pr9060925>.
8. N. H. Huu, D. P. Phuoc, T. N. Huu, and H. T. T. Thu. Optimization of the FDM Parameters to Improve the Compressive Strength of the PLA-copper Based Products. *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 530, no. 1, 2019, <https://doi.org/10.1088/1757-899X/530/1/012001>.
9. A. Dey, D. Hoffman, and N. Yodo, "Optimizing multiple process parameters in fused deposition modeling with particle swarm optimization," *Int. J. Interact. Des. Manuf.*, vol. 14, no. 2, pp. 393–405, 2020, <https://doi.org/10.1007/s12008-019-00637-9>.
10. T. Sai, V. K. Pathak, and A. K. Srivastava. Modeling and optimization of fused deposition modeling (FDM) process through printing PLA implants using adaptive neuro-fuzzy inference system (ANFIS) model and whale optimization algorithm. *J. Brazilian Soc. Mech. Sci. Eng.*, vol. 42, no. 12, 2020, <https://doi.org/10.1007/s40430-020-02699-3>.
11. D. Impens and R. J. Urbanic, "Assessing the impact of post-processing variables on tensile and compression characteristics for 3D printed components," *IFAC-PapersOnLine*, vol. 28, no. 3, pp. 652–657, 2015, <https://doi.org/10.1016/j.ifacol.2015.06.156>.
12. P. Yadav, A. Sahai, and R. S. Sharma, "Strength and Surface Characteristics of FDM-Based 3D Printed PLA Parts for Multiple Infill Design Patterns," *J. Inst. Eng. Ser. C*, vol. 102, no. 1, pp. 197–207, 2021, <https://doi.org/10.1007/s40032-020-00625-z>.
13. M. Xia, B. Nematollahi, and J. Sanjayan, "Influence of binder saturation level on compressive strength and dimensional accuracy of powder-based 3D printed geopolymer," *Mater. Sci. Forum*, vol. 939, pp. 177–183, 2018, <https://doi.org/10.4028/www.scientific.net/MSF.939.177>.
14. S. A. Oudah, H. B. Al-attraqchi, and N. A. Nassir. The Effect of Process Parameters on the Compression Property of Acrylonitrile Butadiene Styrene Produced by 3D Printer. vol. 40, no. May 2021, pp. 189–194, 2022.

15. S. Patibandla and A. Mian, "Layer-to-layer physical characteristics and compression behavior of 3D printed polymer metastructures fabricated using different process parameters," *J. Elastomers Plast.*, vol. 53, no. 5, pp. 386–401, 2021, <https://doi.org/10.1177/0095244320939995>.
16. R. Hernandez, D. Slaughter, D. Whaley, J. Tate, and B. Asiabanpour. Analyzing the tensile, compressive, and flexural properties of 3D printed ABS P430 plastic based on printing orientation using fused deposition modeling. *Solid Free. Fabr. 2016 Proc. 27th Annu. Int. Solid Free. Fabr. Symp. - An Addit. Manuf. Conf. SFF 2016*, pp. 939–950, 2016.
17. Z. Raheem. (PDF) Designation: D695 – 15 Standard Test Method for Compressive Properties of Rigid Plastics 1. no. July, 2019, <https://doi.org/10.1520/D0695-15> <https://www.astm.org/d0695-15.html>.
18. I. Journal. Improving the Performance of FDM Machine Objects by using Optimization Techniques.



Heat Exchanger Optimization for Thermoelectric Generating Power System

Kyeong-ho Shin, Yunho Jung, and Joo-Hyung Kim^(✉)

Laboratory of Intelligent Devices and Thermal Control, Department of Mechanical Engineering,
Inha University, Inha-Ro 100, Michuhol-Gu, Incheon 22212, Republic of Korea
joohyung.kim@inha.ac.kr

Abstract. Thermoelectric generator (TEG) is commonly used to harvest or scavenge the remnant energy from waste heat sources, such as automotive exhaust systems, cooling line of power plants, ocean temperature difference, human body, and industrial processes. The performance of TEG depends on the thermoelectric material and structural configuration. To obtain the maximum power from TEG module, the optimized design of heat exchanger is essential to realize the working power generating system. Internal structure of the heat exchanger plays a critical role to determine the efficiency of TEG by enhancing heat transfer between the hot and cold sides of the TEG module. We studied the effect of internal fin structure in heat exchanger with hot water feeding condition to obtain the best performance of TEG module array using computational simulations and compared with the experiments. From the experiments, we achieved 2 times higher power generation efficiency with a reasonable fin structure in heat exchanger, compared to those bare exchanger. From this study, the optimized fin structure in heat exchanger in TEG system will enhance the net power output of TEG system for waste energy source, which will promote renewable energy transform to help climate change issues.

Keywords: Thermoelectric generator (TEG) · Fin structure · Heat exchanger · Heat distribution · COMSOL multiphysics

1 Introduction

During the last century, economic development in many parts of the world has been associated with the growth of energy consumption, which was mainly driven using fossil fuels. Therefore, this excessive dependence on fossil fuels has caused several problems such as greenhouse gas emissions, air pollution issues including particulate matter and gradual depletion of fuel resources. The Intergovernmental Panel on Climate Change (IPCC) has identified an increase of greenhouse gas concentrations as the cause of global average temperature rise since the 20th century [1].

To reduce greenhouse gas concentrations, international organizations have recently begun to regulate gas emissions. The EU 2050 roadmap (European Climate Foundation, 2010) aims to reduce greenhouse gas emissions by 80% by 2050. Looking at a few

details, we can't ignore the energy harvesting technology in particular, the reduction in energy use in buildings (950 TWh/year) and industry (450 TWh/year), and the use of renewable energy [2].

To achieve the goal, the government introduces several important strategies to reduce dependence on fossil fuels and utilize renewable energy than conventional power system emitting CO₂ emissions. Energy harvesting technology is one of the popular methods, which can harvest the waste or remnant energy that is wasted from our daily life.

The large-scale thermal energy—not fully utilized yet—can be fully recovered and recycled. This discarded heat energy is typically converted to a stable temperature gradient and thermodynamically converted to useful work through the two heat sources of the heat engine. However, these thermal energy resources are low-temperature and unstable, leading the difficulty to be used efficiently. Energy harvesting from low-temperature heat source can only produce a low rate of energy conversion and difficult to be adapted to industrial application.

Numerous techniques are currently under investigation for energy harvesting. Among there are waste heat boilers, run around coil (RAC), direct and indirect contact condensation recovery, membrane distillation, heat pumps, heat recovery steam generators (HRSGs), heat pipe systems, organic Rankine cycle (ORC), Kalina cycle, piezoelectric, thermionic, thermophotovoltaic (TPV), and thermoelectric generator (TEG) [3].

TEGs utilize the Seebeck effect, in which different metals in contact generate electricity when there is a temperature difference between them [4]. These generators are typically used in relatively low or medium temperature ranges (80 to 500 °C) [5], and the net power generated from TEG increases as the temperature difference increases. The net power generated varies depending on how heat is supplied, so improving the device that supplies heat from the heat source to the TEG module can enhance the efficiency of the thermoelectric power generation system.

This paper focuses on a method for optimizing the fin structure of heat exchangers for the efficient use of thermal energy resources using in a low-temperature difference energy harvesting technology centered on low-temperature sources.

2 Heat Exchanger Design and Fin Structure

2.1 Fin Design in Heat Exchanger

We suggest that it is possible to derive the optimal heat exchanger design by using a simulation program before manufacturing the heat exchanger. The geometry of the 2D simulation was designed to model the flow field inside the heat exchanger, allowing for observation of the behavior of the internal fluid (see Fig. 1). To investigate the effects of different types of fins on fluid flow delay, a vertical fin structure, horizontal-vertical mixed fins, horizontal fins, and vertical fins were selected. To evaluate the influence of fin shape, trapezoid-shaped fins in front of a half section were chosen. In addition, a without fin structure was selected to compare the effects of the presence of fins. According to the fin array in the heat exchanger, the hot water can stay and transfer the heat from water toward the surface of heat exchanger, which can be transformed to the heat flux as the input energy to TEG module. Therefore, the surface temperature of heat exchanger will be important parameter to design the TEG based energy power system.

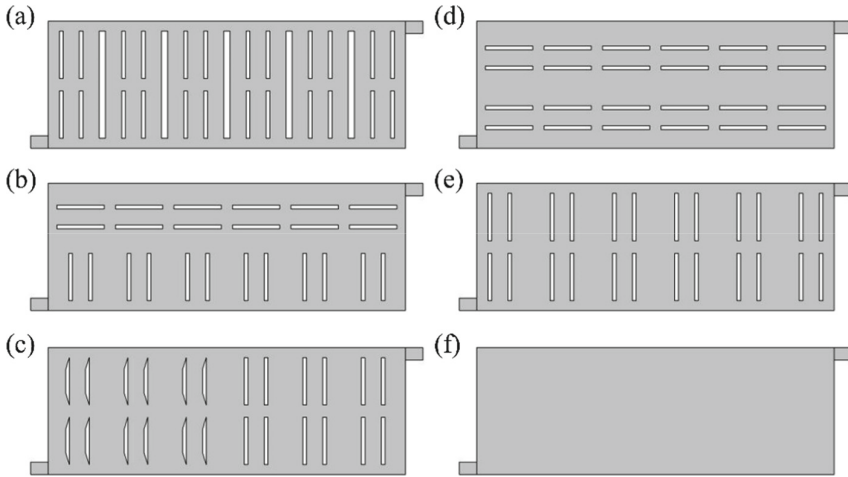


Fig. 1. 2D models of the flow field inside the heat exchangers with (a) vertical fin structure, (b) horizontal-vertical mixed fins, (c) trapezoid-shaped fins in front of a half section, (d) horizontal fins, (e) vertical fins, (f) without fin structure

Two types of simulations were conducted to investigate various thermal phenomena and estimate methods for more effective heat transfer via COMSOL Multiphysics. A 2D flow field analysis was performed to examine flow phenomena, while a 3D thermal analysis was carried out to study heat transfer phenomena and surface heat distribution between the heat exchanger and the fluid.

2.2 2D Flow Simulation in Heat Exchangers

In order to evaluate the fin structure and distribution in heat exchanger, thermal distribution inside the heat exchanger by hot water feeding was simulated. The model was simplified by visualization of the temperature gradient of the hot feeding water from pipe inlet and relatively cooled water in exit, where the temperature distribution of water in exchanger is time dependent variable condition by heat absorption into TEG modules attached on both surfaces of the heat exchanger.

Pre-Process. The material was set by water liquid of built-in material data. Mesh was generated using an automatic normal mesh (see Table 1 and Fig. 2). In the Multiphysics simulation, non-isothermal flow was modeled using the Reynolds-averaged Navier-Stokes $k-\omega$ method, taking into consideration the viscous flow by the interior wall of heat exchanger. The initial value of the fluid field was set to 20 °C, which is room temperature.

The inlet condition was set to 70 °C with care not to exceed the boiling point of water, and the speed was set at 5.94 m/s to simulate via actual flow rate from the bath circulator. The boundary condition in exit was set to zero-gauge pressure. To simulate heat diffusion over time, a time-dependent study was used for 10 s.

Table 1. Mesh quality of 2D simulation

Model	Number of elements	Average element quality
(a)	43538	0.8133
(b)	39156	0.8168
(c)	38443	0.8189
(d)	37786	0.8120
(e)	39516	0.8152
(f)	21870	0.9237

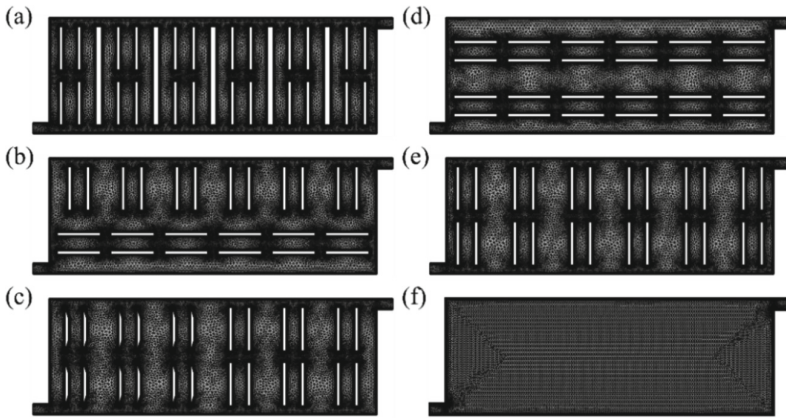


Fig. 2. Automatic normal mesh of the flow field inside the heat exchangers with (a) vertical fin structure, (b) horizontal-vertical mixed fins, (c) trapezoid-shaped fins in front of a half section, (d) horizontal fins, (e) vertical fins, (f) without fin structure

Post-Process. According to the results, differences in heat distribution caused by the arrangement of the fins. The time that the fluid stayed at each point and the resulting heat diffusion varied depending on the arrangement of the fins (see Fig. 3).

For clarity, the average temperature of the entire region was calculated and plotted as a graph (see Fig. 4). In the graph, the effect of heat distribution is divided into two groups. Near the one-second zone, it was observed that one group had a faster increase in temperature compared to the other group. It was hypothesized that the heat exchanger model with slower heat diffusion would be less efficiency.

In the case of the without fin structure model, it was observed that the average temperature increased rapidly in the first section before one second, but then the gradient decreased sharply. While it was assumed that the efficiency of this model may be higher, it was also noted that the gradient of the heat distribution rapidly decreases as it moves backward because of the fluid, which is the heat source, does not stay in the heat exchanger sufficiently and exits quickly due to the absence of fins. Therefore, to

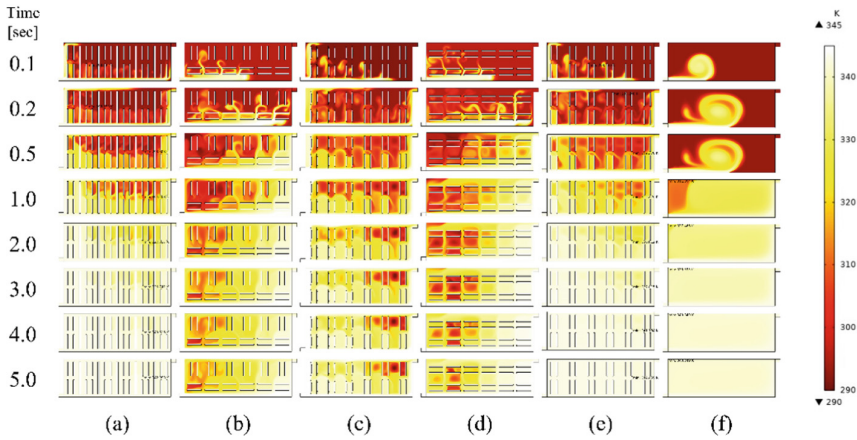


Fig. 3. Image of simulation of heat diffusion in flow field of six different heat exchanger shapes

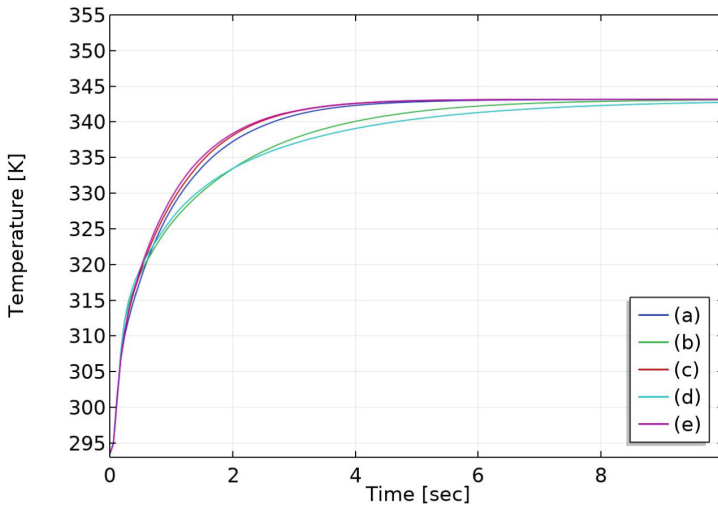


Fig. 4. Graph of average temperature per time in flow field of five different heat exchanger shapes without fin structure

understand the impact of these unique results on performance, actual experiments were conducted (Fig. 5).

2.3 3D Heat Distribution in Heat Exchangers

To check the thermal distribution of the part that comes into contact plate with the thermoelectric module, thermal behavior of heat exchangers via fluid heat source was simulated by 3D non-isothermal model.

Pre-Process. The 3D non-isothermal simulation was used to study the thermal behavior of the heat exchanger, which included the geometry of the heat exchanger and the

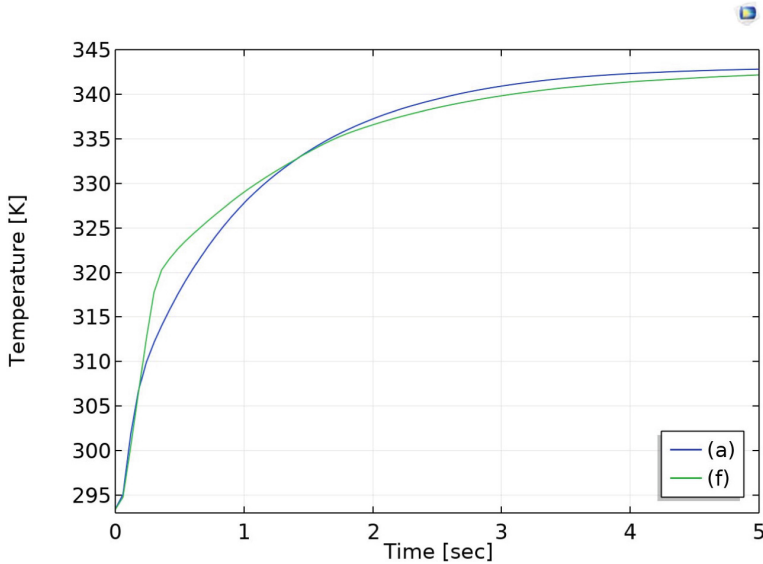


Fig. 5. Temperature profile with and without inner fin structure in heat exchanger

internal flow field (see Fig. 6). The material for the fluid was set to water, the same as in the 2D simulation, while the heat exchanger material was set to aluminum. Mesh was generated using an automatic normal mesh (see Table 2 and Fig. 7). The without fin structure model, which was less complicated, produced a similar average element quality with fewer elements.

For simulation of the heat exchanger, a laminar flow method was used for the non-isothermal flow to reduce the solving time and computational resource requirements. The initial values for the heat exchanger and fluid fields were set to 20 °C and 80 °C, respectively. The inlet was set at a fluid temperature of 80°C with a maximum velocity of 5.94 m/s using a step function to ensure convergence, and the outlet was set to zero-gauge pressure. To simulate heat transfer, two boundary conditions were set for the contact surface: one at a constant temperature of 0 °C and another at the heat flux of 2000W per area. Other surfaces were set to thermal insulation (see Fig. 8). A time-dependent study was used to observe the thermal behavior over 10 s.

Post-Process (2D Thermal Behavior between Fluid and Heat Exchanger) The thermal behavior images obtained from the 2D vertical slice of the simulation were compared to investigate the correlation between the shape of the heat exchanger and the thermal behavior. The cross-section of the heat exchanger at its center was observed, and the temperature distribution over 30 seconds was confirmed. In this case, the boundary condition of the contact surface was set to 0 °C.

After comparing the thermal behavior images obtained from the 2D vertical slice, it was confirmed that the heat distribution of the heat source was generally even and stable across all models, except for the without fin structure model in which the temperature of the heat source fluctuated due to the influence of the contact surface (see Fig. 9).

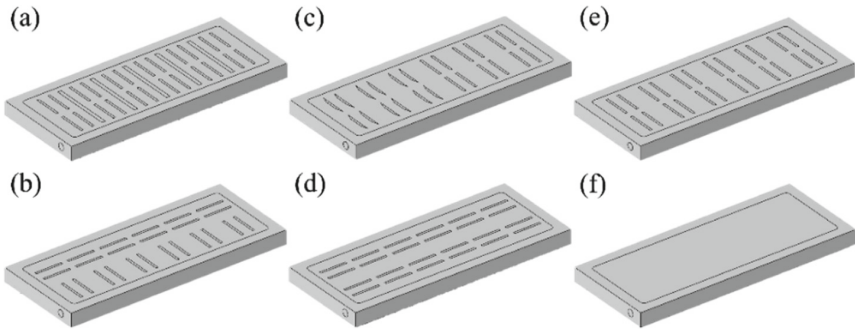


Fig. 6. 3D models of the heat exchangers with (a) vertical fin structure, (b) horizontal-vertical mixed fins, (c) trapezoid-shaped fins in front of a half section, (d) horizontal fins, (e) vertical fins, (f) without fin structure

Table 2. Mesh quality of 3D simulation

Model	Number of elements	Average element quality
(a)	1233901	0.6551
(b)	1109044	0.6563
(c)	1042922	0.6531
(d)	1089899	0.6553
(e)	1000267	0.6555
(f)	397791	0.6704

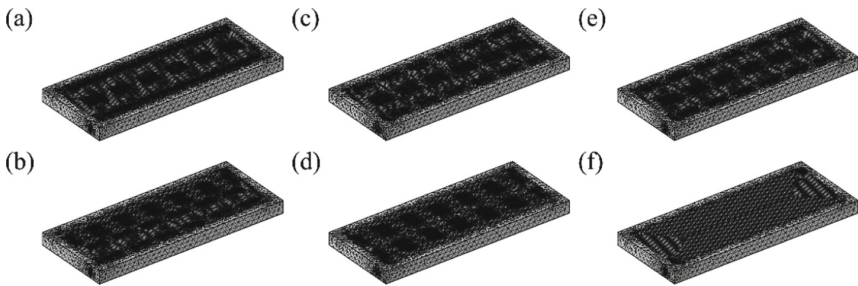


Fig. 7. Automatic normal mesh of the heat exchangers with (a) vertical fin structure, (b) horizontal-vertical mixed fins, (c) trapezoid-shaped fins in front of a half section, (d) horizontal fins, (e) vertical fins, (f) without fin structure

The without fin structure model showed lower temperatures of the internal fluid heat source and non-uniform heat distribution compared to other models. As suggested in the 2D simulation, we intend to conduct a simulation analysis of the without fin structure model with unique results through experiments.

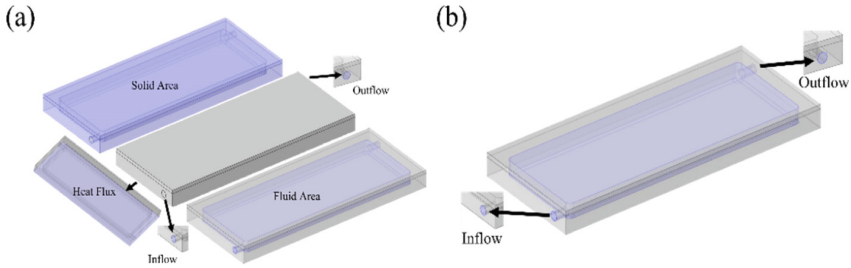


Fig. 8. Boundary condition areas of heat exchanger: (a) heat transfer, (b) laminar flow

Post-Process (Temperature Distribution of Contact Surface). The heat transfer performance according to the shape of the heat exchanger was investigated by comparing the heat distribution image of the contact surface of the simulation (see Fig. 10). The temperature distribution on the contact surface was observed over 10 seconds, with a boundary condition of 2 kW heat flux.

According to the results, compared with the 2D flow field simulation result, it can be confirmed that the group in which heat is evenly distributed within 10 s and the group in which heat is unevenly distributed are the same. When the heat source fluid diffusion was relatively slow, the heat on the surface was also found to be distributed slowly.

In the without fin structure model, which was classified as a specific case in the previous simulation, it was confirmed that the heat distribution on the surface was slow. It is estimated that the lack of heat transfer occurs partially without this effect, unlike the relatively even distribution where the fins induce the flow direction to the vulnerable point.

In addition, it was observed that thermal diffusion occurs preferentially in a section with fins during the initial 0.1 s to 0.5 s. This is due to an increase in the contact cross-sectional area with the heat source. However, it is important to note that having too many fins can interrupt the flow of the heat source fluid, resulting in lower efficiency.

3 Experimental Test

3.1 Test Apparatus

To evaluate the effectiveness of heat exchangers in a laboratory-scale thermoelectric power generation system and compare them with simulations, selected six different heat exchangers were produced and tested. The efficiency of each shape was compared using a both heat exchangers generate system (see Fig. 11).

In the experiments, the temperature was set to 80 °C, and the heat source was water circulating through a liquid sink heat exchanger. The thermoelectric modules were cooled using water coolant to 0 °C. The temperature difference (ΔT) was set to 80 °C, and a series connection of 12 thermoelectric modules was tested (see Fig. 12).

All the experiments used same thermoelectric modules (TG 12–8, Marlow Industries) and same thermal interface material (TIM). Thermal conductivity of TIM is 9.0 W/m-K (SC1001, YOUNGYIEL PRECISION). For supplied the heat source and coolant, heat

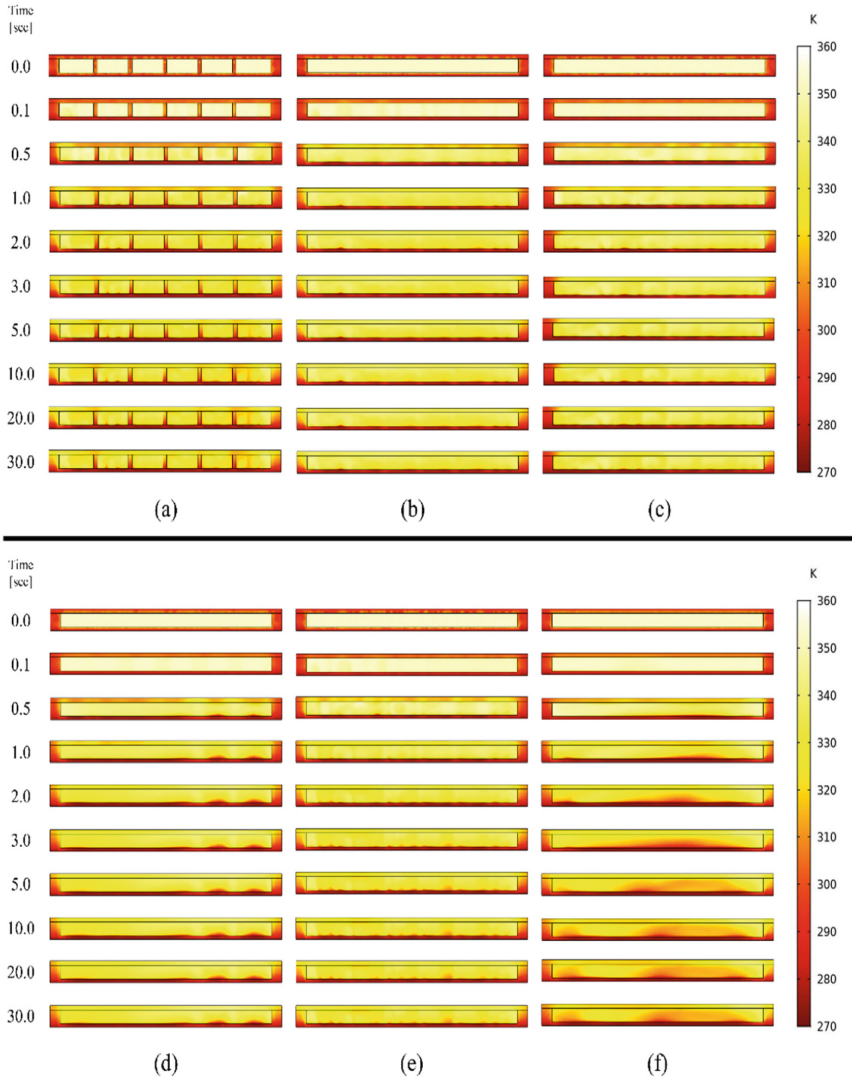


Fig. 9. The thermal behavior images obtained from the 2D vertical slice of six different heat exchanger shapes

exchanger was connected in refrigerating bath circulator (RW-3025G, JEIO TECH). Detection of temperature was performed by thermocouple (T-type, Omega).

The density of the heat source water, flow rate, and heat capacity were assumed to be constant properties to calculate the input heat transfer rate. The input heat transfer rate of the heat source was calculated using the following formula:

$$Q_{in} = \dot{m} \cdot c_{water} \cdot \Delta T \tag{1}$$

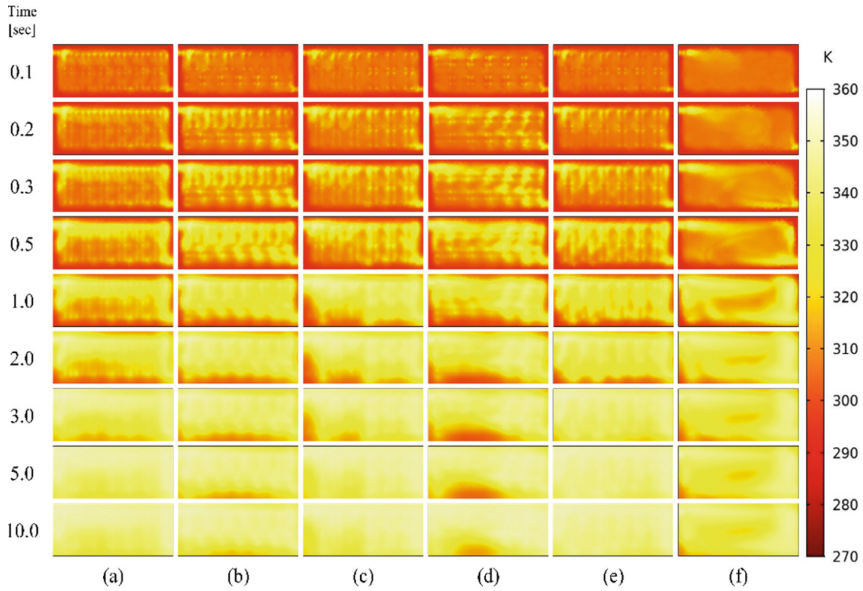


Fig. 10. Temperature distribution image of contact surface of six different heat exchanger shapes

where, \dot{m} is the mass rate of hot water into heat exchange, c_{water} is the heat capacity of hot feeding water. The ΔT was estimated by measuring the temperature difference between the inlet and outlet of the high-temperature heat exchanger. In addition, the output power P of the TEG with hot water heat exchange was calculated using the following formula:

$$P_{output} = V \cdot I \quad (2)$$

By measured the induced voltage V and current I from TEG module, we can estimate power generation efficiency η_{TEG} by following equation:

$$\eta_{TEG} = \frac{P_{output}}{Q_{in}} \times 100 \quad (3)$$

Using this efficiency, the performance of each heat exchanger can be compared.

3.2 Results

Six different heat exchanger shapes were manufactured to test their effectiveness in a low-temperature heat source. The shapes included vertical fin structure, horizontal-vertical mixed fins, trapezoid-shaped fins in front of a half section, horizontal fins, vertical fins, and without fin structure (see Fig. 13). To ensure the reliability of the results, the experiments were performed three times, with each experiment lasting half an hour. The tolerance of temperature was $\pm 1^\circ\text{C}$ for each heat exchanger, and the ΔT between the heat exchanger was calculated using logarithmic mean temperature difference (LMTD).

The result is consistent with the simulation conducted under the temporarily assumed conditions to identify trends. It can be seen that the groups with efficiency above 1%,

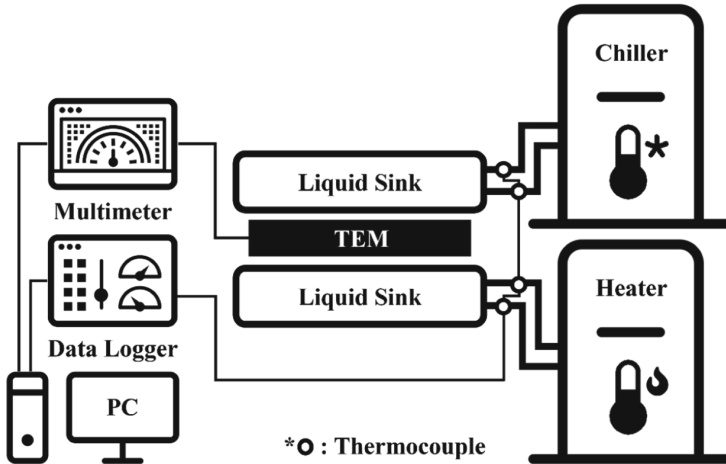


Fig. 11. Schematics of experiment system using both liquid sink heat exchangers

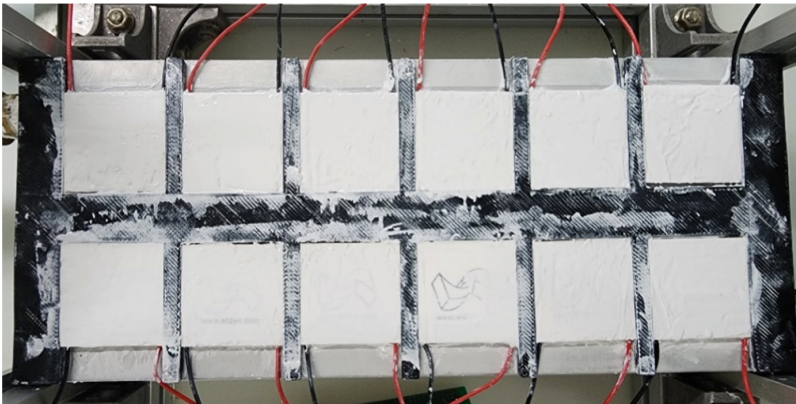


Fig. 12. 12 series connection TEG in heat exchangers

including a vertical fin structure with an efficiency of 1.23%, trapezoid-shaped fins in front of a half section with an efficiency of 1.06%, and vertical fins with an efficiency of 1.07%, and the groups with efficiency below 1%, including horizontal-vertical mixed fins with an efficiency of 0.85% and horizontal fins with an efficiency of 0.99%, are divided (see Table 3), which is the same as the groups divided in the simulation (see Fig. 5). In addition, it was confirmed that the power generation efficiency of the without fin structure model had a performance difference of about twice as much, with a power generation efficiency of 0.64%, compared to the models with fins, which had a maximum power generation efficiency of 1.23%, which is also inferred from the simulation (see Fig. 6). Based on the experimental and simulation results, it can be observed that the power generation significantly decreases due to phenomena where the heat source cannot stay

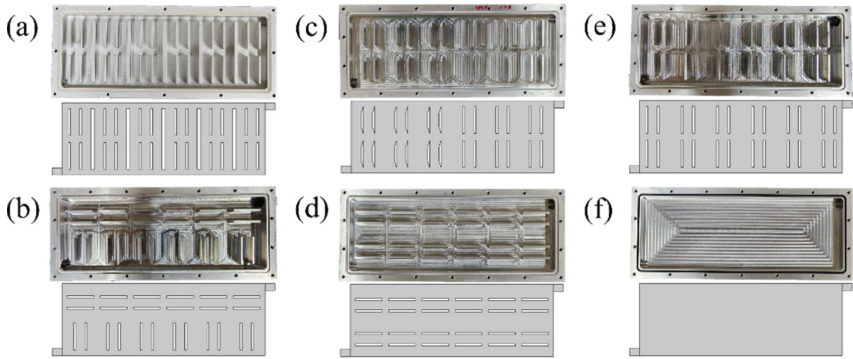


Fig. 13. Shape of various heat exchangers with (a) vertical fin structure, (b) horizontal-vertical mixed fins, (c) trapezoid-shaped fins in front of a half section, (d) horizontal fins, (e) vertical fins, (f) without fin structure

sufficiently in the heat exchanger and is discharged when there are no fins, as compared to the simulation of the structure model without fins (see Figs. 3 and 9).

Table 3. Efficiency results of six different heat exchangers

Model	ΔT [°C]	Generated power [W]	Input Q by hot side [W]	Efficiency [%]
(a)	82.11	34.27	2481.39	1.23
(b)	82.08	24.06	2839.68	0.85
(c)	81.68	35.74	3387.11	1.06
(d)	82.26	21.93	2205.98	0.99
(e)	81.85	34.68	3242.49	1.07
(f)	82.18	17.54	2728.78	0.64

Therefore, fin increases the time for the heat source to stay inside the heat exchanger, and it can be optimized by appropriately adjusting the smooth heat distribution and fluid residual time according to the shape and arrangement of fin.

4 Conclusions

In this paper, we conducted a computational simulation to investigate the effect of the internal fin array on the TEG power generation of heat exchanger by hot water feeding system. The simulation results were validated by manufacturing an actual heat exchanger and measuring the power generation performance.

The simulation was conducted under temporarily assumed conditions to identify trends, where six different heat exchanger shapes were analyzed under identical boundary conditions to determine the variation in temperature distribution and thermal behavior.

These differences were then used to predict the TEG power generation based on the actual heat transfer performance, which can give the guideline to design the optimized heat exchangers for waste water energy from a power plant.

In the actual experiments, the impact of temporarily interrupting the flow of fins inside the heat exchanger on heat transfer performance was examined by comparing the power generation efficiency of the without fin structure models to other models. In addition, to validate the simulation results, we compared the performance predictions obtained through simulation with actual performance, and proposed a method based on simulation to predict the performance of future optimal heat exchanger fin structures.

In the simulation, the vertical fin structure, trapezoid-shaped fins in front of a half section, and vertical fins models were predicted to have the best efficiency, with the vertical fin structure having the highest efficiency of 1.23%, which was confirmed in the actual experiment.

However, the simulation method used in this paper assumed various boundary conditions, so it is challenging to predict the performance with a single simulation result, and simulation data of a comparable heat exchanger model is required for accurate predictions.

In future research, experiments will be conducted using methods to mitigate the influence of the surrounding environment, and the performance differences according to various methods of applying Thermal Interface Material (TIM) will be investigated. In addition, we plan to utilize a new simulation approach to predict heat exchanger performance and develop a power generation prediction model using thermoelectric physics.

References

1. Griggs, D. J., Noguer, M.: Climate change 2001: the scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change. *Weather* 57(8), 267–269 (2002)
2. Barbet-Massin, M., Thuiller W., & Jiguet, F.: The fate of European breeding birds under climate, land-use and dispersal scenarios. *Global Change Biology* 18(3), 881-890 (2012).
3. Jouhara, H., Khordehgah, N., Almahmoud, S., Delpech, B., Chauhan, A., & Tassou, S. A.: Waste heat recovery technologies and applications. *Thermal Science and Engineering Progress* 6, 268-289 (2018).
4. He, W., Zhang, G., Zhang, X., Ji, J., Li, G., & Zhao, X.: Recent development and application of thermoelectric generator and cooler. *Applied Energy* 143, 1–25 (2015).
5. Markets and Markets. (2019). *Thermoelectric Generators Market*.
6. Esarte, J., Min, G., & Rowe, D. M.: Modelling heat exchangers for thermoelectric generators. *Journal of Power Sources* 93(1-2), 72-76 (2001).
7. Lesage, F. J., Sempels, É. V., & Lalande-Bertrand, N.: A study on heat transfer enhancement using flow channel inserts for thermoelectric power generation. *Energy Convers Manage* 75, 532-541 (2013).

Automatic Control



A Physicomimetic Approach to Distributed Intelligence and Control of Autonomous Watercraft

Nikolaos I. Xiros¹(✉) and Erdem Aktosun²(✉)

¹ Boysie Bollinger School of Naval Architecture and Marine Engineering, University of New Orleans, 2000 Lakeshore Dr, New Orleans, LA 70148, USA

nxiros@uno.edu

² Department of Shipbuilding and Ocean Engineering, İzmir Katip Çelebi University, Havaalanı Sosesi Cd, İzmir 35620, Turkey

erdem.aktosun@ikcu.edu.tr

Abstract. A square law of celestial mechanics is investigated as a driving virtual force law toward the development of swarms with distributed, i.e. not centralized, control and intelligence. It has been argued that decentralization of the control structure and task can enhance the robustness of the controllers involved and improve scalability. The approach in this work tackles the problem by adopting the concept of virtual forces which in this case follow the square law of attraction e.g. a nonlinear spring. The concept is adapted to the time and space, i.e. distance scales relevant to small, low-cost watercraft to be applicable. Various scenarios are investigated. All watercraft are assumed to have identical physical characteristics and dynamic responses but with small fluctuations and perturbations superimposed. Watercraft dynamics are simplified to model forward motion and rotation about the yaw axis; more detailed and precise models can also be used if desired or required. The local-loop controllers, which conventionally are called autopilots, ensure that the acceleration of the craft is adhering to the one dictated by the virtual, square law force field. Synthesis of the square law for the virtual force field involves amongst other feasibility of the derived governing dynamics as well as energy consumption, time to terminal conditions, etc.

1 Introduction

Unmanned Surface Vehicles (USVs) are self-contained unmanned untethered vessels that can transit on the surface of the water autonomously or through remote control. Unlike conventional manned surface vessels that are usually large and costly to build and operate, USVs are typically smaller in size and lower cost resulting from the reduced payload requirement due to being unmanned. In manned vessels, much of the volume and weight is necessary to support the activities (such as control, navigation, maintenance, and mission-related tasks), and sustainment (such as berthing, feeding, and entertainment) of the human occupants that recursively increases the size, volume, and power requirements. USVs have no such requirements and therefore are typically many times smaller and more efficient than manned surface vessels.

In the last two decades, significant effort has been invested in the development of Unmanned Underwater Vehicles (UUVs), while only a small effort has focused on Unmanned Surface Vessels/Autonomous Surface Vessels (USVs/ASVs). The major efforts in the design of USVs have focused on two areas: platforms for hydrographic data acquisition [1–3], and signal relay platforms that provide positioning and communications capabilities through the air-sea interface for UUVs [3–5].

The work presented here is part of a larger project that aims to develop a new concept for guidance and control of swarms of marine, or more general, vehicles based on a novel concept termed as robust probabilistic control. In particular, a low-cost solution to tackle the problem is currently under development that addresses the problem of navigation with minimal information exchange requirements.



Fig. 1. (a) The surface vehicle without most of her outfit (left) (b) unmanned underwater vehicle (middle) & (c) rotorcraft (right)

In this sense, a multi-purpose Autonomous Surface Vehicle (ASV) that is a low-cost mobile surface platform (Fig. 1a) has been designed and developed. The system is integrated with a motion measurement package to aid in navigation, and control, and to enhance dynamic performance. This single ASV can also be outfitted with acoustic communication systems to provide position updates and allow underwater vehicles, like in Fig. 1b, or agents to communicate while in transit and surveying. It is also possible to interact with the underwater vehicle to change the mission through an operator communicating with the USV via an RF uplink from shore or a distant vessel [6] or even a hovering aircraft like e.g. the extremely low-cost rotorcraft shown in Fig. 1c.

1.1 Potential Application Framework Formulation

The majority of tactical scenarios envisioned for performing autonomous underwater cueing for targets or objects (like proud mine-like objects, especially in large littoral areas), involve swarms (flotillas) of unmanned marine vehicles, underwater and possibly surface or even aerial. In specific, the proposed strategy for countermine operations relies on heterogeneous swarms in the sense that they consist of at least two types (tiers) of vehicles as shown in Fig. 2(a) A Tier-1 surface vehicle which can be a greater-scale version of the boat in Fig. 1a. This type of vehicle will carry radar, hull-mounted sonar,

compass, magnetometer(s) and proximity sensors, cameras (infrared or optical), GPS receiver(s), AIS and ECDIS as well as have the capacity to carry and deploy the platform of various sonar types which are unmanned underwater vehicles similar to the one in Fig. 1b or towed non-self-propelled platforms (“fish”). Further, the Tier-1 vehicles will carry considerable computing and communication assets necessary for the mission. (b) Tier-2 vehicles may be low-cost aerial units like the one shown in Fig. 1c outfitted with low-cost sensors like camera, GPS and compass as well as appropriate actuation for steering and throttle control.

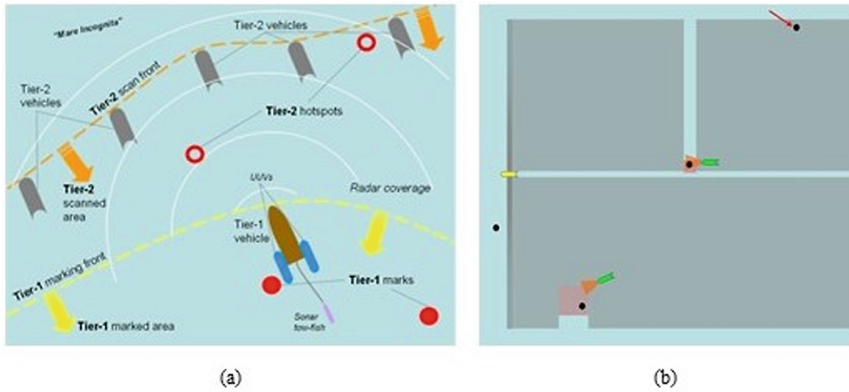


Fig. 2. An approach to mine detection and marking (a) 2-tier swarm structure, up and (b) 2D field of operations, down.

Various strategies have been proposed concerning the level of cooperation between the various vehicles as well as diversification of the flotilla in the sense of the number and types of vehicles used. However, up to now there is no general framework to comparatively assess these strategies and schemes. In this end, a unified performance index (or metric) is needed encompassing various quantities, reflecting mission accomplishment as well as the associated economic costs of system acquisition and operation.

Although, economic costs are more or less straightforward to define, mission accomplishment assessment needs in most cases some additional work. On the other hand, a generic performance index that comes to mind when military operations, financial or strategic decisions are of interest is probabilities associated with various events. Returning to the object cueing scenario and limiting our investigation here to a 2D rectangular area (or field) as shown in Fig. 2b, without significant loss of generality, the following spatial function (map) is introduced for any point (x,y) in our 2D area.

$f(x,y) = 0$ if target is **not** present at (x,y) ; **1** if target **is** present at (x,y) .

Furthermore, we assume that f is time-invariant. Therefore, as time goes on any object cueing system covers an increasingly larger portion of the field and assigns 0 or 1 to each one of the points ‘visited’; however, for each decision made there is an associated *probability of error* $p_e(x,y; t)$. It is noted here that the probability of error is time dependent because each point in the field may be ‘revisited’ as many times needed

and by any type of vehicles. Therefore, a reasonable setup of the object cueing problem is minimize t_0 for which it holds $p_e(x,y; t) < p_{e0}$ for $t > t_0$ and all (x,y) in the field.

If the economic costs of system acquisition and operation are included in the minimization problem above, then the overall problem can be approached by multi-objective optimization methods.

Solving the problem stated above requires the determination of various system parameters including but not limited to number of vehicles and vehicle types, vehicle speed, number of sensors and sensor types, level of cooperation, field coverage scheme etc. Evidently, additional constraints will be introduced to the problem due to the physics and engineering involved as e.g. vehicle speed ceiling, coverage scheme limitations due to maneuverability limitations etc.

A simple model for taking into account the main features of a possible configuration is shown in Fig. 2b. In this simplified version, there are two types of vehicles regardless of their implementation: Search vehicles (SV, shown in yellow) and Classification vehicles (CV, shown in green). Also, search sensors relying on the echoing principle (e.g. sonar) are considered.

SVs are equipped with non-classifying, long-range, side-scanning sensors. They provide with field scans of resolution decreasing with respect to the distance from the vehicle. As shown in Fig. 2b, the area marked as a result of an echo signal is increasing with the round trip time of the signal. Furthermore, this area is in any case larger than the area covered by the target itself as there are a number of uncertainties involved, including the positioning uncertainty of the SV. By assuming a 'line-of-sight' (LOS) mode of operation for the sensors, any target detected causes obstruction of view; therefore, the area 'behind' a detected target cannot be scanned unless the SV approaches the obstructed area from another angle or during another pass.

The probability of error for any point *inside the region scanned* by a specific SV, v , up to time t is given by the following relationship.

$$p_e(x, y; v@t) = f(x, y) \times p_{miss}(\mathbf{d} < (x, y), \text{traj}v@t>) \\ + [1 - f(x, y)] \times \sum \{ [1 - p_{miss}(\mathbf{d} < (x_k, y_k), \text{traj}\{v;t\}>)] \\ \times p_{1|0}(x, y; x_k, y_k; \text{traj}\{v@t\}) \}$$

In the above, $\mathbf{d} < (x,y), \text{traj}\{v@t\} >$ stands for the Euclidean distance of a point (x,y) from the trajectory, $\text{traj}\{v@t\}$, of the specific SV, v , up to time t ; if not defined is set to infinity.

Distribution $p_{miss}(\mathbf{d} < (x,y), \text{traj}\{v@t\} >)$ refers to the probability density function to *miss* a target in the manner indicated by the red arrow in Fig. 2, provided that point (x,y) lies in the area encompassed by a target; $p_{miss}(\mathbf{d} < (x,y), \text{traj}\{v@t\} >)$ is expected to be an increasing function of $\mathbf{d} < (x,y), \text{traj}\{v@t\} >$ and actually to saturate to unity if this distance lies above a threshold. Also, p_{miss} depends on the capabilities of the sensors mounted on the SV.

Distribution $p_{1|0}$ stands for the probability density function that a point (x,y) belongs to the marking area assigned to the k -th target, if detected by SV v until time t . It is reasonable to assume that $p_{1|0}$ is a decreasing function of the distance of point (x,y) from the center location of the k -th target (the larger the distance between the two points the

less probable to be included in the same ‘patch’) and an increasing function of $d < (x_k, y_k), \text{traj}\{v;t\} >$ because of the ‘spreading’ caused by reduced sensor resolution in the far field.

$$p1|0(x, y; x_k, y_k; \text{traj } v@t) = p1|0(d < (x, y), (x_k, y_k) >, d < (x_k, y_k), \text{traj } v;t >)$$

In the above, $d < (x_1, y_1), (x_2, y_2) >$ stands for the Euclidean distance between two points on the plane.

It would be the summation in the relationship for $p_e(x, y; v@t)$ spans all the targets in the field of interest. If (x, y) does not lie in the region already scanned by the specific SV, v , until time t , it is easy to verify that $p_e(x, y; v@t) = 1$ if $f(x, y) = 1$ and $p_e(x, y; v@t) = 0$ if $f(x, y) = 0$ and no target in some vicinity of (x, y) has been scanned yet. Although, it could have been defined differently, these facts make intuitive sense.

The relationship for $p_e(x, y; v@t)$ is valid provided some underlying assumptions hold. The targets must be distributed *reasonably sparse* so that the patches marked for two or more neighboring objects by the SVs do not overlap. Furthermore, the probability of ‘spontaneous triggering’ of the sensors, in the sense that they detect a target at a point where none is present in the vicinity, is considered negligible. Although these assumptions simplify the expression for $p_e(x, y; v@t)$, they can be dropped, if needed, by appropriate extensions.

A way to tackle such requirements expressed in a probabilistic framework is to consider techniques *Robust Stochastic Control*. Indeed, stochastic control deals with systems where the objective is to minimize or maximize the probability of an event related to a dynamic system or random signal by manipulating one or more of a driving signal. In conventional stochastic control, e.g. [7–9], the full dynamics of the process or system under control is considered. Given recent advances in self-organization, however, an alternative path could be followed. Specifically, the swarm’s vehicles will be driven by local, decentralized and in effect distributed deterministic control laws that require the knowledge by each vehicle of its own motion or other variables as well as those of the immediately neighboring units, but in no case the state variables of the entire formation. This scheme is usually implemented as a set of “virtual forces” exerted between neighboring vehicles, resembling the forces arising between the molecules or particles of a solid, liquid or gaseous substance; thereof the term physicomimetics or artificial physics. A critical step needed towards the direction of applying this technique, which is referred to as Robust Probabilistic Control is the migration from the concept of trajectories employed to describe the time evolution of the phase (state) space of a dynamical system to that of the *phase space probability distribution function* (or *measure*), $\rho(q, p; t)$ of the dynamical system. Consider a dynamical system with canonical coordinates q_i and conjugate momenta p_i , where $i = 1, \dots, n$. Then the phase space distribution $\rho(q, p)$ determines the probability $\rho(q, p)d^n q d^n p$ that the system will be found in the infinitesimal phase space volume $d^n q d^n p$. The concept and its applications will be investigated in full detail in a follow-up work.

Probabilistic measures of mission success like the ones introduced earlier can in turn be linked to individual vehicle dynamics through Liouville's Theorem. This theoretical tool depicts the dynamics of the collective behavior of an autonomous swarm on the basis of the phase-space probability distribution measure concept, instead of some more conventional approach based on individual vehicle trajectories. Specifically, Liouville's theorem governs the evolution of the distribution function $\rho(q,p;t)$ in time according to the following partial differential equation:

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \sum_{i=1}^n \left(\frac{\partial\rho}{\partial q_i} \dot{q}_i + \frac{\partial\rho}{\partial p_i} \dot{p}_i \right) = 0 \quad (1)$$

As a direct consequence of Liouville's theorem, the following equation can be obtained for the distribution function of a swarm's phase space.

$$\frac{\partial\rho}{\partial t} = [H, \rho] \quad (2)$$

where H is the Hamiltonian and the Poisson bracket between two functions of the canonical coordinates is defined by the following.

$$[f, g] = \sum_{i=1}^n \left(\frac{\partial f}{\partial q_i} \cdot \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \cdot \frac{\partial g}{\partial q_i} \right) \quad (3)$$

The Newtonian equation of motion for each swarm vehicle will, in effect, assume the following form.

$$\vec{\mathbf{a}}_i = \frac{1}{m_{i,\text{virtual}}} \cdot \sum_{\substack{j \\ i \neq j}} \vec{\mathbf{F}}_{ij,\text{virtual}} \quad (4)$$

In the above, \mathbf{a}_i is the acceleration of the i -th vehicle, $m_{i,\text{virtual}}$ the virtual mass assigned to the vehicle by the artificial physics, and $\mathbf{F}_{ij,\text{virtual}}$ the virtual force exerted on the i -th vehicle by the j -th neighbor per the physicomimetic interaction law governing the swarm dynamics.

The challenges related with swarm vehicle navigation and guidance are complex and can only be tackled using a comprehensive and non-trivial framework for the design of the local-loop controller that can guarantee compliance of the vehicle to artificial physics dynamics. The actual equation of motion of an arbitrary vehicle, assuming a material point model, is as follows.

$$\vec{\mathbf{a}}_i = \frac{1}{m_i} \cdot \left[\sum \vec{\mathbf{F}}_{\text{exogenous}} + \sum \vec{\mathbf{F}}_{\text{control}} \right] \quad (5)$$

In the above, m_i is the actual mass of the i -th vehicle, $\Sigma \mathbf{F}_{\text{exogenous}}$ the resultant exogenous force exerted on the i -th vehicle by the environment and $\Sigma \mathbf{F}_{\text{control}}$ the resultant control force exerted on the i -th vehicle by thrusters, control surfaces etc. To make vehicle

dynamics comply to artificial physics, as defined in the previous task, the resultant control force has to be as follows.

$$\sum \vec{\mathbf{F}}_{\text{control}} = \frac{m_i}{m_{i,\text{virtual}}} \cdot \sum_{\substack{j \\ i \neq j}} \vec{\mathbf{F}}_{ij,\text{virtual}} - \sum \vec{\mathbf{F}}_{\text{exogenous}} \quad (6)$$

The above clearly shows the issues that may arise if the physicomimetic control law generates excessively high forces or the virtual mass value is small compared to the actual mass value of the vehicle. This can lead to excessive control force requirements that can make the artificial physics scheme infeasible. This is why the local-loop controller design is a non-trivial aspect and needs to be investigated at least for one type or class of practical vehicles.

As in other similar cases, two or more subsystems either linear or nonlinear and either lumped or distributed, are coupled through processes that commonly are highly nonlinear. In our case, an appropriate identification method is employed, as explained later on, which comes as a generalization of the identification methodology for linear systems in phase-plane. Since the physicomimetic governing control laws may need to be continuously reconfigured, the option to adapt accordingly the gains of the local-loop controllers will be investigated in order to implement them in a follow-up effort. Such an approach will further increase system responsiveness and adaptability. Such adaptability enables for broad real-time reconfiguration of an autonomous swarm which in turn could convert routinely multi-sortie missions, due to the need for off-line reconfiguration, to single-sortie ones not needing real-time oversight. The above work clearly demonstrates the need for vehicle dynamics identification for local-loop controller tuning and adaptation which will be addressed in future work.

2 Methodology

The methodology is now investigated in the case of a single vehicle. The physicomimetic control law is that of a nonlinear spring with a quadratic force law. One end of the virtual spring is attached to the origin which is the homing point the vehicle needs to reach. The free end of the spring is attached to the center of mass of the vehicle. A second rotational nonlinear spring-like mechanism applies to adjusting the heading of the vehicle, so that the centerline (longitudinal stern-bow) axis of the vehicle boat gets aligned with that of the onboard radius vector pointing from the vehicle's center to the origin (target).

Figure 3 shows the supervisory controller with a discrete-time model. We use this model to implement the control law with the feedback information. As can be seen, the supervisory controller generates forward speed and yaw rate setpoints. The feedback controller then, if designed properly, makes sure that the setpoints issued by the supervisory controller are implemented and followed. This is achieved by using the thrusters and maneuvering assets of the boat. We now focus on the algorithm generating the setpoints, while in a future work the feedback controller synthesis will be addressed too.

In particular, consider the setup shown in Fig. 4, the vehicle (boat) is shown on a plane map at a certain location determined by the x- and y-coordinate (North-South and

East-West axis respectively). The vehicle is assumed to have (significant) speed along its centerline axis only pointing forward. The heading of the vehicle i.e. the angle of its forward velocity with respect to the North semiaxis is the third and final degree of freedom considered. As a result, the feedback information to implement the control law is the x-and y-coordinate as well as the compass reading for heading angle ψ . Using these, the onboard radius r and its angle with respect to the semiaxis pointing to the North can be calculated as the angle ϕ in the range $-\pi \leq \phi < \pi$, calculated as follows.

$$r = \sqrt{x^2 + y^2}, \cos \phi = \frac{x}{r}, \sin \phi = \frac{y}{r} \quad (7)$$

Also, in this plane setup, the coordinates of the vehicle's velocity vector can be calculated as follows, where v is the velocity (magnitude) of the vehicle.

$$v_x = v \cos \psi, v_y = v \sin \psi \quad (8)$$

In effect, the following approximate numerical integration equations over time can be used to simulate the vehicle kinematics:

$$x(t + \Delta t) = x(t) + v_x(t) \Delta t \quad (9)$$

$$y(t + \Delta t) = y(t) + v_y(t) \Delta t \quad (10)$$

$$\psi(t + \Delta t) = \psi(t) + \omega(t) \Delta t \quad (11)$$

In the above, t is time and Δt an appropriately small integration time step; omega (ω) stands for the rotational speed of the vehicle about an axis perpendicular to the xy-plane and with origin at the center of the vehicle.

Based on this framework, angle θ between the onboard radius vector pointing from the vehicle to the origin (target) and the vehicle's velocity vector can be calculated as follows (see Fig. 4).

$$\theta = \left(\frac{\pi}{2} - \phi \right) + \frac{\pi}{2} + \psi = \psi - \phi + \pi \quad (12)$$

Then a control law for a Supervisory Vehicle Controller, issuing forward speed and heading velocity setpoints, v_{set} and ω_{set} correspondingly can be introduced as follows.

$$v_{set} = \begin{cases} g(\theta) \left(\frac{r}{R} \right)^2 v_{max}, & 0 \leq r < R \\ g(\theta) v_{max}, & r \geq R > 0 \end{cases}, \omega_{set} = \frac{\omega_{max}}{2} f(\theta) \quad (13)$$

$$g(\theta) = 1 - \frac{|\theta|}{\pi}, f(\theta) = \text{sgn}(\theta)(\cos \theta - 1), \text{sgn}(\theta) = \begin{cases} +1, & \theta > 0 \\ 0, & \theta = 0 \\ -1, & \theta < 0 \end{cases} \quad (14)$$

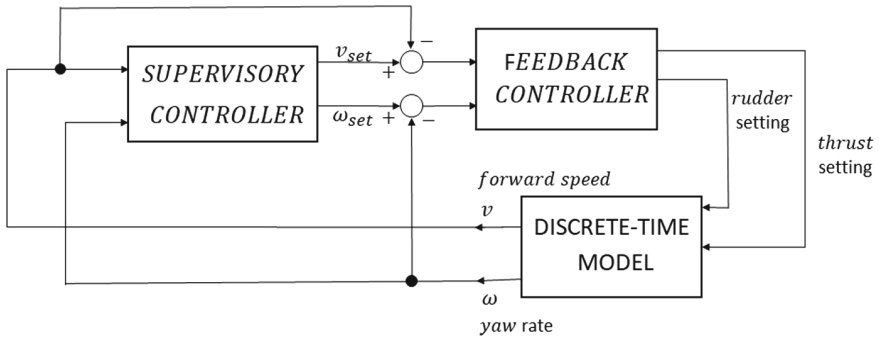


Fig. 3. Supervisory controller with discrete-time model

It is easy to see that $0 < v_{set} \leq v_{max}$ and $-\omega_{max} \leq \omega_{set} \leq \omega_{max}$; values v_{max} and ω_{max} represent the maximum forward and rotational speed values that the vehicle can attain. Distance value R is a user-defined range; this can be defined as e.g. twenty times the vehicle's length and in any case, it is sufficiently large so the vehicle can attain its maximum forward speed without the risk of overshooting the target.

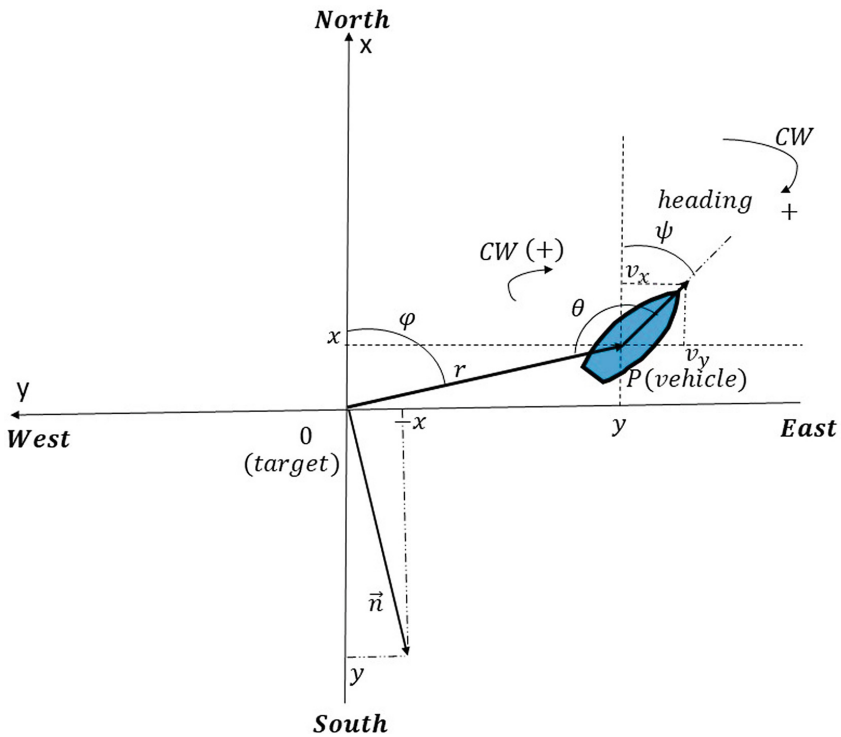


Fig. 4. Geometry of vehicle kinematics with the target point at the origin.

Since the control law is nonlinear, its performance can only be assessed through simulation, asymptotic analysis and qualitative arguments. Simulation results will be presented in the next section. For asymptotic analysis the following facts are noted.

If $\theta \approx \pm\pi$, then

$$v_{set} \approx 0, \omega_{set} \approx -\omega_{max} \text{sgn}(\theta) \quad (15)$$

If $\theta \approx \pm\pi/2$, then

$$v_{set} \approx \begin{cases} \left(\frac{r}{R}\right)^2 \frac{v_{max}}{2}, & 0 \leq r < R \\ \frac{v_{max}}{2}, & r \geq R \end{cases}, \omega_{set} \approx -\frac{\omega_{max}}{2} \text{sgn}(\theta) \quad (16)$$

If $\theta \approx \pm 0$, then

$$v_{set} \approx \begin{cases} \left(\frac{r}{R}\right)^2 v_{max}, & 0 \leq r < R \\ v_{max}, & r \geq R \end{cases}, \omega_{set} \approx \left(-\frac{\omega_{max}}{4}\right) |\theta| \theta \quad (17)$$

Given these observations, and assuming vehicle is initially at some arbitrary point and with arbitrary heading, it is first attempted to align the centerline axis with the onboard radius. Only then the vehicle is allowed to attain forward speed depending upon its distance from the target and relative to range R . As the vehicle gets closer to the target, forward speed decreases quadratically with distance, and corrections are applied so that the centerline remains aligned to the onboard radius.

3 Results and Discussion

The simulations based on psychomimetic control law were conducted in an attempt to model the behavior of a swarm of boats, using a methodology that had previously only been discussed with a single boat. To conduct the simulations, a total of 60 boats were selected, and the simulations themselves ran for a duration of 10 min. To begin the simulations, a set of initial conditions were established. These included the distance of each boat from its starting point, as well as the initial angles of both the φ and the heading angles of each boat. In addition, all boats were initially at rest, meaning that they had no forward speed at the beginning of the simulation. Throughout the course of the 10-min simulation, the movements and interactions of the boats with one another were tracked and analyzed. This allowed researchers to gain insights into how a swarm of boats might behave in a real-world scenario, and to identify any patterns or trends that could be useful in developing new strategies for navigating and controlling large groups of vessels.

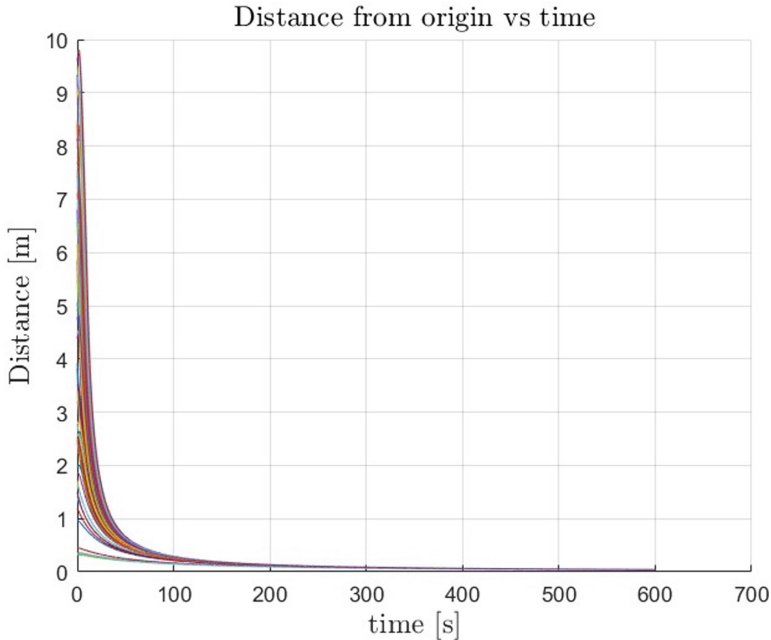


Fig. 5. Distance from origin with time for the boats' swarm

In Fig. 5, we can observe the distance that each of the 60 boats covered from their starting positions to their intended target. The initial conditions for the boats were chosen randomly. As evident from the figure, all 60 boats were able to reach the same target within a simulation period of 10 min.

Figure 6 displays the phase portraits of a group of boats acting as a swarm, with the θ angle as the primary axis. The figure shows the starting positions of the boats, which are determined by their distance from each other and the angle formed by that distance and their heading. The plot reveals that the phase portraits are spread out across a wide area, indicating a significant amount of dispersion among the boats.

Figure 7 is a visual representation of selected samples from the simulations, showcasing individual plots of six out of the total 60 samples. These individual plots provide a detailed depiction of the boat's movement during the simulation, including its forward speed and yaw rate, as well as important angles and trajectory. The plot effectively illustrates how each boat moves through the simulation, making it a valuable tool for analysis and understanding of the simulation results.

The results of our analysis indicate that when multiple boats are present, their trajectories do not intersect. This means that even though the boats may approach each other, they never cross paths and collide with one another. For example, when we examine a particular trajectory of a boat in the diagram, we can see that it converges smoothly with the trajectories of other boats without any collision occurring. To provide more information about the movement of boats, we can consider how a boat moves toward its destination. Initially, the boat starts from an arbitrary location and follows a curvy

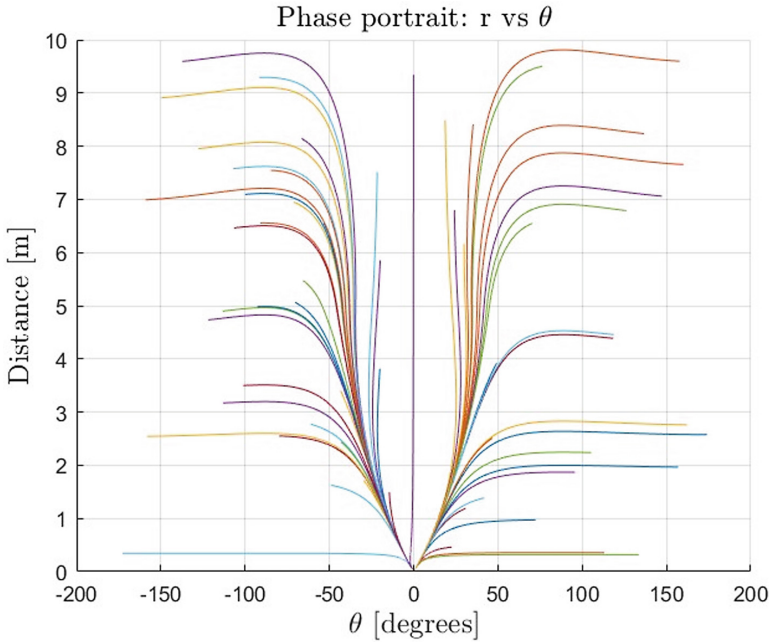


Fig. 6. Phase portraits of boats' swarm

path to reach its target. Our simulations reveal that the boat can move faster if it is fully aligned with and pointing toward the origin. This alignment ensures that the boat moves directly toward its intended target and does not waste energy moving away from it. On the other hand, if the boat is pointing away from the origin, it cannot move forward as efficiently because it has to use energy to move away from its intended target before it can start making progress toward it. As a result, the boat's progress is slower, and it may take longer to reach its destination.

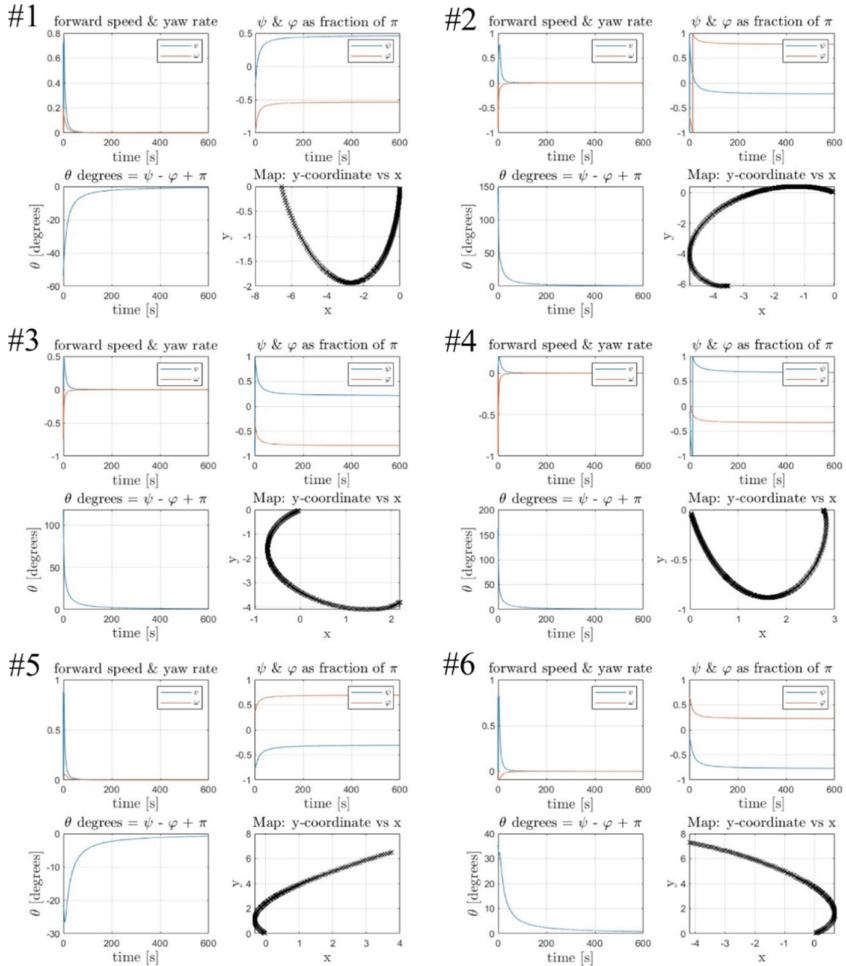


Fig. 7. 6 samples cases among all 60 cases

4 Conclusion

The present work presents research related to homing of an autonomous watercraft swarm using a distributed control law. A swarm of 60 boats is investigated in a simulation. Each vehicle is outfitted with a physicomimetic control law. The boats are assumed to have the same physical characteristics and responses, and their tracks are simplified to only consider motion and rotation around the yaw axis. The simulation investigates how the swarm of boats can be more energy-efficient through autonomous control. In future studies, the focus may shift to tuning local controllers and identifying boat dynamics in greater detail.

References

1. Chaumet-Lagrange M, Loeb H, Ygorra S. Design of an autonomous surface vehicle (ASV). In: Proceedings of OCEANS'94. 1994. p. I–120.
2. Manley JE. Development of the autonomous surface craft "aces". In: Oceans' 97 MTS/IEEE Conference Proceedings. 1997. p. 827–32.
3. DSOR. CARAVELA Development of a Long-Range Autonomous Oceanographic Vessel in Dynamic Systems and Ocean Robotics lab [Internet]. 1998. Available from: https://welcome.isr.tecnico.ulisboa.pt/projects_cat/dsor/
4. Oliveira P, Pascoal A, Rufino M, Sebastião L, Silvestre C. The DELFIM Autonomous Surface Craft. Rep (December 1999). 1999;
5. Oliveira P, Pascoal A, Kaminer I. A nonlinear vision based tracking system for coordinated control of marine vehicles. IFAC Proc Vol. 2002;35(1):295–300.
6. Leonessa A, Beaujean P-P, Driscoll F. Development of a small, multi-purpose, autonomous surface vessel. 2003.
7. Xiros NI, Logis E, Gasparis E, Tsolakidis S, Kardasis K. Theoretical and experimental investigation of unmanned boat electric propulsion system with PMDC motor and waterjet. J Mar Eng & Technol. 2009;8(2):27–43.
8. Xiros NI, Logis E, Charitos G. An automatic steering system for robust disturbance rejection. IASME Trans. 2004;1(2).
9. Fossen TI. Guidance and control of ocean vehicles. Univ Trondheim, Norway, Print by John Wiley & Sons, Chichester, England, ISBN 0 471 94113 1, Dr Thesis. 1999;



Distributed Supervisor Control for a Large-Scale Waterway Lock System

Fotis N. Koumboulis¹, Dimitrios G. Fragkoulis², Maria P. Tzamtzi¹ (✉),
and Panayotis G. Totomis²

¹ Department of Digital Industry Technologies, School of Science, National and Kapodistrian
University of Athens, Euripus Campus, 34400 Euboea, Greece

mtzamtzi@dind.uoa.gr

² Core Department, National and Kapodistrian University of Athens, Euripus Campus, 34400
Euboea, Greece

Abstract. In the present paper, a cyber-physical system, namely the WaterWay Lock (Lock III) in Tilburg, the Netherlands, will be studied. The discrete event model of the system, in the Ramadge Wonham framework, will be presented. The subsystems of the waterway lock system are the sensors and the actuators of the outgoing traffic lights, the incoming traffic lights, the gate paddles, the culverts, and the gates. The total number of the subsystems is fifty-two. The desired behavior of the systems will be expressed in the form of five rules and is formulated as one hundred and twelve supervisors derived by a general two-state supervisor form. This way, the implementation of the distributed control architecture is greatly facilitated.

Keywords: Cyber-physical systems · Supervisory control · Infrastructural systems

1 Introduction

Due to the advanced technologies adopted, modern systems are expected to be more and more effective [1, 2]. So, they are more difficult to be controlled and the control design becomes more complex. Modern systems including, among others, mechanical, electrical, and electronic systems interconnected through communication systems and coordinated/regulated by control units, with appropriate embedded control algorithms, are called, in general, cyber-physical systems (CPS). The term CPS covers, except large scale industrial and urban systems, traffic related systems [3].

The Ramadge-Wonham (RW) framework of the supervisory control theory uses the discrete event model of the system in order to design appropriate supervisors, being as possible less restrictive [4–7]. From the control point of view, the main idea of the RW framework is to disable appropriate controllable events, while guaranteeing that the system under control is nonblocking. This way, the designed supervisor architecture is implemented as a part of the CPS and is responsible for the coordination of all CPS components/devices [8–10]. Finally, the supervisory scheme is implemented either to

a central control unit managing all signals (sensors and actuators) or to local control devices, managing the signals of the components [11–17]. Usually in the central control unit a special designed software, e.g., Supervisory Control and Data Acquisition (SCADA), or manufacturing execution system (MES), is used [11, 18]. On the other hand, local control units [12–17] are usually Programmable Logic Controllers (PLC) or Remote Terminal Units (RTU).

In [18, 19], the WaterWay Lock (Lock III) in Tilburg, the Netherlands, is studied through the state diagrams of its subsystems. Also, in [18], a rule-based logic supervisory control scheme is proposed. Here, the cyber-physical system of this waterway, will also be studied. The subsystems of the waterway lock are the outgoing traffic lights, the incoming traffic lights, the paddles, the culverts, and the gates. Five behavioral rules, regarding the safe and unobstructed functionality of the system, are imposed. Here, the behavioral rules will be realized by appropriate supervisors, produced as specializations of a general two-state supervisor automaton. This characteristic facilitates the implementation of the proposed supervisory scheme [14]. The development of the parametric general two-state supervisor is the contribution of the paper. A distributed and modular control architecture in the Ramadge Wonham framework, using the automata of the supervisors will be implemented to control the total automaton of the system, namely the synchronous product of the automata of all subsystems. The resulting controlled automaton is physically realizable (see [20]) and nonblocking. The present control scheme, as compared to the one in [18], has the advantage of being fully systemic and satisfies desired system properties.

2 Model of the Water Way Lock

The Waterway Lock studied here is the one chamber lock located in Lock III in Tilburg, the Netherlands. As already mentioned in Sect. 1, Lock III has already been studied in [18]. The main subsystems of the Waterway Lock are the traffic lights, the gates, the paddles, and the culverts. The system has two types of traffic lights. The first includes two-lamp traffic lights and the second includes three-lamp traffic lights. The first type is for the outgoing traffic lights and the second type is for the incoming traffic lights. There are two incoming and two outgoing traffic lights at each entrance and exit, respectively. The two lamp traffic lights have one red and one green aspect, the red aspect indicates that the approximating vessel must wait, and the green aspect indicates that the approximating vessel can move forward. The three lamp traffic lights have two red aspects and one green aspect. The red aspect indicates that the vessel must wait. The green aspect indicates that the vessel can move forward. The red-green aspect (one red lamp and the green lamp are turned on) indicates that the vessel can move forward in a while and finally the red-red aspect (both red lamps are turned on and the green lamp is turned off) indicates that the lock is not working. The subsystem of the gates are watertight doors sealing the inside chamber of the lock from the water. The water from the chamber can be emptied through appropriate paddles installed in the gates at the downstream side. On the contrary, the chamber can be filled with water via appropriate culverts installed in the upstream side. Summing up, the system has four (4) Outgoing traffic lights, four (4) Incoming traffic lights, two (2) Paddles, two (2) Culverts and four (4) Gates. For each subsystem, except

the gates, there is installed one actuator. Each gate has three actuators (one for the direction, one for the pressure and one for the speed). Hence, overall, the system has twenty-four (24) actuator subsystems and twenty-eight (28) sensor subsystems.

Let $G_{O,i}$ be the automaton of the outgoing traffic lights actuator. Let ${}^rG_{O,i}$ and ${}^gG_{O,i}$ be the automata of the outgoing traffic lights sensors of the red and green aspects, respectively. Let $G_{I,i}$ be the automaton of the incoming traffic lights actuator. Let ${}^rG_{I,i}$, ${}^{rr}G_{I,i}$ and ${}^gG_{I,i}$ be the automata of the incoming traffic lights sensors of the red, red-red and green lamps, respectively. Let $G_{C,i}$ and ${}^sG_{C,i}$ be the automata of the actuator and the sensor of the culvert, respectively. Let $G_{P,i}$ and ${}^sG_{P,i}$ be the automata of the actuator and the sensor of the paddle, respectively. Finally, let $G_{D,i}$, $G_{T,i}$, $G_{A,i}$ and ${}^sG_{D,i}$ be the automata of the gate direction actuator, the gate pressure actuator, the gate speed actuator and the gate sensor, respectively. The range of the index i depends upon the automaton and will be specified later. In order to present the models of the automata, they will be grouped in to five different types, with respect to the number of their states and their initial state. The indices of the gates installed in the downstream of the Lock are 1 and 2, while the indices of the gates installed in the upstream of the Lock are 3 and 4.

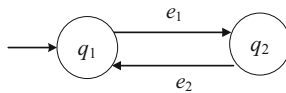


Fig. 1. State diagram of a two-state automaton

Table 1. Subsystems in the form of Fig. 1

Components of two-state state diagram			$q_j; j \in \{1, 2\}$	$e_j; j \in \{1, 2\}$
Subsystems	$i \in \{1, 2, 3, 4\}$	$G_{O,i}$	$q_{O,i,j}$ red/green aspect	$e_{O,i,j}$ commands to green/red aspect
		${}^rG_{O,i}$	${}^r q_{O,i,j}$ off/on red lamp	${}^r e_{O,i,j}$ signal on/off
		${}^gG_{O,i}$	${}^g q_{O,i,j}$ off/on green lamp	${}^g e_{O,i,j}$ signal on/off
		${}^rG_{I,i}$	${}^r q_{I,i,j}$ off/on red lamp	${}^r e_{I,i,j}$ signal on/off
		${}^{rr}G_{I,i}$	${}^{rr} q_{I,i,j}$ off/on red-red lamp	${}^{rr} e_{I,i,j}$ signal on/off
		${}^gG_{I,i}$	${}^g q_{I,i,j}$ off/on green lamp	${}^g e_{I,i,j}$ signal on/off
		$i \in \{1, 2\}$	$G_{C,i}$	$q_{C,i,j}$ closed/opened culvert
${}^sG_{C,i}$	${}^s q_{C,i,j}$ no flow/flow		${}^s e_{C,i,j}$ signal flow/no flow	

In Fig. 1, the state diagram of a general two-state automaton is presented, and in Table 1, the subsystems, having a two-state structure are presented. It is noted that in the state diagrams, the small vertical line in the begin of an arrow denotes that the event triggering the transition is controllable.

All events of $\mathbf{G}_{O,i}$ and $\mathbf{G}_{C,i}$ are controllable. All events of ${}^r\mathbf{G}_{O,i}$, ${}^s\mathbf{G}_{O,i}$, ${}^r\mathbf{G}_{I,i}$, ${}^{rr}\mathbf{G}_{I,i}$, ${}^s\mathbf{G}_{I,i}$ and ${}^s\mathbf{G}_{C,i}$ are uncontrollable. In $\mathbf{G}_{O,i}$, ${}^s\mathbf{G}_{O,i}$, ${}^s\mathbf{G}_{I,i}$, ${}^{rr}\mathbf{G}_{I,i}$, $\mathbf{G}_{C,i}$ and ${}^s\mathbf{G}_{C,i}$ the marked state is the 1st state, while in ${}^r\mathbf{G}_{O,i}$ and ${}^r\mathbf{G}_{I,i}$ the marked state is the 2nd state. In all automata, the initial state is the first state except ${}^r\mathbf{G}_{O,i}$ and ${}^r\mathbf{G}_{I,i}$, where the initial state is the second state.

In Fig. 2, the state diagram of a 3-state automaton with a “central” initial state is presented. In Table 2, the subsystems, having the structure of Fig. 2, are presented. It is mentioned that all events of these subsystems are controllable.

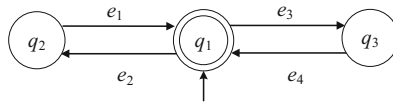


Fig. 2. State diagram of a 3-state automaton with a “central” initial state

Table 2. Subsystems in the form of Fig. 2

			Components of three-state diagram of type 1	
			$q_j; j \in \{1, 2, 3\}$	$e_j; j \in \{1, 2, 3, 4\}$
Subsystems	$i \in \{1, 2\}$	$\mathbf{G}_{P,i}$	$q_{P,i,j}$ idle /opens/ closes	$e_{P,i,j}$: commands to stop ($j \in \{1, 4\}$)/start ($j \in \{2, 3\}$) opening ($j \in \{1, 2\}$)/closing ($j \in \{3, 4\}$), respectively
	$i \in \{1, 2, 3, 4\}$	$\mathbf{G}_{D,i}$	$q_{D,i,j}$ idle /opens/ closes	$e_{D,i,j}$: commands to stop ($j \in \{1, 4\}$)/start ($j \in \{2, 3\}$) opening ($j \in \{1, 2\}$)/closing ($j \in \{3, 4\}$), respectively
		$\mathbf{G}_{T,i}$	$q_{T,i,j}$ idle / pressure high/ pressure low	$e_{T,i,j}$: command to deactivate ($j \in \{1, 4\}$)/activate ($j \in \{2, 3\}$) high pressure ($j \in \{1, 2\}$)/low pressure ($j \in \{3, 4\}$), respectively

In Fig. 3, the state diagram of a 3-state automaton with a “left” initial state is presented. In Table 3, the subsystems, having the structure of Fig. 3, are presented. It is mentioned that all events of these subsystems are uncontrollable.

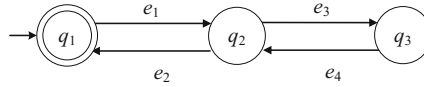


Fig. 3. State diagram of a three-state automaton with “left” initial state

Table 3. Subsystems in the form of Fig. 3

			Components of three-state diagram of type 2	
			$qj; j \in \{1, 2, 3\}$	$ej; j \in \{1, 2, 3, 4\}$
Subsystems	$i \in \{1, 2\}$	${}^S G_{P,i}$	${}^S q_{P,i,j}$ closed/ between/ opened	${}^S e_{P,i,j}$ signal not closed/ closed/ opened/ not opened
	$i \in \{1, 2, 3, 4\}$	${}^S G_{D,i}$	${}^S q_{D,i,j}$ closed/ between/ opened	${}^S e_{D,i,j}$ signal not closed/ closed/ opened/ not opened

The state diagram of $G_{I,i}$ is presented in Fig. 4, where $i \in \{1, 2, 3, 4\}$.

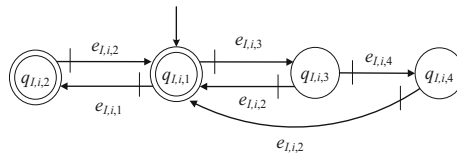


Fig. 4. State diagram of the actuator of the incoming traffic lights

The state $q_{I,i,1}$ is the red aspect, $q_{I,i,2}$ is the red-red aspect, $q_{I,i,3}$ is the red-green aspect and $q_{I,i,4}$ is the green aspect. The event $e_{I,i,1}$ is the command of the red-red aspect, $e_{I,i,2}$ is the command of the red aspect, $e_{I,i,3}$ is the command of the red-green aspect and $e_{I,i,4}$ is the command of the green aspect. All events of $G_{I,i}$ are controllable.

The state diagram of $G_{A,i}$ is presented in Fig. 5, where $i \in \{1, 2, 3, 4\}$.

The state $q_{A,i,1}$ is idle, $q_{A,i,2}$ opens in high speed, $q_{A,i,3}$ opens in low speed, $q_{A,i,4}$ closes in high speed and $q_{A,i,5}$ closes in low speed. The event $e_{A,i,1}$ is the command to open in high speed, $e_{A,i,2}$ is the command to stop opening in high speed, $e_{A,i,3}$ is the command to open in low speed, $e_{A,i,4}$ is the command to stop opening in low speed, $e_{A,i,5}$ is the command to close in high speed, $e_{A,i,6}$ is the command to stop closing in high speed, $e_{A,i,7}$ is the command to close in low speed, and $e_{A,i,8}$ is the command to stop closing in low speed. All events of $G_{A,i}$ are controllable.

The state diagrams of all subsystems are in accordance to those in [18].

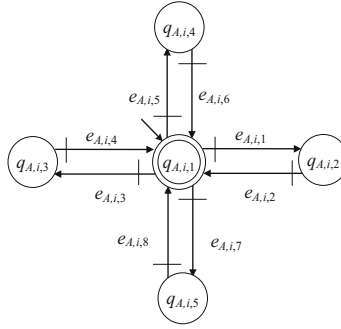


Fig. 5. State diagram of the actuator of the gate's speed

3 Supervisor Design

In order to secure the safe functionality of the waterway lock, a number of specification rules is imposed. The desired rules, introduced in [18], are the following:

1. A paddle at the downstream side of the Lock or a culvert at the upstream side of the Lock is not allowed to open if
 - a. any culvert at the opposite side or any paddle at the opposite side, respectively, is not closed,
 - b. a gate at the opposite side of the Lock is not closed.
2. A gate at the downstream side of the Lock or a gate at the upstream side of the Lock is not allowed to open if
 - a. a culvert at the upstream side of the Lock or a paddle at the downstream side of the Lock, respectively, is not closed,
 - b. A gate at the upstream side of the Lock or a gate at the downstream side of the Lock, respectively, is not closed.
3. A gate is not allowed to close if
 - a. an incoming traffic light does not display the red lamp on,
 - b. an outgoing traffic light does not display the red aspect.
4. An incoming traffic light is not allowed to display the green aspect if
 - a. an outgoing traffic light displays the green aspect,
 - b. the gates at the same side are not completely open.
5. An outgoing traffic light is not allowed to display the green aspect if
 - a. an incoming traffic light displays the green lamp on,
 - b. the gates at the same side are not completely open.

The general form of a supervisor automaton, expressed by a 6-tuple, (see [13, 14] and [20–23]) is $\mathbf{S} = (\mathbb{Q}_S, \mathbb{E}_S, f_S, \mathbb{H}_S, x_{S,0}, \mathbb{Q}_{S,m})$. The set of the states is $\mathbb{Q}_S = \{q_{S,1}, q_{S,2}\}$. The alphabet is $\mathbb{E}_S = \mathbb{E}_{S,1} \cup \mathbb{E}_{S,2} \cup \mathbb{E}_{S,3} \cup \mathbb{E}_{S,4}$. The initial state is $x_{S,0} = q_{S,1}$. The set of the marked states is $\mathbb{Q}_{S,m} = \mathbb{Q}_S$. The values of the transition functions are $f_S(q_{S,1}, e) = q_{S,1}, \forall e \in \mathbb{E}_{S,1}, f_S(q_{S,1}, e) = q_{S,2}, \forall e \in \mathbb{E}_{S,2}, f_S(q_{S,2}, e) = q_{S,2}, \forall e \in \mathbb{E}_{S,3}$ and $f_S(q_{S,2}, e) = q_{S,1}, \forall e \in \mathbb{E}_{S,4}$. The active event sets of \mathbf{S} are $\mathbb{H}_S(q_{S,1}) = \mathbb{E}_{S,1} \cup \mathbb{E}_{S,2}$ and $\mathbb{H}_S(q_{S,2}) = \mathbb{E}_{S,3} \cup \mathbb{E}_{S,4}$. The closed and the marked behavior of \mathbf{S} is $\mathbb{L}(\mathbf{S}) = \mathbb{L}_m(\mathbf{S}) = (\mathbb{E}_{S,1}^* \mathbb{E}_{S,2} \mathbb{E}_{S,3}^* \mathbb{E}_{S,4})^*$. In Fig. 6, the state diagram of \mathbf{S} is presented.

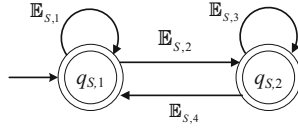


Fig. 6. State diagram of S

In what follows all the supervisors will be designed based on the supervisor automaton of Fig. 6.

The first rule is realized via four parametric supervisors. The first two supervisors restrict the actions of the paddles' actuators, and the last two supervisors restrict the actions of the culverts' actuators. The supervisors of the paddles' actuators are denoted by ${}^1S_P(i, j)$. The alphabets of the transitions of ${}^1S_P(i, j)$ are

$$\begin{aligned} {}^1\mathbb{E}_{S,P,1}(i, j) &= \begin{cases} \{e_{P,i,2}, {}^s e_{C,j,2}\}, & \text{if } i, j \in \{1, 2\} \\ \{e_{P,i,2}, {}^s e_{D,j,2}\}, & \text{if } (i \in \{1, 2\}) \wedge (j \in \{3, 4\}) \end{cases}, \\ {}^1\mathbb{E}_{S,P,2}(i, j) &= \begin{cases} \{{}^s e_{C,j,1}\}, & \text{if } i, j \in \{1, 2\} \\ \{{}^s e_{D,j,1}\}, & \text{if } (i \in \{1, 2\}) \wedge (j \in \{3, 4\}) \end{cases}, \\ {}^1\mathbb{E}_{S,P,3}(i, j) &= \begin{cases} \{{}^s e_{C,j,1}\}, & \text{if } i, j \in \{1, 2\} \\ \{{}^s e_{D,j,1}\}, & \text{if } (i \in \{1, 2\}) \wedge (j \in \{3, 4\}) \end{cases}, \\ {}^1\mathbb{E}_{S,P,4}(i, j) &= \begin{cases} \{{}^s e_{C,j,2}\}, & \text{if } i, j \in \{1, 2\} \\ \{{}^s e_{D,j,2}\}, & \text{if } (i \in \{1, 2\}) \wedge (j \in \{3, 4\}) \end{cases}. \end{aligned}$$

The alphabets of ${}^1S_P(i, j)$ are

$${}^1\mathbb{E}_{S,P}(i, j) = \begin{cases} \{e_{P,i,2}, {}^s e_{C,j,1}, {}^s e_{C,j,2}\}, & \text{if } i, j \in \{1, 2\} \\ \{e_{P,i,2}, {}^s e_{D,j,1}, {}^s e_{D,j,2}\}, & \text{if } (i \in \{1, 2\}) \wedge (j \in \{3, 4\}) \end{cases}$$

The supervisors of the culverts' actuators are denoted by ${}^1S_C(i, j)$. The alphabets of the transitions of ${}^1S_C(i, j)$ are

$$\begin{aligned} {}^1\mathbb{E}_{S,C,1}(i, j) &= \begin{cases} \{e_{C,i,1}, {}^s e_{P,j,2}\}, & \text{if } i, j \in \{1, 2\} \\ \{e_{C,i-2,1}, {}^s e_{D,j-2,2}\}, & \text{if } (i, j \in \{3, 4\}) \end{cases}, \\ {}^1\mathbb{E}_{S,C,2}(i, j) &= \begin{cases} \{{}^s e_{P,j,1}\}, & \text{if } i, j \in \{1, 2\} \\ \{{}^s e_{D,j-2,1}\}, & \text{if } (i, j \in \{3, 4\}) \end{cases}, \\ {}^1\mathbb{E}_{S,C,3}(i, j) &= \begin{cases} \{{}^s e_{P,j,1}\}, & \text{if } i, j \in \{1, 2\} \\ \{{}^s e_{D,j-2,1}\}, & \text{if } (i, j \in \{3, 4\}) \end{cases}, \\ {}^1\mathbb{E}_{S,C,4}(i, j) &= \begin{cases} \{{}^s e_{P,j,2}\}, & \text{if } i, j \in \{1, 2\} \\ \{{}^s e_{D,j-2,2}\}, & \text{if } (i, j \in \{3, 4\}) \end{cases}. \end{aligned}$$

The alphabets of ${}^1S_C(i, j)$ are

$${}^1\mathbb{E}_{S,C}(i, j) = \begin{cases} \{e_{C,i,1}, {}^s e_{P,j,1}, {}^s e_{P,j,2}\}, & \text{if } i, j \in \{1, 2\} \\ \{e_{C,i-2,1}, {}^s e_{D,j-2,1}, {}^s e_{D,j-2,2}\}, & \text{if } (i, j \in \{3, 4\}) \end{cases}$$

The second rule is realized via three parametric supervisors. The supervisors using paddles' sensors signals are denoted by ${}^2\mathbf{S}_P(i, j)$, ($i \in \{3, 4\}, j \in \{1, 2\}$). The alphabets of the transitions of ${}^2\mathbf{S}_P(i, j)$ are ${}^2\mathbb{E}_{S,P,1}(i, j) = \{e_{D,i,2}, {}^s e_{P,j,2}\}$, ${}^2\mathbb{E}_{S,P,2}(i, j) = \{{}^s e_{P,j,1}\}$, ${}^2\mathbb{E}_{S,P,3}(i, j) = \{{}^s e_{P,j,1}\}$ and ${}^2\mathbb{E}_{S,P,4}(i, j) = \{{}^s e_{P,j,2}\}$. The alphabets of ${}^2\mathbf{S}_P(i, j)$ are ${}^2\mathbb{E}_{S,P}(i, j) = \{e_{D,i,2}, {}^s e_{P,j,1}, {}^s e_{P,j,2}\}$. The supervisors using culverts' sensors signals are denoted by ${}^2\mathbf{S}_C(i, j)$. The alphabets of the transitions of ${}^2\mathbf{S}_C(i, j)$, ($i \in \{1, 2\}, j \in \{1, 2\}$) are ${}^2\mathbb{E}_{S,C,1}(i, j) = \{e_{D,i,2}, {}^s e_{C,j,2}\}$, ${}^2\mathbb{E}_{S,C,2}(i, j) = \{{}^s e_{C,j,1}\}$, ${}^2\mathbb{E}_{S,C,3}(i, j) = \{{}^s e_{C,j,1}\}$ and ${}^2\mathbb{E}_{S,C,4}(i, j) = \{{}^s e_{C,j,2}\}$. The alphabets of ${}^2\mathbf{S}_C(i, j)$ are ${}^2\mathbb{E}_{S,C}(i, j) = \{e_{D,i,2}, {}^s e_{C,j,1}, {}^s e_{C,j,2}\}$. The supervisors using gates' sensors signals are denoted by ${}^2\mathbf{S}_D(i, j)$, ($(i \in \{1, 2\}, j \in \{3, 4\}) \vee (i \in \{3, 4\}, j \in \{1, 2\})$). The alphabets of the transitions of ${}^2\mathbf{S}_D(i, j)$ are ${}^2\mathbb{E}_{S,D,1}(i, j) = \{e_{D,i,2}, {}^s e_{D,j,2}\}$, ${}^2\mathbb{E}_{S,D,2}(i, j) = \{{}^s e_{D,j,1}\}$, ${}^2\mathbb{E}_{S,D,3}(i, j) = \{{}^s e_{D,j,1}\}$ and ${}^2\mathbb{E}_{S,D,4}(i, j) = \{{}^s e_{D,j,2}\}$. The alphabets of ${}^2\mathbf{S}_D(i, j)$ are ${}^2\mathbb{E}_{S,D}(i, j) = \{e_{D,i,2}, {}^s e_{D,j,1}, {}^s e_{D,j,2}\}$.

The third rule is realized via two parametric supervisors. The supervisors using incoming traffic sensors signals are denoted by ${}^3\mathbf{S}_I(i, j)$, ($i, j \in \{1, 2, 3, 4\}$). The alphabets of the transitions of ${}^3\mathbf{S}_I(i, j)$ are ${}^3\mathbb{E}_{S,I,1}(i, j) = \{{}^r e_{I,i,1}, e_{D,j,3}\}$, ${}^3\mathbb{E}_{S,I,2}(i, j) = \{{}^r e_{I,i,2}\}$, ${}^3\mathbb{E}_{S,I,3}(i, j) = \{{}^r e_{I,i,2}\}$ and ${}^3\mathbb{E}_{S,I,4}(i, j) = \{{}^r e_{I,i,1}\}$. The alphabets of ${}^3\mathbf{S}_I(i, j)$ are ${}^3\mathbb{E}_{S,I}(i, j) = \{e_{D,j,3}, {}^r e_{I,i,1}, {}^r e_{I,i,2}\}$. The supervisors using outgoing traffic sensors signals are denoted by ${}^3\mathbf{S}_O(i, j)$, ($i, j \in \{1, 2, 3, 4\}$). The alphabets of the transitions of ${}^3\mathbf{S}_O(i, j)$ are ${}^3\mathbb{E}_{S,O,1}(i, j) = \{{}^r e_{O,i,1}, e_{D,j,3}\}$, ${}^3\mathbb{E}_{S,O,2}(i, j) = \{{}^r e_{O,i,2}\}$, ${}^3\mathbb{E}_{S,O,3}(i, j) = \{{}^r e_{O,i,2}\}$ and ${}^3\mathbb{E}_{S,O,4}(i, j) = \{{}^r e_{O,i,1}\}$. The alphabets of ${}^3\mathbf{S}_O(i, j)$ are ${}^3\mathbb{E}_{S,O}(i, j) = \{e_{D,j,3}, {}^r e_{O,i,1}, {}^r e_{O,i,2}\}$.

The fourth rule is realized via two parametric supervisors. The supervisors using outgoing traffic sensors signals are denoted by ${}^4\mathbf{S}_O(i, j)$, ($i, j \in \{1, 2, 3, 4\}$). The alphabets of the transitions of ${}^4\mathbf{S}_O(i, j)$ are ${}^4\mathbb{E}_{S,O,1}(i, j) = \{e_{I,i,4}, {}^s e_{O,j,2}\}$, ${}^4\mathbb{E}_{S,O,2}(i, j) = \{{}^s e_{O,j,1}\}$, ${}^4\mathbb{E}_{S,O,3}(i, j) = \{{}^s e_{O,j,1}\}$ and ${}^4\mathbb{E}_{S,O,4}(i, j) = \{{}^s e_{O,j,2}\}$. The alphabets of ${}^4\mathbf{S}_O(i, j)$ are ${}^4\mathbb{E}_{S,O}(i, j) = \{e_{I,i,4}, {}^s e_{O,i,1}, {}^s e_{O,i,2}\}$. The supervisors using gate sensors signals are denoted by ${}^4\mathbf{S}_D(i, j)$, ($(i, j \in \{1, 2\}) \vee (i, j \in \{3, 4\})$). The alphabets of the transitions of ${}^4\mathbf{S}_D(i, j)$ are ${}^4\mathbb{E}_{S,D,1}(i, j) = \{{}^s e_{D,j,4}\}$, ${}^4\mathbb{E}_{S,D,2}(i, j) = \{{}^s e_{D,j,3}\}$, ${}^4\mathbb{E}_{S,D,3}(i, j) = \{e_{I,i,4}, {}^s e_{D,j,3}\}$ and ${}^4\mathbb{E}_{S,D,4}(i, j) = \{{}^s e_{D,j,4}\}$. The alphabets of ${}^4\mathbf{S}_D(i, j)$ are ${}^4\mathbb{E}_{S,D}(i, j) = \{e_{I,i,4}, {}^s e_{D,j,3}, {}^s e_{D,j,4}\}$.

The fifth rule is realized via two parametric supervisors. The supervisors using incoming traffic sensors signals are denoted by ${}^5\mathbf{S}_I(i, j)$, ($i, j \in \{1, 2, 3, 4\}$). The alphabets of the transitions of ${}^5\mathbf{S}_I(i, j)$ are ${}^5\mathbb{E}_{S,I,1}(i, j) = \{e_{O,i,1}, {}^s e_{I,j,2}\}$, ${}^5\mathbb{E}_{S,I,2}(i, j) = \{{}^s e_{I,j,1}\}$, ${}^5\mathbb{E}_{S,I,3}(i, j) = \{{}^s e_{I,j,1}\}$ and ${}^5\mathbb{E}_{S,I,4}(i, j) = \{{}^s e_{I,j,2}\}$. The alphabets of ${}^5\mathbf{S}_I(i, j)$ are ${}^5\mathbb{E}_{S,I}(i, j) = \{e_{O,i,1}, {}^s e_{I,i,1}, {}^s e_{I,i,2}\}$. The supervisors using gate sensors signals are denoted by ${}^5\mathbf{S}_D(i, j)$, ($(i, j \in \{1, 2\}) \vee (i, j \in \{3, 4\})$). The alphabets of the transitions of ${}^5\mathbf{S}_D(i, j)$ are ${}^5\mathbb{E}_{S,D,1}(i, j) = \{{}^s e_{D,j,4}\}$, ${}^5\mathbb{E}_{S,D,2}(i, j) = \{{}^s e_{D,j,3}\}$, ${}^5\mathbb{E}_{S,D,3}(i, j) = \{e_{O,i,1}, {}^s e_{D,j,3}\}$ and ${}^5\mathbb{E}_{S,D,4}(i, j) = \{{}^s e_{D,j,4}\}$. The alphabets of ${}^5\mathbf{S}_D(i, j)$ are ${}^5\mathbb{E}_{S,D}(i, j) = \{e_{O,i,1}, {}^s e_{D,j,3}, {}^s e_{D,j,4}\}$.

4 Distributed Supervisory Control

The total system is the shuffle (synchronous product with disjoint alphabets) of the automata of all subsystems, i.e.,

$$\mathbf{G} = \prod_{\lambda=1}^2 \left(\mathbf{G}_{C,\lambda} \parallel \mathbf{G}_{P,\lambda} \parallel^s \mathbf{G}_{C,\lambda} \parallel^s \mathbf{G}_{P,\lambda} \right) \prod_{\lambda=1}^4 \left(\mathbf{G}_{O,\lambda} \parallel \mathbf{G}_{I,\lambda} \parallel \mathbf{G}_{D,\lambda} \parallel \mathbf{G}_{T,\lambda} \parallel \mathbf{G}_{A,\lambda} \parallel^r \mathbf{G}_{O,\lambda} \parallel^s \mathbf{G}_{O,\lambda} \parallel^r \mathbf{G}_{I,\lambda} \parallel^{rr} \mathbf{G}_{I,\lambda} \parallel^s \mathbf{G}_{I,\lambda} \parallel^s \mathbf{G}_{D,\lambda} \right). \quad (1)$$

From (1) it is observed that the total number of the automata of all subsystems is equal to 52, where the twenty-four (24) of them describe actuator subsystems and the twenty-eight (28) describe sensor subsystems.

The controlled automaton of the total system is the synchronous product of the total automaton in (1) and all supervisors developed in Sect. 3, i.e.,

$$\mathbf{G}_c = \mathbf{G} \parallel \left(\prod_{\lambda=1}^4 \left(\prod_{\mu=1}^4 \left({}^3\mathbf{S}_1(\lambda, \mu) \parallel {}^3\mathbf{S}_o(\lambda, \mu) \parallel {}^4\mathbf{S}_o(\lambda, \mu) \parallel {}^5\mathbf{S}_I(\lambda, \mu) \right) \right) \right) \parallel \left(\prod_{\lambda=1}^2 \left(\prod_{\mu=1}^2 \left(\prod_{\nu=4}^5 \left({}^\nu\mathbf{S}_D(\lambda, \mu) \parallel {}^\nu\mathbf{S}_D(\lambda + 2, \mu + 2) \right) \parallel {}^1\mathbf{S}_p(\lambda, \mu + 2) \parallel {}^1\mathbf{S}_p(\lambda, \mu) \parallel {}^1\mathbf{S}_C(\lambda, \mu) \parallel {}^1\mathbf{S}_C(\lambda + 2, \mu + 2) \parallel {}^2\mathbf{S}_C(\lambda, \mu) \parallel {}^2\mathbf{S}_p(\lambda + 2, \mu) \parallel {}^2\mathbf{S}_D(\lambda, \mu + 2) \parallel {}^2\mathbf{S}_D(\lambda + 2, \mu) \right) \right) \right) \quad (2)$$

It is important to mention that $\mathbf{G}_{T,\lambda}$ and $\mathbf{G}_{A,\lambda}$ have disjoint alphabets with all supervisor automata. Thus, they are not restricted by the supervisors. From (2) it is observed that the total number of supervisors realized in the system is equal to 112. From the alphabets of all supervisor automata in (2), the following property is derived. $({}^y\mathbb{E}_{S,x}(i, j) \cap {}^\zeta\mathbb{E}_{S,\chi}(i, \xi)) \subseteq \mathbb{E}_{uc}, \forall x, y, i, j, \zeta, \chi, i, \xi; y \neq \zeta,$

where \mathbb{E}_{uc} is the union of all the uncontrollable event sets of all the subsystems and where $x, \chi \in \{C, P, O, I, D\}$ and $y, \zeta \in \{1, 2, 3, 4, 5\}$. The last two integers are the numbers of the respective rules. In other words, the common event in the alphabets of the supervisors of the different rules are only uncontrollable events. Also, regarding the supervisors, it is observed that $|{}^y\mathbb{E}_{S,x}(i, j) \cap \mathbb{E}_c| = 1 \forall y, x, i, j$, where \mathbb{E}_c is the union of all controllable event sets of all subsystems. In other words, in each supervisor alphabet there is only one controllable event, or equivalently there is only one actuator event as all sensors' events are uncontrollable. So, a set of supervisors restrict the action of only one actuator automaton, in the sense that each supervisor, in this set, restricts the action of this actuator automaton. Thus, the supervisors, belonging to this set, compose a control section controlling only this actuator automaton. Hence, it is concluded that the sensors' events are used only for monitoring purposes and the present supervisory scheme is a distributed control architecture analyzed to control actions per actuator's automata. The following two properties are satisfied: (a) All supervisors are physically realizable (PR) regarding \mathbf{G} through (2), (b) The automaton \mathbf{G}_c is a nonblocking automaton.

The satisfactory performance of the controlled automaton \mathbf{G}_c , is verified by the marked behavior of the controlled automaton \mathbf{G}_c , being the following

$$\begin{aligned} \mathbb{L}_m(\mathbf{G}_c) = & \mathbb{L}_m(\mathbf{G}) \cap \bigcap_{\lambda=1}^4 \left(\bigcap_{\mu=1}^4 \left(\mathbb{L}_m({}^3\mathbf{S}_I(\lambda, \mu)) \right) \cap \mathbb{L}_m({}^3\mathbf{S}_O(\lambda, \mu)) \right) \cap \\ & \cap \mathbb{L}_m({}^4\mathbf{S}_O(\lambda, \mu)) \cap \mathbb{L}_m({}^5\mathbf{S}_I(\lambda, \mu)) \cap \\ & \bigcap_{\lambda=1}^2 \left(\bigcap_{\mu=1}^2 \left(\bigcap_{v=4}^5 \left(\mathbb{L}_m({}^v\mathbf{S}_D(\lambda, \mu)) \right) \cap \mathbb{L}_m({}^v\mathbf{S}_D(\lambda + 2, \mu + 2)) \right) \right) \cap \\ & \cap \mathbb{L}_m({}^1\mathbf{S}_P(\lambda, \mu + 2)) \cap \mathbb{L}_m({}^1\mathbf{S}_C(\lambda + 2, \mu + 2)) \cap \\ & \cap \mathbb{L}_m({}^2\mathbf{S}_P(\lambda + 2, \mu)) \cap \mathbb{L}_m({}^2\mathbf{S}_C(\lambda, \mu)) \cap \\ & \cap \mathbb{L}_m({}^2\mathbf{S}_D(\lambda, \mu + 2)) \cap \mathbb{L}_m({}^2\mathbf{S}_D(\lambda + 2, \mu)) \cap \\ & \cap \bigcap_{\mu=1}^2 \left(\mathbb{L}_m({}^1\mathbf{S}_P(\lambda, \mu)) \right) \cap \mathbb{L}_m({}^1\mathbf{S}_C(\lambda, \mu)). \end{aligned}$$

where ${}^y_{ij}P_x$ are the projections of \mathbb{E}^* to $({}^y\mathbb{E}_{S,x}(i, j))^*$ and $x \in \{C, P, O, I, D\}$ and $y \in \{1, 2, 3, 4, 5\}$. \mathbb{E} is the union of the alphabets of all subsystems.

5 Conclusion

In the present paper, using discrete event system tools, the model of the Waterway Lock, Lock III in Tilburg, the Netherlands, has been presented. The system has 52 subsystems (sensors and actuators). The desired behavior of this large-scale system has been imposed by 5 rules that have been formulated to a set of 112 supervisor automata resulting as special cases of a two-state general supervisor automaton. This characteristic facilitates the implementation of the supervisor scheme through function blocks. It has been shown that the present supervisor scheme is a distributed supervisory control architecture. Finally, the marked behavior of the controlled automaton has analytically been presented.

Acknowledgement. Research supported by the MSc Program in “Advanced Control Systems and Robotics”, National and Kapodistrian University of Athens, Greece.

References

1. Yang, F., Gu, S.: Industry 4.0, a revolution that requires technology and national strategies, *Complex Intelligent Systems*, 7, 1311–1325 (2021).
2. Jiang, Y, Yin, S., Kaynak, O.: Data-Driven Monitoring and Safety Control of Industrial Cyber-Physical Systems: Basics and Beyond, in *IEEE Access*, vol. 6, pp. 47374–47384 (2018).
3. Serpanos, D.: The Cyber-Physical Systems Revolution, in *Computer*, vol. 51, no. 3, pp. 70-73 (2018).
4. Reijnen, F. F. H., Goorden, M. A., Van de Mortel-Fronczak, J. M., Rooda, J. E.: Modeling for supervisor synthesis – a lock-bridge combination case study, *Discrete Event Dynamic Systems*, volume 30, pages 499–532 (2020).
5. Ramadge, P. J., Wonham, W. M.: Supervisory control of a class of discrete event processes, *SIAM journal on control and optimization*, 25(1), 206–230 (1987).
6. Wonham, W. M. and Kai, C.: Supervisory control of discrete-event systems, Springer International Publishing (2019).
7. Cassandras C. G., Lafortune S.: Introduction to Discrete Event Systems, 3rd ed., Springer: Cham, Switzerland (2021).
8. Vahidi, A., Fabian, M., Lennartson, B.: Efficient supervisory synthesis of large systems, *Control Engineering Practice*, vol. 14, no. 10, pp. 1157–1167 (2006).
9. Cai, K., Wonham, W. M.: Supervisor Localization: A Top-Down Approach to Distributed Control of Discrete-Event Systems, in *IEEE Transactions on Automatic Control*, vol. 55, no. 3, pp. 605-618, (2010).
10. Moormann, L., van de Mortel-Fronczak, J. M., Fokkink, W. J., Maessen, P., Rooda, J. E.: Supervisory control synthesis for large-scale systems with isomorphism, *Control Engineering Practice*, 115, 104902 (2021).
11. Beatriz Portilla, N., de Queiroz, M. H., Cury, J. E.: Integration of supervisory control with SCADA system for a flexible manufacturing cell, *2014 12th IEEE International Conference on Industrial Informatics (INDIN)*, Porto Alegre, Brazil, pp. 261–266 (2014)
12. Reijnen, F. F., Goorden, M. A., van de Mortel-Fronczak, J. M., Rooda, J. E.: Modeling for supervisor synthesis—a lock-bridge combination case study, *Discrete Event Dynamic Systems*, vol 30, pp.499-532 (2020).
13. Kouvakas, N.D., Koumboulis, F.N., Fragkoulis, D.G., Souliotis, A.: Modular Supervisory Control for the Coordination of a Manufacturing Cell with Observable Faults, *Sensors*, 23, 163 (2023).
14. Koumboulis, F. N., Fragkoulis, D. G., Kalkanis, I., Fragulis, G. F.: Supervisor Design for a Pressurized Reactor Unit in the Presence of Sensor and Actuator Faults, *Electronics*, 11(16), 2534, (2022).
15. Scotti, W. A. F., Portilla, N. B., de Queiroz, M. H., Cury, J. E. R., Modular and systematic design of supervisory control system integrating PLC, SCADA and task routing for a modular production system, *2015 IEEE International Conference on Automation Science and Engineering (CASE)*, Gothenburg, Sweden, pp. 807–812 (2015).
16. Reijnen, F. F. H., Reniers, M. A., Van de Mortel-Fronczak, J. M., Rooda, J. E.: Structured Synthesis of Fault-Tolerant Supervisory control, *IFAC PapersOnLine* 51-24, pp. 894–901 (2018).
17. Fokkink, W., Goorden, M., van de Mortel-Fronczak, J., Reijnen, F., Rooda, J.: Supervisor Synthesis: Bridging Theory and Practice, in *Computer*, vol. 55, no. 10, pp. 48-54 (2022).
18. Reijnen, F. F. H., Goorden, M. A., Van de Mortel-Fronczak, J. M., Rooda, J. E.: Supervisory control synthesis for a waterway lock, *IEEE Conference on Control Technology and Applications (CCTA)*, August 27–30, Kohala Coast, Hawaii, USA (2017).

19. Goorden, M., van de Mortel-Fronczak, J., Reniers, M., Fokkink, W., Rooda, J.: The Impact of Requirement Splitting on the Efficiency of Supervisory Control Synthesis. In: Larsen, K., Willems, T. (eds) *Formal Methods for Industrial Critical Systems. FMICS 2019. Lecture Notes in Computer Science*, vol 11687. Springer, Cham (2019).
20. Koumboulis, F. N., Fragkoulis, D. G., Arapakis, S.: Supervisor design for an assembly line in the presence of faults, 27th IEEE International Conference on Emerging Technology and Factory Automation, Stuttgart, Germany, September 6–9 (2022).
21. Koumboulis, F. N., Fragkoulis, D. G., Menexis, A. N.: Supervisory Control for Flexibility of Production Manufacturing Processes, IEEE 21st International Conference on Intelligent Engineering Systems 2017 (INES 2017), Larnaca, October 20–23, Cyprus (2017).
22. Koumboulis, F. N., Fragkoulis, D. G., Ioannou, K. A.: Control of Router Nodes in Production Manufacturing Processes, International Conference on Modern Circuits and Systems Technologies (MOCAS), May 7–9, Thessaloniki, Greece (2018).
23. Koumboulis, F. N., Fragkoulis, D. G., Diveris, G. K.: Function Supervisors for Storage Systems, International Conference on Modern Circuits and Systems Technologies (MOCAS), May 7–9, Thessaloniki, Greece (2018).



Supervisory Control of Multiple Product Flow for a Flexible Manufacturing System

Fotis N. Koumboulis¹, Dimitrios G. Fragkoulis², Maria P. Tzamtzi¹ (✉),
and Dimitrios Papadopoulos²

¹ Department of Digital Industry Technologies, School of Science, National and Kapodistrian University of Athens, Euripus Campus, 34400 Euboea, Greece

mtzamtzi@dind.uoa.gr

² Core Department, National and Kapodistrian University of Athens, Euripus Campus, 34400 Euboea, Greece

Abstract. A general parametric modeling and control framework of multi-product manufacturing systems is proposed. The manufacturing system is of a general and parametric form covering a wide range of manufacturing architectures. The system consists of a parametric number of manufacturing machines, a parametric number of product transfer systems, a parametric number of buffers, and a parametric set of input and output product stacks. The model of each machine and transfer system is provided in the form of DES. Behavioral rules are translated to regular languages. For each language, a two-state supervisor automaton will be designed.

Keywords: Discrete event systems · Supervisory control · Flexible manufacturing systems

1 Introduction

In the framework of Supervisory Control Theory (SCT) of Discrete Event Systems (DES) (see [3, 4]), a common method to control manufacturing systems is the use of interfaces among devices and components of the system, limiting their interaction, towards providing local verification of global properties (e.g., see [5–7]). The hierarchical interfaces (see [1, 2] and [5–9]) as well as abstractions of the mode of the manufacturing system (see [10, 11]), are widely applied. All SCT results, mentioned above, have been derived for manufacturing systems with a prespecified number of product processing units and a prespecified number of transfer units, as well as a prespecified number of buffers (see [1, 12, 13] and the references within). The approaches in the above results do not face the general manufacturing system problem, being parametric with respect to the number of treated products and the participating devices/machines.

In the present paper, a general multi-product manufacturing system will be considered. The system consists of a parametric set of manufacturing machines, i.e., Milling Machines, Painting Machines, Lathes, and Assembly Machines; a parametric set of product transfer systems, i.e., Robotic Manipulators, and Conveyor Belts; a parametric set of

single slot buffers, and a parametric set of input and output product stacks. The models of all above manufacturing and transfer subsystems have been presented in the simple DES forms, given in [1]. Here, the DES models are in the form of parametric 6-tuples (see [14–17]). Using these subsystems' models, a general manufacturing system of any architecture is introduced. This is the first contribution of the paper. A supervisory control scheme, based on specific behavioral rules, is developed after the translation of the desired behavioral rules to a set of parametric regular languages. Then, the supervisor automata of the above parametric languages are derived by developing a general and parametric supervisor form, being suitable for PLC implementation using parametrized library elements and Function Blocks (see [18]). This is the second contribution of the paper.

In Sect. 2, DES models of the machines, the product transfer missions, and the respective product transfer systems are presented. In Sect. 3, a set of local supervisors restricting the product transfer systems to their mission are proposed. In Sect. 4, the desired behavior of the system is expressed in the form of a set of regular languages. In Sect. 5, general parametric supervisor automata have been developed to realize all regular languages, via appropriate specializations.

2 Modelling of the Subsystems of the Parametric Manufacturing System

In Table 1 the nomenclature of the present parametric model is presented.

Remark 1 It is important to mention that all buffers have only one product slot. Also, it holds that $\mathbb{V}_{I,T}(\eta) \cup \mathbb{V}_{I,M}(\eta) \neq \emptyset$, which means that each buffer is the physical entity playing the role of the product input spot for at least one transfer mission or machine with no cyclic product flow. Moreover, it holds that $\mathbb{V}_{O,T}(\eta) \cup \mathbb{V}_{O,M}(\eta) \neq \emptyset$, which means that each buffer is the physical entity playing the role of the product output spot for at least one machine or transfer mission. Finally, the sets $\mathbb{V}_{I,M}(\eta)$, $\mathbb{V}_{O,M}(\eta)$, $\mathbb{V}_{I,T}(\eta)$ and $\mathbb{V}_{O,T}(\eta)$ contain only the indices of the machines or transfer missions associated with the respective buffer V_η . The input and output product stacks are considered to have enough many slots. Additionally, the input product stacks always have a product available to feed, while the output product stacks always have the required space to receive a product.

2.1 The Models of the Manufacturing Machines

2.1.1 A General Model for All Machines

The general DES model of M_i in 6-tuple form (see [1, 12] and [13]) is defined to be $\mathbf{G}_{M,i} = (\mathbb{Q}_{M,i}, \mathbb{E}_{M,i}, f_{M,i}, \mathbb{H}_{M,i}, x_{M,i,0}, \mathbb{Q}_{M,i,m})$. The set of the states is

$$\mathbb{Q}_{M,i} = \begin{cases} \{q_{M,i,1}, q_{M,i,2}\}, & \text{if } i \in \mathbb{I}_M \cup \mathbb{I}_P \\ \{q_{M,i,1}, q_{M,i,2}, q_{M,i,3}\}, & \text{if } i \in \mathbb{I}_A \\ \bigcup_{\lambda=1}^{n_i+1} \{q_{M,i,\lambda}\}, & \text{if } i \in \mathbb{I}_L \end{cases}$$

Table 1. Nomenclature of the parametric manufacturing system

Symbol	Definition
n	Number of types of products
m	Number of manufacturing machines
ρ	Number of product transfer systems
μ	Number of buffers
ℓ	Number of product transfer missions
θ	Number of input product stacks
ζ	Number of output product stacks
i	Index of the manufacturing machines
y	Index of the product transfer systems
η	Index of the buffers
r	Index of the input product stack
o	Index of the output product stack
j	Index of the product transfer missions
M_i	Manufacturing machine
W_y	Product transfer system
V_η	Single slot buffer
T_j	Product transfer mission
n_i	Number of the types of products manufactured by M_i
\mathbb{I}	Set of indices of the machines
$\mathbb{I}_M, \mathbb{I}_P, \mathbb{I}_L, \mathbb{I}_A$	Set of the indices of the milling machines, painting machines, lathes and assembly machines, respectively
\mathbb{Y}	Set of the indices of the product transfer systems
$\mathbb{Y}_B, \mathbb{Y}_R$	Set of the indices of the conveyor belts and robotic manipulators, respectively
${}^{i,k}B_{I,M}$	Product input spot of M_i , being an input product stack or a buffer
${}^iB_{O,M}$	Product output spot of M_i , being an output product stack or a buffer
$\mathbb{V}_{I,M}(\eta)$	Set of the pairs of indices (i, k) of the ${}^{i,k}B_{I,M}$ implemented by V_η
$\mathbb{V}_{O,M}(\eta)$	Set of the indices i of the machines outputting to V_η
$\mathbb{W}(y)$	Set of the indices of the product transfer missions of W_y
${}^jB_{I,T}$	Product input spot of T_j , being an input product stack or a buffer
${}^jB_{O,T}$	Product output spot of T_j , being an output product stack or a buffer

(continued)

Table 1. (continued)

Symbol	Definition
$\mathbb{V}_{I,T}(\eta)$	Set of indices j of all the transfer missions with product input spots ${}^jB_{I,T}$, corresponding to the physical entity V_η
$\mathbb{V}_{O,T}(\eta)$	Set of indices j of all the transfer missions with product output spots ${}^jB_{O,T}$, corresponding to the physical entity V_η

where n_i is the number of different products that the i -th machine manufactures in the case where the i -th machine is a lathe or an assembly machine. The initial state is $x_{M,i,0} = q_{M,i,1}$. The set of the marked states is $\mathbb{Q}_{M,i,m} = \{q_{M,i,1}\}$. The alphabet is

$$\mathbb{E}_{M,i} = \begin{cases} \{e_{M,i,1}, e_{M,i,2}\}, & \text{if } i \in \mathbb{I}_M \cup \mathbb{I}_P \\ \bigcup_{\lambda=1}^{n_i+2} \{e_{M,i,\lambda}\}, & \text{if } i \in \mathbb{I}_A \\ \bigcup_{\lambda=1}^{2n_i} \{e_{M,i,\lambda}\}, & \text{if } i \in \mathbb{I}_L \end{cases}$$

In Fig. 1, the state diagram of $G_{M,i}$, where $i \in \mathbb{I}_M \cup \mathbb{I}_P$, is depicted. In Fig. 2, the state diagram of $G_{M,i}$, where $i \in \mathbb{I}_A$, is depicted. The state diagram of $G_{M,i}$, where $i \in \mathbb{I}_L$, is depicted in Fig. 3. All state diagrams are in accordance to those in the special case of manufacturing machines presented in [1, 12] and [13].

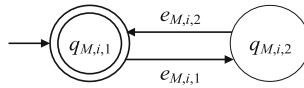


Fig. 1. State diagram of $G_{M,i}$ for $i \in \mathbb{I}_M \cup \mathbb{I}_P$

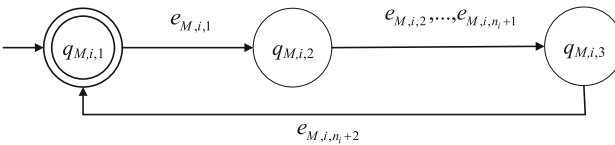


Fig. 2. State diagram of $G_{M,i}$ for $i \in \mathbb{I}_A$

2.1.2 States and Events of a Milling Machine

If the i -th machine is a milling machine, it holds that $i \in \mathbb{I}_M$. Milling machines have two states. The first, $q_{M,i,1}$, is the case where the milling machine is in idle state. The second, $q_{M,i,2}$, is the case where the milling machine manufactures a product. The alphabet of the milling machine is $\mathbb{E}_{M,i} = \{e_{M,i,1}, e_{M,i,2}\}$. The event $e_{M,i,1}$ is the command to load a product from the product input spot to the milling machine and start manufacturing. The event $e_{M,i,2}$ is the signal indicating that the manufacturing of the product has been

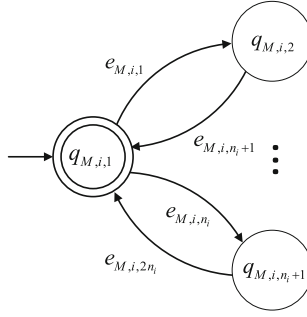


Fig. 3. State diagram of $G_{M,i}$ for $i \in \mathbb{I}_L$

completed and the product is loaded to the product output spot. The controllable event set is $\mathbb{E}_{M,i,c} = \{e_{M,i,1}\}$ and the uncontrollable event set is $\mathbb{E}_{M,i,uc} = \{e_{M,i,2}\}$.

2.1.3 States and Events of a Painting Machine

If the i -th machine is a painting machine, it holds that $i \in \mathbb{I}_P$. Painting machines have two states. The first, $q_{M,i,1}$, is the case where the painting machine is in idle state. The second, $q_{M,i,2}$, is the case where the painting machine paints a product. The alphabet of the painting machine is $\mathbb{E}_{M,i} = \{e_{M,i,1}, e_{M,i,2}\}$. The event $e_{M,i,1}$ is the command to load a product from the product input spot to the painting machine and start painting. The event $e_{M,i,2}$ is the signal indicating that the painting of the product has been completed and the product is loaded to the product output spot. The controllable event set is $\mathbb{E}_{M,i,c} = \{e_{M,i,1}\}$ and the uncontrollable event set is $\mathbb{E}_{M,i,uc} = \{e_{M,i,2}\}$.

2.1.4 States and Events of a Lathe

If the i -th machine is a lathe, it holds that $i \in \mathbb{I}_L$. Lathes have $n_i + 1$ states. The first, $q_{M,i,1}$, is the case where the lathe is in idle state. Each of the next n_i states, $q_{M,i,\lambda+1}$, where $\lambda \in \{1, \dots, n_i\}$, is the case where the lathe manufactures the product of type λ (manufacturing mode λ). The alphabet of the lathe is $\mathbb{E}_{M,i} = \bigcup_{\lambda=1}^{n_i} \{e_{M,i,\lambda}\}$. The event $e_{M,i,\lambda}$, where $\lambda \in \{1, \dots, n_i\}$, is the command to load a product of type λ from the product input spot ${}^{i,\lambda}B_{I,M}$ to the lathe and start the appropriate manufacturing. The event $e_{M,i,n_i+\lambda}$ is the signal indicating that the manufacturing of the product of type λ has been completed and the product is loaded to the product output spot. The controllable event set is $\mathbb{E}_{M,i,c} = \bigcup_{\lambda=1}^{n_i} \{e_{M,i,\lambda}\}$ and the uncontrollable event set is $\mathbb{E}_{M,i,uc} = \bigcup_{\lambda=n_i+1}^{2n_i} \{e_{M,i,\lambda}\}$.

2.1.5 States and Events of an Assembly Machine

If the i -th machine is an assembly machine, it holds that $i \in \mathbb{I}_A$. Assembly Machines have 3 states. The first, $q_{M,i,1}$, is the case where the assembly machine is in idle state. The second, $q_{M,i,2}$, is the case where the assembly machine has been initialized. The third, $q_{M,i,3}$, is the case where the assembly machine assembles a product. The alphabet

of the assembly machine is $\mathbb{E}_{M,i} = \bigcup_{\lambda=1}^{n_i+2} \{e_{M,i,\lambda}\}$. The event $e_{M,i,1}$ is the command to initialize the assembly machine. The command $e_{M,i,\lambda+1}$, where $\lambda \in \{1, \dots, n_i\}$, is the command to load a product of type λ from the product input spot ${}^{i,\lambda}B_{I,M}$ of the machine to the assembly machine and start the assembly of the product. The event e_{M,i,n_i+2} is the signal indicating that the assembly of a product, of any type, has been completed and the product is loaded to the product output spot. The controllable event set is $\mathbb{E}_{M,i,c} = \bigcup_{\lambda=1}^{n_i+1} \{e_{M,i,\lambda}\}$ and the uncontrollable event set is $\mathbb{E}_{M,i,uc} = \{e_{M,i,n_i+2}\}$.

2.2 The Models for Product Transfer

2.2.1 A General Model of All Transfer Missions

The general DES model of T_j is $\mathbf{R}_{T,j} = (\mathbb{Q}_{T,j}, \mathbb{E}_{T,j}, f_{T,j}, \mathbb{H}_{T,j}, x_{T,j,0}, \mathbb{Q}_{T,j,m})$. The set of the states is $\mathbb{Q}_{T,j} = \{q_{T,j,1}, q_{T,j,2}\}$, where $q_{T,j,1}$ is the stage of the mission where no product is carried out, and $q_{T,j,2}$ is the stage of the mission where a product is carried out. The initial state is $x_{T,j,0} = q_{T,j,1}$. The set of the marked states is $\mathbb{Q}_{T,j,m} = \{q_{T,j,1}\}$. The alphabet of the automaton is $\mathbb{E}_{T,j} = \{e_{T,j,1}, e_{T,j,2}\} \cup \tilde{\mathbb{E}}_j$, where $e_{T,j,1}$ is the command to load a product from a transfer product input spot to the transfer system, executing the j -th transfer mission, and $e_{T,j,2}$ is the command to unload a product from the transfer system executing the j -th transfer mission, to the respective product output spot. The event set $\tilde{\mathbb{E}}_j$ contains all commands of the respective product transfer system that do not trigger transitions in the automaton $\mathbf{R}_{T,j}$, i.e., they do not belong to the active event sets of $\mathbf{R}_{T,j}$. The controllable events set is $\mathbb{E}_{T,j,c} = \mathbb{E}_{T,j}$ and the uncontrollable event set is $\mathbb{E}_{T,j,uc} = \emptyset$.

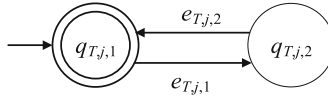


Fig. 4. State diagram of $\mathbf{R}_{T,j}$

2.2.2 The General Model of the Product Transfer Systems

The general DES model of W_y is $\mathbf{G}_{W,y} = (\mathbb{Q}_{W,y}, \mathbb{E}_{W,y}, f_{W,y}, \mathbb{H}_{W,y}, x_{W,y,0}, \mathbb{Q}_{W,y,m})$. The set of the states is $\mathbb{Q}_{W,y} = \{q_{W,y,1}, q_{W,y,2}\}$, where $q_{W,y,1}$ denotes the case where the product transfer system is empty, and $q_{W,y,2}$ denotes the case where the product transfer system carries a product. The initial state of the automaton is $x_{W,y,0} = q_{W,y,1}$. The set of the marked states is $\mathbb{Q}_{W,y,m} = \{q_{W,y,1}\}$. The alphabet is $\mathbb{E}_{W,y} = \left(\bigcup_{\lambda \in \mathbb{W}(y)} \{e_{T,\lambda,1}, e_{T,\lambda,2}\} \right) \cup \tilde{\mathbb{E}}_{I,y} \cup \tilde{\mathbb{E}}_{O,y}$, where $\tilde{\mathbb{E}}_{I,y}$ contains all commands to load a product that do not trigger transition to at least one of its missions, and where $\tilde{\mathbb{E}}_{O,y}$ contains all commands to unload products that do not trigger transition to at least one of its missions. Clearly, $\tilde{\mathbb{E}}_j = \tilde{\mathbb{E}}_{I,y} \cup \tilde{\mathbb{E}}_{O,y}; \forall j \in \mathbb{W}(y)$. The controllable event set is $\mathbb{E}_{W,y,c} = \mathbb{E}_{W,y}$ and the uncontrollable event set is

$\mathbb{E}_{W,y,uc} = \emptyset$. The active event sets are $\mathbb{H}_{W,y}(q_{W,y,1}) = \left(\bigcup_{\lambda \in \mathbb{W}(y)} \{e_{T,\lambda,1}\} \right) \cup \tilde{\mathbb{E}}_{I,y}$ and $\mathbb{H}_{W,y}(q_{W,y,2}) = \left(\bigcup_{\lambda \in \mathbb{W}(y)} \{e_{T,\lambda,2}\} \right) \cup \tilde{\mathbb{E}}_{O,y}$. The values of transition functions are $f_{W,y}(q_{W,y,1}, e) = q_{W,y,2}; e \in \left(\bigcup_{\lambda \in \mathbb{W}(y)} \{e_{T,\lambda,1}\} \right) \cup \tilde{\mathbb{E}}_{I,y}$ and $f_{W,y}(q_{W,y,2}, e) = q_{W,y,1}; e \in \left(\bigcup_{\lambda \in \mathbb{W}(y)} \{e_{T,\lambda,2}\} \right) \cup \tilde{\mathbb{E}}_{O,y}$. The state diagram of $\mathbf{G}_{W,y}$ is of a two-state form, similar to that in Fig. 4. The marked and the closed behaviors of $\mathbf{G}_{W,y}$ are $\mathbb{L}_m(\mathbf{G}_{W,y}) = \left(\left(\begin{array}{c} + \\ e \in \mathbb{H}_{W,y}(q_{W,y,1}) \end{array} \right) \left(\begin{array}{c} + \\ e \in \mathbb{H}_{W,y}(q_{W,y,2}) \end{array} \right) \right)^*$ and $\mathbb{L}(\mathbf{G}_{W,y}) = \overline{\mathbb{L}_m(\mathbf{G}_{W,y})}$, respectively.

3 Supervisor Design Restricting the Product Transfer Systems to Their Missions

In order to restrict the operation of the product transfer system W_y , where $y \in \{1, \dots, \rho\}$, to its missions, the automata of the product transfer missions, related to W_y , are used as supervisors of W_y . Particularly, all $\mathbf{R}_{T,j}$, where $j \in \mathbb{W}(y)$, are embedded to the local control unit of W_y , denoted as U_y . All $\mathbf{R}_{T,j}$ restrict W_y to its missions, through the following synchronous product $\mathbf{G}_{W,y,c} = \mathbf{G}_{W,y} \parallel \left(\bigparallel_{j \in \mathbb{W}(y)} \mathbf{R}_{T,j} \right)$, where $\mathbf{G}_{W,y,c}$ is the resulting controlled automaton that can be expressed as the 6-tuple $\mathbf{G}_{W,y,c} = (\mathbb{Q}_{W,y,c}, \mathbb{E}_{W,y,c}, f_{W,y,c}, \mathbb{H}_{W,y,c}, x_{W,y,c,0}, \mathbb{Q}_{W,y,c,m})$. The marked and the closed behaviors are $\mathbb{L}_m(\mathbf{G}_{W,y,c}) = \left(\begin{array}{c} + \\ \lambda \in \mathbb{W}(y) \end{array} (e_{T,\lambda,1} e_{T,\lambda,2}) \right)^*$, $\mathbb{L}(\mathbf{G}_{W,y,c}) = \overline{\left(\begin{array}{c} + \\ \lambda \in \mathbb{W}(y) \end{array} (e_{T,\lambda,1} e_{T,\lambda,2}) \right)^*}$. Clearly, $\mathbb{L}(\mathbf{G}_{W,y,c}) = \overline{\mathbb{L}_m(\mathbf{G}_{W,y,c})}$. Thus, $\mathbf{G}_{W,y,c}$ is a nonblocking automaton.

4 Formal Expressions Preventing Buffer Overflow and Underflow

As proposed in [1, 12] and [13], a critical property of a flexible manufacturing system (FMS), guaranteeing continuous and unobstructed product flow, is the protection of the buffers' overflow and underflow. Overflow takes place when a product transfer system or a machine tries to place a product in a full buffer. Underflow takes place when a product transfer system or a machine tries to pick a product from an empty buffer. Overflow and underflow may appear at any buffer, since each buffer is used both as product input and output spot of respective machines or transfer missions. The specification rules to prevent overflow and underflow of a buffer are the following:

1. If a buffer has a product, then the machines and the transfer missions, being associated with the buffer, are not allowed to place a product on it.

2. If a buffer is empty, then the machines and the transfer missions, being associated with the buffer, are not allowed to pick a product from it.
3. The machines and the transfer missions having the same physical buffer as product output spot are not allowed to process or transfer a product until the physical buffer is emptied by a machine or transfer mission that have this physical buffer as product input spot.

The three rules can be translated to the following two regular languages.

$$\mathbb{K}_I(\eta) = \overline{((c_{O,T}(\eta) + c_{O,M}(\eta))(c_{O,M}(\eta))^*(c_{I,T}(\eta) + c_{I,M}(\eta)))^*}; \eta \in \{1, \dots, \mu\} \quad (1)$$

$$\mathbb{K}_V(\eta) = \overline{((c_{V,T}(\eta) + c_{V,M}(\eta))(c_{I,T}(\eta) + c_{I,M}(\eta)))^*}; \eta \in \{1, \dots, \mu\} \quad (2)$$

where $c_{O,T}(\eta) = \sum_{\lambda \in \mathbb{V}_{O,T}(\eta)} (e_{T,\lambda,2})$, $c_{I,T}(\eta) = \sum_{\lambda \in \mathbb{V}_{I,T}(\eta)} (e_{T,\lambda,1})$, $c_{O,M}(\eta) = \left(\sum_{\lambda \in \mathbb{V}_{O,M}(\eta)} (d_{O,\lambda}) \right)$, $c_{I,M}(\eta) = \left(\sum_{(\lambda,\chi) \in \mathbb{V}_{I,M}(\eta)} (d_{I,\lambda,\chi}) \right)$, $c_{V,M}(\eta) = \left(\sum_{\lambda \in \mathbb{V}_{O,M}(\eta)} (d_{V,\lambda}) \right)$ and $c_{V,T}(\eta) = \sum_{\lambda \in \mathbb{V}_{O,T}(\eta)} (e_{T,\lambda,1})$, with

$$d_{V,\lambda} = \begin{cases} e_{M,\lambda,1}, & \text{if } \lambda \in \mathbb{I}_M \cup \mathbb{I}_P \\ \sum_{\psi=1}^{n_\lambda} e_{M,\lambda,\psi+1}, & \text{if } \lambda \in \mathbb{I}_A \\ \sum_{\psi=1}^{n_\lambda} e_{M,\lambda,\psi}, & \text{if } \lambda \in \mathbb{I}_L \end{cases}, \quad d_{I,\lambda,\chi} = \begin{cases} e_{M,\lambda,1}, & \text{if } \lambda \in \mathbb{I}_M \cup \mathbb{I}_P \\ e_{M,\lambda,\chi+1}, & \text{if } \lambda \in \mathbb{I}_A \\ e_{M,\lambda,\chi}, & \text{if } \lambda \in \mathbb{I}_L \end{cases} \quad (3)$$

$$d_{O,\lambda} = \begin{cases} e_{M,\lambda,2}, & \text{if } \lambda \in \mathbb{I}_M \cup \mathbb{I}_P \\ e_{M,\lambda,n_\lambda+2}, & \text{if } \lambda \in \mathbb{I}_A \\ \sum_{\psi=1}^{n_\lambda} e_{M,\lambda,n_\lambda+\psi}, & \text{if } \lambda \in \mathbb{I}_L \end{cases} \quad (4)$$

In the above formulas, the definition $\sum_{\lambda \in \emptyset} e_\lambda = \varepsilon$, where ε is the empty word, has been used for uniformity purposes. The second rule of V_η , where $\eta \in \{1, \dots, \mu\}$, is covered by the first regular language. The third rule of V_η is covered by the second regular language. The first rule of V_η is covered by the combination of the first and the second regular language. This language, i.e., the language $\mathbb{K}_V(\eta)$, is necessary as the events of the machines that place a product to their output buffer are uncontrollable. Thus, these events cannot be restricted, and the overflow and the underflow of the buffers must be guaranteed through a serial process or/and transfer of the products.

Remark 2 In the present study, initially the buffers are considered to be empty. This is a quite common hypothesis in FMS (indicatively see [1, 3, 12, 13] and the references within).

Remark 3 Underflow does not concern input product stacks, since they always have an available product to feed. Overflow does not concern output product stacks, since they always have enough space to receive a product.

5 A General Parametric Supervisor Automaton Realizing the Desired Languages

In this section, the issue of the realization of the regular languages, presented in Sect. 4, will be studied. Towards this aim, a general two-state supervisor form is proposed, being parametric with respect to the index $\eta \in \{1, \dots, \mu\}$ of the buffers and the regular language realized by the supervisor, i.e., with respect to the index $\chi \in \{I, V\}$. The general supervisor automaton will be introduced in the following 6-tuple form.

$$\mathcal{S}_\chi(\eta) = (\mathbb{Q}_{S,\chi}(\eta), \mathbb{E}_{S,\chi}(\eta), fs_{S,\chi}(\eta), \mathbb{H}_{S,\chi}(\eta), x_{S,\chi,0}(\eta), \mathbb{Q}_{S,\chi,m}(\eta)).$$

The set of the states of $\mathcal{S}_\chi(\eta)$ is $\mathbb{Q}_{S,\chi}(\eta) = \{q_{S,\chi,1}(\eta), q_{S,\chi,2}(\eta)\}$. The alphabet of $\mathcal{S}_\chi(\eta)$ is $\mathbb{E}_{S,\chi}(\eta) = \mathbb{E}_{S,\chi,1}(\eta) \cup \mathbb{E}_{S,\chi,2}(\eta) \cup \mathbb{E}_{S,\chi,3}(\eta) \cup \mathbb{E}_{S,\chi,4}(\eta)$, where the individual event sets will be specified later. The initial state is $x_{S,\chi,0}(\eta) = q_{S,\chi,1}(\eta)$ and the set of the marked states is $\mathbb{Q}_{S,\chi,m}(\eta) = \mathbb{Q}_{S,\chi}(\eta)$. The active event sets of $\mathcal{S}_\chi(\eta)$ are $\mathbb{H}_{S,\chi}(\eta, q_{S,\chi,1}(\eta)) = \mathbb{E}_{S,\chi,1}(\eta) \cup \mathbb{E}_{S,\chi,2}(\eta)$ and $\mathbb{H}_{S,\chi}(\eta, q_{S,\chi,2}(\eta)) = \mathbb{E}_{S,\chi,3}(\eta) \cup \mathbb{E}_{S,\chi,4}(\eta) \cup \mathbb{E}_{O,y}$. The values of the transition function of $\mathcal{S}_\chi(\eta)$ are

$$fs_{S,\chi}(\eta, q_{S,\chi,1}(\eta), e) = q_{S,\chi,1}(\eta); \forall e \in \mathbb{E}_{S,\chi,1}(\eta) \quad (5a)$$

$$fs_{S,\chi}(\eta, q_{S,\chi,1}(\eta), e) = q_{S,\chi,2}(\eta); \forall e \in \mathbb{E}_{S,\chi,2}(\eta) \quad (5b)$$

$$fs_{S,\chi}(\eta, q_{S,\chi,2}(\eta), e) = q_{S,\chi,2}(\eta); \forall e \in \mathbb{E}_{S,\chi,3}(\eta) \quad (5c)$$

$$fs_{S,\chi}(\eta, q_{S,\chi,2}(\eta), e) = q_{S,\chi,1}(\eta); \forall e \in \mathbb{E}_{S,\chi,4}(\eta) \quad (5d)$$

The closed behavior and the marked behavior of $\mathcal{S}_\chi(\eta)$ are $\mathbb{L}(\mathcal{S}_\chi(\eta)) = \overline{\mathbb{L}_m(\mathcal{S}_\chi(\eta))} = \left(\mathbb{E}_{S,\chi,1}^* \mathbb{E}_{S,\chi,2}^* \mathbb{E}_{S,\chi,3}^* \mathbb{E}_{S,\chi,4}^* \right)^*$. In Fig. 5, the state diagram of $\mathcal{S}_\chi(\eta)$ is depicted.

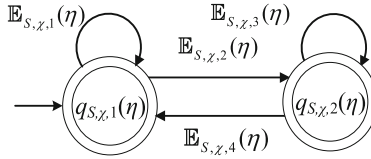


Fig. 5. State diagram of the general supervisor

For $\mathcal{S}_\chi(\eta)$ to realize the two desired languages, the alphabets $\mathbb{E}_{S,\chi,1}(\eta)$, $\mathbb{E}_{S,\chi,2}(\eta)$, $\mathbb{E}_{S,\chi,3}(\eta)$ and $\mathbb{E}_{S,\chi,4}(\eta)$ are specified per language as follows: Regarding $\mathbb{K}_I(\eta)$, let

$$\mathbb{E}_{S,I,1}(\eta) = \emptyset, \mathbb{E}_{S,I,2}(\eta) = c_{O,T}(\eta) + c_{O,M}(\eta), \mathbb{E}_{S,I,3}(\eta) = c_{O,M}(\eta) \quad (6a)$$

$$\mathbb{E}_{S,I,4}(\eta) = c_{I,T}(\eta) + c_{I,M}(\eta) \quad (6b)$$

Thus, $\mathbb{L}(\mathbf{S}_I(\eta)) = \mathbb{L}_m(\mathbf{S}_I(\eta)) = \mathbb{K}_I(\eta)$. Regarding $\mathbb{K}_V(\eta)$, let

$$\mathbb{E}_{S,V,1}(\eta) = \mathbb{E}_{S,V,3}(\eta) = \emptyset, \quad \mathbb{E}_{S,V,2}(\eta) = c_{V,M}(\eta) + c_{V,T}(\eta) \quad (7a)$$

$$\mathbb{E}_{S,V,4}(\eta) = c_{I,T}(\eta) + c_{I,M}(\eta) \quad (7b)$$

Thus, $\mathbb{L}(\mathbf{S}_V(\eta)) = \mathbb{L}_m(\mathbf{S}_V(\eta)) = \mathbb{K}_V(\eta)$.

6 Realization of the Control Architecture

The supervisors $\mathbf{R}_{T,j}$, proposed in Sect. 3 to restrict the commands (controllable events) of the respective product transfer systems, is embedded to the local control unit U_y . In many cases, the platform of the local control unit is a Programmable Logic Controller (PLC). Using $\mathbf{R}_{T,j}$, the controlled product transfer systems $\mathbf{G}_{W,y,c}$ are formed. The rest supervisors, i.e., the supervisors designed in Sect. 5, impose prevention of buffer underflow and overflow by restricting appropriate commands (controllable events) of the machines and the product transfer systems. For each buffer, two supervisors are used. The two supervisors, i.e., $\mathbf{S}_I(\eta)$ and $\mathbf{S}_V(\eta)$ are implemented to a central control unit, e.g., a Supervisory Control And Data Acquisition (SCADA) system). The central control unit acquires all commands and signals from all subsystems of the manufacturing system. The total controlled automaton of the manufacturing system, resulting after the application of all local and central supervisors, is

$$\mathbf{G}_c = \left(\prod_{i=1}^m \mathbf{G}_{M,i} \right) \parallel \left(\prod_{y=1}^{\rho} \left(\mathbf{G}_{W,y} \parallel \left(\prod_{j \in \mathbb{W}(y)} \mathbf{R}_{T,j} \right) \right) \right) \parallel \left(\prod_{\eta=1}^{\mu} (\mathbf{S}_I(\eta) \parallel \mathbf{S}_V(\eta)) \right) \quad (8)$$

The automaton \mathbf{G}_c has the following two properties:

- (i) All supervisors are physical realizable (PR) (see [19]), with respect to $\mathbf{G} = \left(\prod_{i=1}^m \mathbf{G}_{M,i} \right) \parallel \left(\prod_{y=1}^{\rho} \mathbf{G}_{W,y} \right)$, through the synchronous product in (8).
- (ii) The total controlled automaton \mathbf{G}_c is a nonblocking automaton, for continuous product flow to the buffers.

7 Conclusions

In the present paper, the modelling and control of a general parametric multi-product manufacturing system has been tackled. The system consists of manufacturing machines (i.e., Milling Machines, Painting Machines, Lathes and Assembly Machines) and product transfer systems (Robotic Manipulators and Conveyor Belts). The DES models of the machines and transfer systems have been presented in the form of parametric 6-tuples. A supervisory control scheme has been developed based on specific behavioral rules translated to regular languages. The supervisor automata of the behavior languages have been derived by developing a general two-state and parametric supervisor form.

In this paper, the designed supervisory control architecture is as restrictive required by the property that the machines' events commanding product drops to the respective

output buffer are uncontrollable. Clearly, if the manufacturing system is enriched with one or more multi-slot buffers assigned as registers of finished products or machine's output buffers, then a more permissive supervisory control architecture could be derived.

Future concerns of the present work is the design of a supervisory control architecture acting as the system's task coordinator, synchronizing accordingly all machines and product transfer system commands. A perspective of the present work is the computation of the complexity of the supervisors and their PLC implementation. The extension of the present results to the case of general manufacturing processes with multi slot buffers is currently under investigation.

Acknowledgement. Research supported by the MSc Program in "Advanced Control Systems and Robotics", National and Kapodistrian University of Athens, Greece.

References

1. Hill, R. C., Cury, J. E. R., de Queiroz, M. H., Tilbury, D. M., Lafortune, S.: Multi-level hierarchical interface-based supervisory control," *Automatica*, vol. 46, 1152-1164 (2010).
2. Malik, R., Leduc, R.: Hierarchical modelling of manufacturing systems using discrete event systems and the conflict preorder *Discrete Event Dynamic Systems*, vol. 25, 177–201 (2015).
3. Wonham, W.M., Kai, C.: *Supervisory Control of Discrete-Event Systems*; Springer: Cham, Switzerland (2019).
4. Cassandras, C. G., Lafortune, S.: *Introduction to Discrete Event Systems*, 3rd ed., Springer: Cham, Switzerland (2021).
5. Leduc, R. J., Brandin, B. A., Lawford, M., & Wonham, W. M.: Hierarchical interface-based supervisory control_part I: serial case. *IEEE Transactions on Automatic Control*, vol. 50(9), 1322-1335 (2005)
6. Leduc, R. J., Lawford, M., & Wonham, W. M.: Hierarchical interface-based supervisory control part II: Parallel case. *IEEE Transactions on Automatic Control*, vol 50(9), 1336-1348 (2005).
7. Hill, R. C., Tilbury, D. M.: Incremental hierarchical construction of modular supervisors for discrete-event systems. *International Journal of Control*, vol. 81(9), 1364-1381 (2008).
8. Ngo, Q., Seow, K.: A hierarchical consistency framework for real- time supervisory control, *Discrete Event Dyn. Syst.*, vol. 28, no. 3, pp. 375–426 (2018).
9. Komenda, J., Masopust, T.: Hierarchical Supervisory Control Under Partial Observation: Normality, *IEEE Transactions on Automatic Control* (2023). doi: <https://doi.org/10.1109/TAC.2023.3257232>.
10. Su, R., van Schuppen, J. H., Rooda, J. E.: Model abstraction of nondeterministic finite-state automata in supervisor synthesis. *IEEE Transactions on Automatic Control*, vol. 55(11), 2527–2541 (2010)
11. Feng, L., Wonham, W. M.: Computationally efficient supervisor design: abstraction and modularity, *2006 8th International Workshop on Discrete Event Systems*, Ann Arbor, MI, pp. 3-8, (2006).
12. de Queiroz, M. H., Cury, J.E.R.: Modular Multitasking Supervisory Control of Composite Discrete-Event Systems, *IFAC Proceedings Volumes*, vol. 38 (1), pp. 91-96, (2005).
13. de Queiroz, M. H., Cury, J.E.R., Wonham, W.M.: Multitasking Supervisory Control of Discrete-Event Systems, *Discrete Event Dyn Syst*, vol. 15, pp. 375–395 (2005).

14. Koumboulis, F. N., Fragkoulis, D. G., Kalkanas, I., Fragulis, G. F.: Supervisor Design for a Pressurized Reactor Unit in the Presence of Sensor and Actuator Faults, *Electronics*, 11(16), 2534 (2022).
15. Kouvakas, N. D., Koumboulis, F. N., Fragkoulis, D. G., Souliotis, A.: Modular Supervisory Control for the Coordination of a Manufacturing Cell with Observable Faults, *Sensors*, 23, 163 (2023).
16. Koumboulis, F. N., Fragkoulis, D. G., Menexis, A. N.: Supervisory Control for Flexibility of Production Manufacturing Processes, *IEEE 21st Int. Conf. on Intelligent Engineering Systems 2017 (INES 2017)*, Larnaca, Cyprus, October 20–23 (2017).
17. Koumboulis, F. N., Fragkoulis, D. G., Ioannou, K. A.: Control of Router Nodes in Production Manufacturing Processes, *2018 7th Int. Conf. on Systems and Control (ICSC)* (2018).
18. Leitão, H.A.S., Rosso, R.S.U., Leal, A.B., Zoitl, A.: Fault Handling in Discrete Event Systems Applied to IEC 61499. In *Proceedings of the 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETF A)*, Vienna, Austria, 8–11 September (2020).
19. Koumboulis, F. N., Fragkoulis, D. G., Arapakis, S.: Supervisor design for an assembly line in the presence of faults, *27th IEEE International Conference on Emerging Technology and Factory Automation*, Stuttgart, Germany, September 6–9 (2022).
20. Cai, K., Wonham, W. M.: New results on supervisor localization, with case studies. *Discrete Event Dynamic Systems*, 25, 1218 203–226 (2015).



An Oscillator Model Applied to Power Take-Off and Tuned Control for Renewable Energy

Nikolaos I. Xiros¹(✉) and Erdem Aktosun²(✉)

¹ Boysie Bollinger School of Naval Architecture and Marine Engineering, University of New Orleans, 2000 Lakeshore Dr, New Orleans, LA 70148, USA

nxiros@uno.edu

² Department of Shipbuilding and Ocean Engineering, İzmir Katip Çelebi University, Havaalanı Sosesi Cd, İzmir 35620, Turkey

erdem.aktosun@ikcu.edu.tr

Abstract. The square-law model inspired by its electronic circuit namesake is adopted to describe the dynamics of an electromechanical system that can be used as a power takeoff subsystem in renewable power generation plants where the energy source is near-periodic reciprocation along an axis. Typical cases of such processes in the maritime environment include many types of ocean energy converters as well as several hydrokinetic energy converters that convert water current translational energy into the oscillatory motion of bluff bodies located in-stream with mechanisms like e.g. vortex-induced vibrations (VIV). In the latter case specifically, the steady fluid flow causes through a vortex pattern an in-stream bluff body, typically a cylinder outfitted oftentimes with appendages, to oscillate in a near periodic pattern. From a dynamical system perspective, this can be described as a linear or more appropriately positive feedback oscillator. The most common form of feedback oscillator is an amplifying process connected in a feedback loop with its output fed back into its input through a frequency selective filter to provide positive feedback. When the steady flow to the amplifying process is initially engaged, fluctuations and noise in the loop provide a non-zero source to get oscillations started. The noise travels around the loop and is amplified and filtered until very quickly it converges on a sine wave at a single frequency. Such dynamic behavior occurs not only in electronic circuits but also in renewable energy harvesting processes of importance to ocean engineering applications.

1 Introduction

Renewable ocean energy refers to the energy obtained from the physical and chemical characteristics of the ocean, including its movement. This type of energy can be produced from various sources, such as waves, tides, and currents, as well as temperature differences in the ocean. According to the National Renewable Energy Laboratory, if the U.S. fully harnesses its ocean energy resources, it could generate an amount of electricity equivalent to over 50% of what the country produced in 2019. Government officials and researchers in the United States anticipate that ocean energy will initially be utilized to supply energy and water for island and coastal communities, as well as for offshore

operations. These stakeholders believe that using ocean energy for these purposes may aid in the advancement of technology and help make ocean energy more competitive in other markets. Different types of devices can be used to harness ocean energy resources, and the design and size of these devices depend on various factors, including the type of energy they harness and the amount of power they need to generate. The design and scale of devices used to capture the energy of moving water vary depending on whether they are designed to harness waves, tides, or ocean currents [1].

The use of oil, coal, and other fossil fuels to provide energy has caused increasingly serious environmental problems in recent years, and the fact that these resources are gradually being depleted has made research into clean and renewable energy more critical and strategic [2]. In parallel with this approach, renewable clean energy production from rivers, seas, and oceans has become an important research topic in recent years. Energy systems obtained from vortex-induced vibrations (VIV), on the other hand, are a new research area, and significant progress is still anticipated in this regard [3]. Compared to traditional hydroelectric power plants, it has been observed that these types of systems have a long service life, relatively low maintenance costs, and high energy density [4].

VIVs are a nonlinear dynamic fluid-structure interaction observed in many engineering fields (such as offshore structures, towing cables, mooring systems, ship appendages, buildings, bridges, power cables, pipelines, antennas, chimneys, heat exchangers, etc.). There are extensive studies examining the physics of the subject [5–7]. When bluff bodies such as circular cylinders are exposed to flows at a certain value and above of Reynolds numbers, two boundary layer separations can be observed. These separated boundary layers usually cause the formation of vortices that detach from the body and follow each other periodically, often at both ends of the body when exposed to flows with certain values and above of Reynolds numbers. The instantaneous pressure differences and forces caused by these periodic vortex shedding on the body force it to vibrate. Assuming that these bodies are elastically connected in the direction of vortex shedding (i.e., perpendicular to the flow direction), large-amplitude movements occur in the range of values where the shedding frequency of the vortices approaches the natural frequency of the body. These amplitudes are often visible on engineering structures. This area of large amplitudes is known as the synchronization region. In low-density fluids such as air, these amplitudes tend to be narrow-banded because structural responses, physical properties, and natural frequencies of the bodies are more dominant than the vortex shedding. However, vibrations in dense fluids such as water contribute significantly to the frequency characteristics and result in large, broad-amplitude vibration responses. The effect of these vibrations often leads to negative consequences such as structural failure, structural fatigue, noise generation, and operational disturbances, especially in slender and long ocean engineering structures such as towed cables, risers, and mooring systems. However, in recent years, energy harvesting applications from VIV have become increasingly important. The VIV energy harvester (VIVEH) has gained considerable interest due to its distinctive synchronization region, which serves to safeguard the harvester while generating a substantial amount of power. The typical VIVEHs consist of a cylindrical bluff body that is combined with an energy conversion structure [8, 9]. Therefore, the accurate prediction of motion characteristics such as amplitudes and

frequencies of objects moving due to flow-induced vibrations has a significant impact on these energy-harvesting systems [4, 10].

One of the efficient and fast method to make an accurate prediction of motions could be direct numerical simulations attempting to solve the complex problem of coupled fluid and structure interactions, but they are computationally intensive and not suitable for obtaining fast solutions, especially for high Reynolds number flows. The other method could be reduced order models which offer a faster solution, but they often suffer from inaccuracies in representing the fluid wake. For example, wake oscillator models mimic the dynamic characteristics of vortex-induced vibrations but lack a first-principles derivation, requiring empirical coefficients for model training. As a result, wake models may not perform well outside of their trained data range. Semi-empirical techniques for predicting cross-flow motion VIV have been found to be more accurate by utilizing simple structural models along with an empirical database obtained from forced motion experiments. These methods directly incorporate real measured forces into the prediction of VIV coupling, leading to improved accuracy in the predictions [11]. We used a comprehensive force database by measuring forces on cylinders that undergo controlled combined in-line and cross-flow motion in a uniform free stream that aims to capture a wide range of potential vibrations in our experimental database, using prescribed sinusoidal combined in-line and cross-flow motions based on observations from previous free vibration experiments conducted in [12, 13]. To achieve a more detailed and extensive database, we utilized a unique automated experimental apparatus to gather data on forced combined in-line and cross-flow motions at a specific Reynolds number [14] to investigate a model of VIV oscillation applied to power take-off and tuned control for renewable energy.

Obtaining energy from vortex-induced vibrations has the potential to provide significant amounts of clean energy from rivers and oceans due to the synchronization zone mentioned earlier. Mechanisms that generate energy in this way produce a considerable amount of power [8]. The flow field caused by the vortices can be estimated using Particle Image Velocimetry (PIV) measurement instrument [15]. For systems such as VIV that can only be explained physically using nonlinear equations, accurately predicting the performance of energy generation and producing an efficient model is a fundamental technological challenge. To validate these methods, an efficient database is needed despite the numerous academic publications on energy obtained from VIVs. There is limited published information on high-resolution measurements, systematic and careful experimental studies, and documentation for energy harvesting systems from flow-induced vibrations of low-mass cylinders [3, 16].

In this study, the main goals of the control strategy are eventually intended to efficiently harness energy from VIV while ensuring that the system and its parts operate within a safe range. We investigate estimating dominant frequency from motions and forces responses from a chosen a specific experimental run from comprehensive force database [14]. The intended modes of operation include, but are not restricted to: maximizing energy extraction up to the full power rating of the system, ensuring energy extraction at a safe level when VIV oscillation exceeds the system's capacity or energy demand, controlling energy extraction, transitioning to idle or offline mode, and incorporating a fail-safe mode that activates in the event of a malfunction, placing the system

in a secure state. Specifically, the control objective that was thoroughly studied will establish a sturdy control plan that can extract the highest amount of energy from VIV, up to the system's full rated power. To achieve this goal, the combined motions of the circular cylinder must be analyzed before selecting the design of the controller.

2 Experimentation of Force Responses from 2-DOF Forced Motion of a Circular Cylinder in a Free Stream

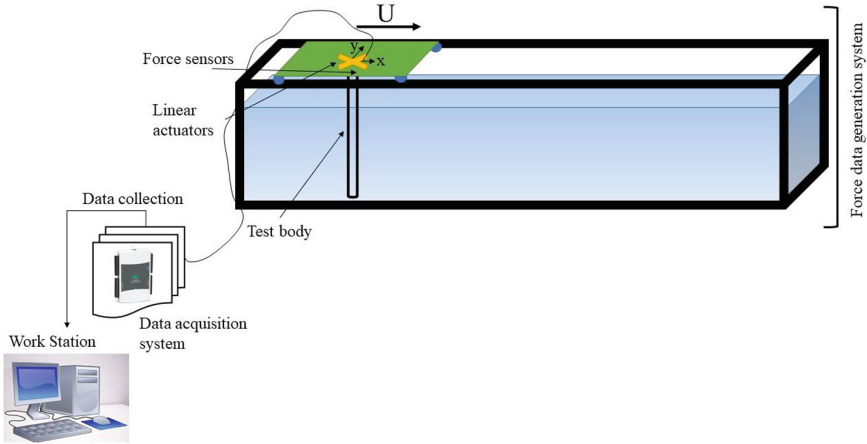


Fig. 1. A diagram illustrating the experimental arrangement is presented, depicting how the test cylinder is undergoing prescribed motion within the x-y plane in a forced motion.

The nonlinearity of the interaction between the fluid and structural system does not necessarily ensure that the motions follow sinusoidal patterns in both directions despite the fact that the combination of cross-flow and in-line vortex-induced vibrations (VIV) usually results in a cylinder's motion resembling figure-eight and crescent shapes (Lissajous patterns) when observed from a fixed reference frame attached to the carriage. In our analysis, we make the assumption that the primary motion of the body in both the in-line and cross-flow directions can be approximated as sinusoidal, with a single amplitude and frequency governing these motions to maintain simplicity and consistency with definitions in [12, 13]. Furthermore, we assume that the frequency of in-line motions is twice that of cross-flow motions, enabling the use of a single frequency parameter to characterize the body's motion in accordance with experimental observations. By adopting this approach, we can reduce the number of variables involved in defining the motion of the body, resulting in forced motions that can be characterized as in [14] (Fig. 2).

$$y = A_y \sin(\omega_y t) \quad (1)$$

$$x = A_x \sin(2\omega_y t + \theta) \quad (2)$$

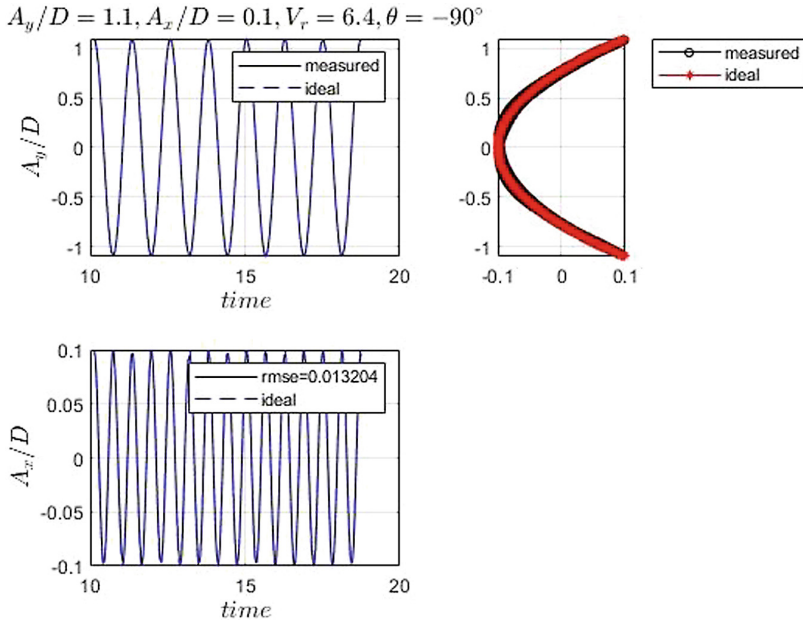


Fig. 2. An illustrative run of the time traces for cross-flow and in-line motions and the orbital motion of the cylinder as well as the error calculation between ideal and measured motions: nondimensional cross-flow amplitude (A_y/D) = 1.1, nondimensional in-line motion amplitude (A_x/D) = 0.1, reduced velocity (V_r) = 6.4, and phase angle (θ) = -90° .

The relative phase difference, denoted as θ , between the in-line and cross-flow motions determines the orbital movement of the cylinder in relation to the fixed reference frame attached to the carriage. Subsequently, the motion of the cylinder can be described using non-dimensional amplitudes (A_y/D and A_x/D), a single reduced velocity in the cross-flow direction (V_r), and the phase angle (θ), as expressed by the following equation:

$$V_r = \frac{2\pi U}{\omega_y D} \quad (3)$$

In the experimental setup, a small test cylinder with low mass (diameter $D = 3.85$ cm, submerged length $L = 57.25$ cm, mass $m = 0.2679$ kg) was subjected to prescribed motions while being towed at a constant velocity in a flow tank. The Reynolds number was maintained at 7620. The experiments were conducted at the Experimental Fluid Mechanics flow tank located at the Narragansett Bay Campus of the University of Rhode Island, which has dimensions of 107 cm width, 280 cm length, and 75 cm depth. Hydrodynamic forces were measured using ATI Gamma SI-65-10 and SI-130-10 six-axis force sensors.

The non-dimensional parameters governing the body's motion were varied, including in-line amplitude (ranging from 0.1 to 0.5 in increments of 0.1), cross-flow amplitude (ranging from 0.1 to 1.6 in increments of 0.25), reduced velocity (ranging from 4 to

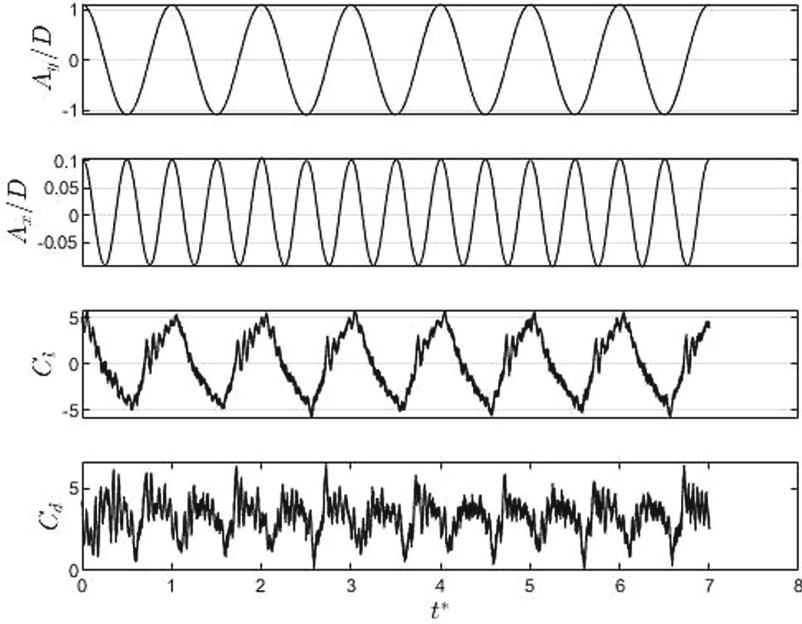


Fig. 3. An illustrative run of the time traces for cross-flow and in-line motions, as well as the time-dependent lift and drag forces, can be described for the following parameters: nondimensional cross-flow amplitude (A_y/D) = 1.1, nondimensional in-line motion amplitude (A_x/D) = 0.1, reduced velocity (Vr) = 6.4, and phase angle (θ) = -90° .

8 in increments of 0.2), and phase difference between in-line and cross-flow motions (ranging from -180° to 180° in increments of 30°). This resulted in a total of 9555 experiments, comprising the experimental database for the single Reynolds number of 7620. The experimental setup, including the location of the force sensor and coordinate system, is illustrated in Fig. 1, looking down the length of the tank.

In Fig. 2, a time series plot displays data from a specific run obtained from the force database. The plot reveals that the root mean square error, which measures the difference between the ideal prescribed motion and the measured motion, is minimal. Additionally, the plot illustrates that the cylinder is exhibiting a crescent-shaped motion in this particular run.

We obtained a time series of force and motion measurements for different parameter combinations. An example of such measurements for a specific run is shown in Fig. 3, which illustrates the normalized motion of the cylinder in the cross-flow and in-line directions, as well as the measured normalized lift and drag forces. It is important to note that the cylinder oscillations were limited to a number of cycles in the cross-flow direction, depending on the reduced velocity of the specific run, due to the limited length of the experimental tank.

Figure 4 illustrates the results of power spectral analysis conducted on the lift and drag force coefficients at a specific point, using data from a selected run in the force database.

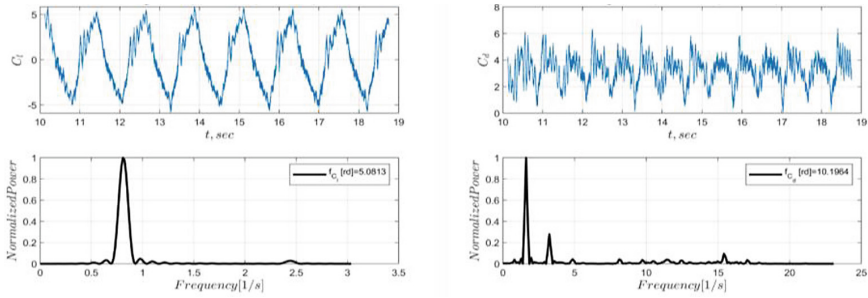


Fig. 4. Normalized the power spectral density of lift and drag force coefficients, as well as their respective peak frequencies at: $(A_y/D) = 1.1$, $(A_x/D) = 0.1$, $(V_r) = 6.4$, $(\theta) = -90^\circ$.

The figure indicates that the first harmonic frequency is the dominant frequency for the lift force coefficient, while the first frequency also dominates the drag force coefficient.

In order to determine the frequency of the circular cylinder using the Square-Law Detector algorithm, it is important to analyze the interplay between the cylinder's motions and forces. This analysis is crucial in selecting the design of the controller. A specific run from the forced experimental data is chosen for this purpose because it is within the free vibration range and has a high cross-flow amplitude [14]. The high cross-flow amplitude is particularly important for energy harvesting from VIV processes. Therefore, by examining this specific run, it becomes possible to estimate the frequency and design a suitable controller for the system.

3 Results and Discussion

One of the most important aspects in the controller implementation for motion control and power management purposes is VIV motion and force oscillation frequency acquisition. A straightforward approach is to use 6 degrees of freedom force and torque sensor, like the ATI F/T sensor: gamma SI-65-5 device [17]. The ATI Multi-Axis Force/Torque Sensor system is capable of measuring force and torque in all six components. It includes a transducer, a high-flex cable that is shielded, and an intelligent data acquisition system, which can be connected to an Ethernet/DeviceNet interface or F/T controller. These Force/Torque sensors are widely used in various industries for activities such as product testing and robotic assembly. They are also used in various research fields with many other applications. The response of the ATI Multi-Axis Force/Torque Sensor consists of volt measurements which is seen the right side of Fig. 5.

The ATI SI-65-5 and SI-130-10 force sensors are utilized in forced vibration experiments to determine the hydrodynamic forces applied to a cylinder. The SI-65-5 force sensor is capable of measuring forces up to 65 N in both the x and y directions, while the SI-130-10 force sensor can measure up to 130 N in both directions.

An analog nonlinear signal processing approach of the crossflow motion signal (y/D) is proposed next. The approach is shown in block diagram form in Fig. 6 and its implementation in MATLAB/Simulink[®] as seen in Fig. 7 using the experimental data series.

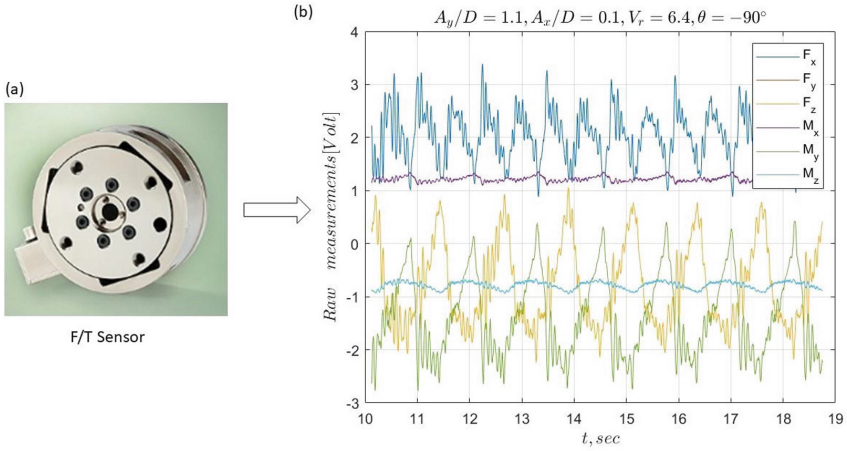


Fig. 5. Multi-axis force/torque sensor (a), output raw signals (b).

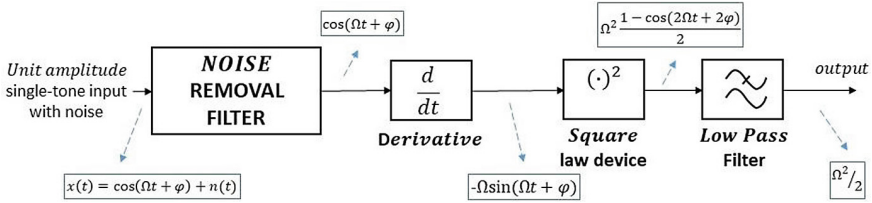


Fig. 6. Square-law detector signal processing architecture.

The rationale for the method above is explained below. The waveform signal on the left-hand side of the block diagram above needs to have a predominant frequency in the form of a monochromatic signal. It must also not have any bias (DC offset).

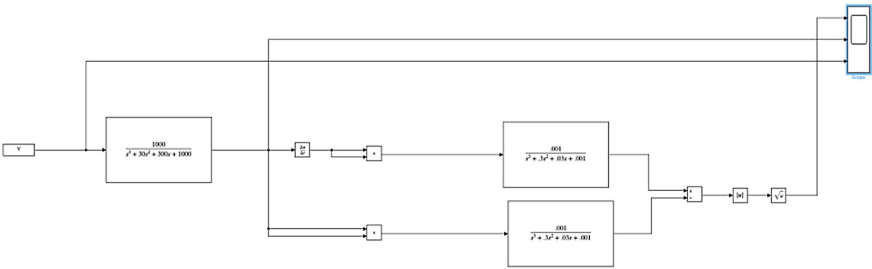


Fig. 7. Cross-flow motion signal processing to obtain cross-flow motion frequency.

A square-law detector starts with passing the filtered signal through a differentiator to obtain its derivative with respect to time. The derivative then is squared and passed through a low-pass filter to reject any high frequency components due to the second-order

signal processing applied. If the signal has not unit amplitude some further processing is required to obtain a frequency estimate of the signal at the input port. Consequently, a square-law detector can track an input frequency, or it can generate a frequency that is a multiple of the input frequency. The former property is used for demodulation, and the latter property is used for indirect frequency synthesis.

The processing architecture is further explained in the case of a unit-amplitude monochromatic input signal.

$$\xi(t) = \cos(\Omega t + \psi) \Rightarrow \dot{\xi}(t) = -\Omega \sin(\Omega t + \psi) \quad (4)$$

In the above ψ is an arbitrary, unknown phase and Ω capital the unknown frequency to be determined. Driving the derivative of the input signal through the quadratic block yields the following signal.

$$\xi_2(t) = (\dot{\xi}(t))^2 = \Omega^2 \sin^2(\Omega t + \psi) = \Omega^2 \frac{1 - \cos 2(\Omega t + \psi)}{2} \quad (5)$$

In effect, driving the above signal through a low-pass filter (LPF) the high frequency component (sinusoidal of frequency double Ω) is eliminated and the bias (“DC”) term which is equal to half frequency Ω squared survives.

In the figures that follow the theoretical process above is employed for the VIV experimental data series. The filter used is a simple second-order one with unity DC gain. Better filters like e.g. 10th-order Butterworth filters with appropriate cutoff frequencies can be used to achieve improved results. As seen after a transient interval, the output of the signal processing scheme is 5.2 rad/s for cross flow motion frequency and lift force coefficient.

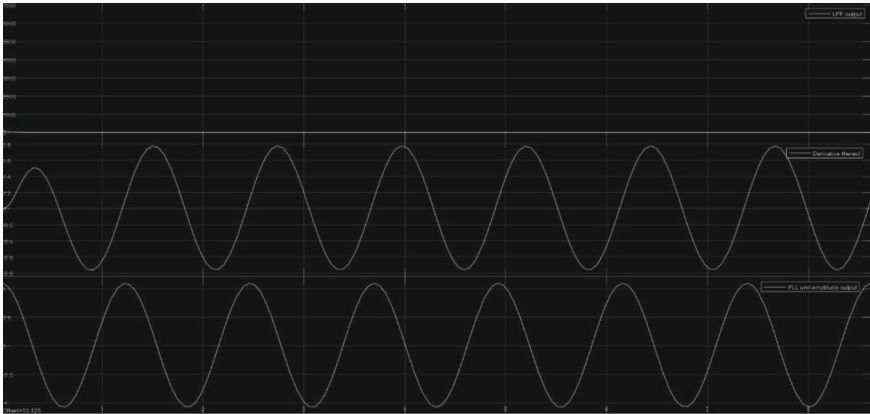


Fig. 8. Analog signal processing for cross flow motion amplitude to obtain the frequency of a signal: sinusoidal input (bottom), derivative thereof (middle) and LPF output signal (top).

Figure 8 illustrates the outcomes of determining frequency through the square-law detector using cross-flow motion. The lower figure in Fig. 8 displays the original signals

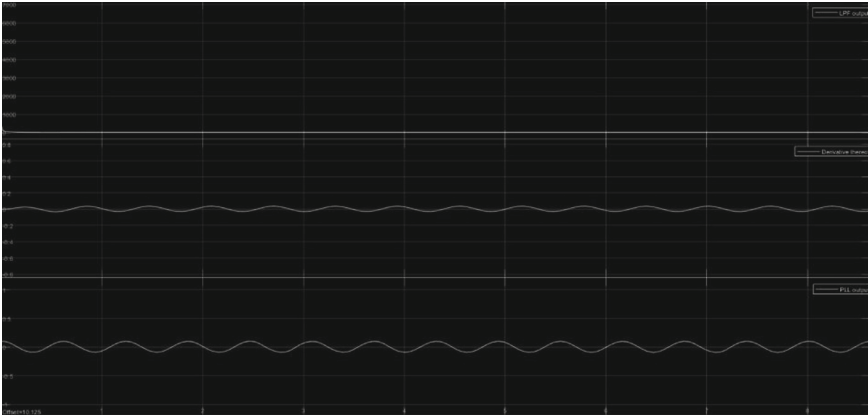


Fig. 9. Analog signal processing for inline motion amplitude to obtain the frequency of a signal: sinusoidal input (bottom), derivative thereof (middle) and LPF output signal (top).

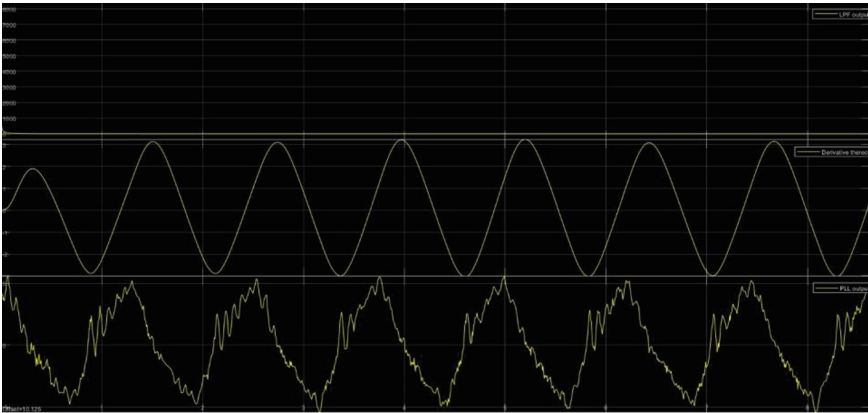


Fig. 10. Analog signal processing for lift force coefficient to obtain the frequency of a signal: noisy input signal (bottom), derivative thereof (middle) and LPF output signal (top).

captured, including noise. The middle plot shows the filtered signal derived from the initial measurements. The top plot demonstrates the frequency determination result, which matches the peak frequency analysis obtained from the fast Fourier transform method. Likewise, Fig. 9 exhibits the results of frequency determination from the square-law detector method using in-line motion. The bottom figure in Fig. 9 presents the original measured signals with noise. The middle plot displays the filtered signal obtained from the primary measurements. The top plot indicates the frequency determination outcome, which matches the peak frequency analysis derived from the fast Fourier transform method. Furthermore, Fig. 10 displays the outcomes of frequency determination from the square-law detector method using lift force coefficient. The bottom figure in Fig. 10 shows the original signals captured, including noise. The middle plot shows the filtered signal obtained from the initial measurements. The top plot indicates the frequency

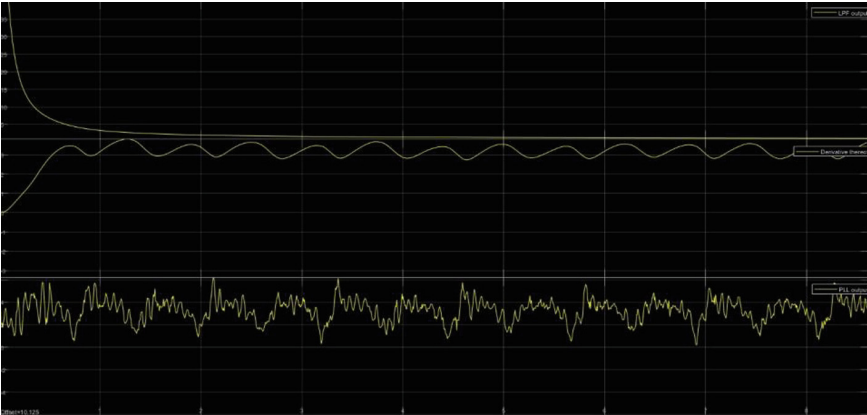


Fig. 11. Analog signal processing for drag force coefficient to obtain the frequency of a signal: noisy input signal (bottom), derivative thereof (middle) and LPF output signal (top).

determination result, which matches the peak frequency analysis obtained from the fast Fourier transform method. Finally, Fig. 11 illustrates the results of frequency determination from the square-law detector method using drag force coefficient. The bottom figure in Fig. 11 displays the original measured signals with noise. The middle plot shows the filtered signal derived from the initial measurements. The top plot demonstrates the frequency determination outcome, which does not converge with the peak frequency analysis obtained from the fast Fourier transform method.

Without the need for Fourier transform, the system can estimate frequency in real-time as the data comes in after the run. It processes the incoming data and produces an accurate frequency estimate. This approach is much faster and more efficient compared to the demanding Fast Fourier Transform method. Therefore, this system provides a more effective and streamlined solution for frequency estimation.

To improve the accuracy of the frequency determination, we need to devote more attention to refining the design of the leading filter. As the level of noise in the signal increases, it becomes increasingly important to ensure that the filter is able to handle this effectively, and this will require more detailed optimization. We observed that the performance of the leading filter varied across different cases. In cases where there was no noise or a reasonable amount of noise, the filter worked well and was able to produce a signal with an accurate frequency. However, when the noise level was excessive, the filter was not effective and was unable to produce an accurate signal. When we examined the lift coefficient filter, we found that it was not entirely clean but was still able to produce a reasonable signal in the presence of some noise. The frequency was determined accurately in this case. However, when we looked at the drag force coefficient, we found that there was excessive noise in the signal, which made it difficult for the leading filter to produce an accurate signal. Despite the fact that the filtered signals appeared sinusoidal in most cases, this was not true for the drag force coefficient. Next step, it will be necessary to devote more attention to tuning the system,

particularly the leading filter, to ensure that it can handle signals with high levels of noise and produce accurate frequency determinations.

For example, in order to have adaptive control [18], it is crucial to be capable of recognizing the frequency of the system. This is significant for the control system because the frequency varies based on the incoming velocity in real-world situations. To accurately determine the frequency value, we require an efficient algorithm other than the fast Fourier transform.

4 Conclusion and Future Work

The focus of this study is to develop a control strategy that can efficiently extract energy from Vortex-Induced Vibration (VIV) while maintaining the safety of the system and its components. To achieve this, the dominant frequency is estimated by the square-law detector method by analyzing the motions and forces responses in a specific experimental run from a comprehensive forced motion database. In future work, the following items could be considered for further investigation. The leading filter design needs to be further refined to obtain more accurate frequency estimates, especially for excessively noisy signals arising in the power takeoff system for VIV energy harvesting.

References

1. GAO UG. Renewable Ocean Energy [Internet]. 2021 [cited 2023 May 4]. Available from: <https://www.gao.gov/assets/gao-21-533sp.pdf>
2. Kabir E, Kumar P, Kumar S, Adelodun AA, Kim K-H. Solar energy: Potential and future prospects. *Renew Sustain Energy Rev* [Internet]. 2018 Feb;82:894–900. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1364032117313485>
3. Wang J, Zhao G, Zhang M, Zhang Z. Efficient study of a coarse structure number on the bluff body during the harvesting of wind energy. *Energy Sources, Part A Recover Util Environ Eff* [Internet]. 2018 Aug 3;40(15):1788–97. Available from: <https://www.tandfonline.com/doi/full/https://doi.org/10.1080/15567036.2018.1486916>
4. Bernitsas MM, Raghavan K, Ben-Simon Y, Garcia EMH. VIVACE (Vortex Induced Vibration Aquatic Clean Energy): A New Concept in Generation of Clean and Renewable Energy From Fluid Flow. *J Offshore Mech Arct Eng* [Internet]. 2008 Nov 1;130(4). Available from: <https://asmedigitalcollection.asme.org/offshoremechanics/article/doi/https://doi.org/10.1115/1.2957913/471896/VIVACE-Vortex-Induced-Vibration-Aquatic-Clean>
5. Bearman PW. Vortex Shedding from Oscillating Bluff Bodies. *Annu Rev Fluid Mech* [Internet]. 1984 Jan;16(1):195–222. Available from: <https://www.annualreviews.org/doi/https://doi.org/10.1146/annurev.fl.16.010184.001211>
6. Williamson CHK, Govardhan R. Vortex-induced vibrations. *Annu Rev Fluid Mech* [Internet]. 2004 Jan 21;36(1):413–55. Available from: <https://www.annualreviews.org/doi/https://doi.org/10.1146/annurev.fluid.36.050802.122128>
7. Sarpkaya T. A critical review of the intrinsic nature of vortex-induced vibrations. *J Fluids Struct* [Internet]. 2004 May;19(4):389–447. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0889974604000350>
8. Leclercq T, de Langre E. Vortex-induced vibrations of cylinders bent by the flow. *J Fluids Struct* [Internet]. 2018 Jul;80:77–93. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0889974617307764>

9. Wang J, Geng L, Ding L, Zhu H, Yurchenko D. The state-of-the-art review on energy harvesting from flow-induced vibrations. *Appl Energy* [Internet]. 2020 Jun;267:114902. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0306261920304141>
10. Lee JH, Xiros N, Bernitsas MM. Virtual damper–spring system for VIV experiments and hydrokinetic energy conversion. *Ocean Eng* [Internet]. 2011 Apr;38(5–6):732–47. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0029801810002842>
11. Chaplin JR, Bearman PW, Cheng Y, Fontaine E, Graham JMR, Herfjord K, et al. Blind predictions of laboratory measurements of vortex-induced vibrations of a tension riser. *J Fluids Struct* [Internet]. 2005 Nov;21(1):25–40. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0889974605001374>
12. JAUVTIS N, WILLIAMSON CHK. The effect of two degrees of freedom on vortex-induced vibration at low mass and damping. *J Fluid Mech* [Internet]. 2004 Jun 25;509:23–62. Available from: http://www.journals.cambridge.org/abstract_S0022112004008778
13. Dahl JM, Hover FS, Triantafyllou MS. Two-degree-of-freedom vortex-induced vibrations using a force assisted apparatus. *J Fluids Struct* [Internet]. 2006 Aug;22(6–7):807–18. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0889974606000466>
14. Aktosun E, Gedikli ED, Dahl JM. Observed wake branches from the 2-DOF forced motion of a circular cylinder in a free stream. *J Fluids Struct* 124;104035. Available from: <https://doi.org/10.1016/j.jfluidstructs.2023.104035>
15. Goushcha O, Elvin N, Andreopoulos Y. Interactions of vortices with a flexible beam with applications in fluidic energy harvesting. *Appl Phys Lett* [Internet]. 2014 Jan 13;104(2):021919. Available from: <http://aip.scitation.org/doi/https://doi.org/10.1063/1.4861927>
16. Ma X, Zhou S. A review of flow-induced vibration energy harvesters. *Energy Convers Manag* [Internet]. 2022 Feb;254:115223. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S019689042200019X>
17. Sensor F. ATI Industrial automation. 2023.
18. Xiros NI, Aktosun E. Stabilization of Neural Network Models for VIV Force Data Using Decoupled, Linear Feedback. *J Mar Sci Eng*. 2022 Feb;10(2):272.

Data-Driven Engineering



Simulation-Based Data-Driven Wind Engineering—Analyzing the Influence of Building Proximity and Skyways on Pedestrian Comfort

Knut Erik Teigen Giljarhus^(✉)  and Trond-Ola Hågbo 

Department of Mechanical and Structural Engineering and Materials Science, University of Stavanger, PB 8600, 4036 Stavanger, Norway

knut.e.giljarhus@uis.no

Abstract. Pedestrian wind comfort and safety is an important factor in urban city planning. Building structures and details can significantly impact the wind environment, causing undesirable accelerations of the wind close to the ground. Commonly used methods to analyze the wind environment include wind tunnel experiments and computational fluid dynamics simulations. Recently, machine learning techniques have also been proposed as an alternative method. These methods require a large amount of training data to build a prediction model, and existing simulation models can be suitable for the generation of such data. This work demonstrates a methodology to perform a large number of computational fluid dynamics simulations to generate training data for wind engineering problems. The methodology is then applied to analyze the influence of building proximity on pedestrian wind comfort. Additionally, the influence of a skyway, or pedestrian bridge, between the buildings is included. A comprehensive set of simulations is performed varying the surrounding building size and the skyway elevation, span and height. A hybrid approach is then applied, where a data-driven approach is used to identify dominating parameters and to build a fast predictive model, while the simulation data is used to gain further insight into the physical phenomena.

Keywords: Computational wind engineering · Building proximity · Skyway

1 Introduction

There is increased interest in machine-learning models to predict physical phenomena, due to their ability to cope with large amounts of data and fast prediction speed. In the design of urban environments, machine learning models are promising for interactive design where real-time feedback is important, and for generative design where a large number of designs must be evaluated quickly [1]. In the context of wind engineering and urban wind flow, machine-learning models have recently been applied to predict wind pressures on buildings [2, 3] and wind fields [4, 5].

With the rise of interest in machine learning models, there is also an increased need to develop methods for efficiently generating training data for these models. Computational

Fluid Dynamics (CFD) simulations can be a useful tool for generating training data for machine learning models in the context of urban wind flow. CFD simulations have successfully been used to investigate flows in urban environments [6–8]. Hence, they can be utilized to generate large datasets of wind flow behavior in an automated manner. Using simulations to generate training data can overcome some of the limitations of using real-world data, such as the cost and difficulty of obtaining accurate full-scale measurements in an urban environment. Furthermore, CFD simulations can provide a controlled and repeatable data source that can be used to train machine learning models with a high degree of confidence.

In this work, we illustrate how a CFD model can be utilized to efficiently generate a large amount of training data for machine learning models. As a case study for the methodology, we investigate the speedup in the gap between two buildings and the influence of a skyway between the buildings. It is well known that buildings in close proximity can cause the wind to be funneled into the gap between them, creating areas of turbulence and high wind speeds that can be uncomfortable for pedestrians [9, 10]. Skyways, which are elevated walkways that connect buildings (see Fig. 1), can also influence wind flow and pedestrian comfort. Skyways can alter wind speed and direction by acting as a barrier to the airflow. If located near the ground, they can accelerate flow underneath it and worsen the pedestrian comfort at ground level. Despite their prevalence, the impact of skyways has not been previously analyzed in the scientific literature.

Here, we use the proposed framework to simulate 625 combinations of building sizes without skyways and 2500 combinations with added skyways to quantify the impact the skyways have on the pedestrian-level wind speeds. We then use the generated data to build regression models for the maximum wind speed based on the building parameters to demonstrate one possible use of the data in machine learning models.



Fig. 1. Examples of skyways between buildings. © Knut Erik Teigen Giljarhus

2 Methods

2.1 Geometry Setup

The geometry setup is illustrated in Fig. 2. The buildings are cuboid with dimensions $W \times L \times H$, and the gap width between the buildings is denoted G . Note that symmetry is exploited so that only half the domain is simulated. The skyway is also a cuboid, with a constant width of 3 m and a height of S . The elevation from the ground level to the bottom of the skyway is denoted E . We choose five discrete values for each building parameter and two values for the skyway, as shown in Table 1. For the cases without skyway, this gives a total of $5^4 = 625$ unique cases. For the cases with skyway, the total number of cases is 2500.

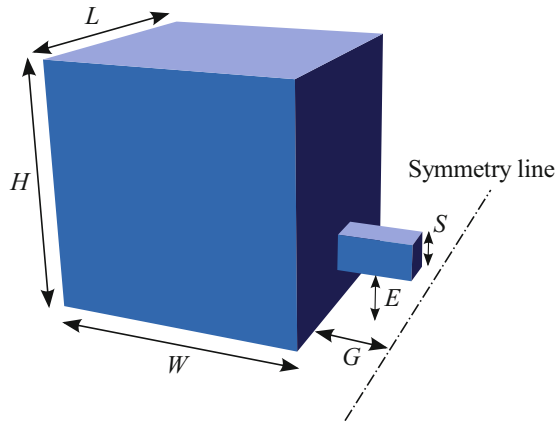


Fig. 2. Geometry setup with of building and skyway with parameters.

Table 1. Geometric parameters and values considered in this study.

Parameter	Symbol	Unit	Value
Building length	L	m	10, 20, 40, 60, 80
Building width	W	m	10, 20, 40, 60, 80
Building height	H	m	10, 20, 40, 60, 80
Gap width	G	m	6, 12, 18, 24, 30
Skyway elevation	E	m	3, 6
Skyway height	S	m	3, 6

2.2 Computational Setup

To simulate the fluid flow, the Reynolds-averaged, incompressible Navier-Stokes equations are solved,

$$\frac{\partial u_i}{\partial x_i} = 0 \quad (1)$$

$$u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial P}{\partial x_i} + \nu_{\text{eff}} \frac{\partial^2 u_i}{\partial x_j^2} \quad (2)$$

Here, u_i represents the averaged velocity components, P is the dynamic pressure and ρ the fluid density. The effective viscosity, ν_{eff} , is a combination of the fluid kinematic viscosity and the turbulent eddy viscosity. The turbulent eddy viscosity is found through the solution of two additional transport equations, with the so-called realizable k-epsilon turbulence model [11, 12]. We use the software package OpenFOAM v8 for the simulations [13].

We generate the computational mesh using the snappyHexMesh, a hex-dominant, unstructured mesh generator. Two levels of 1:2 refinement are used near the buildings, and two additional ones are used in the gap. Additionally, the mesh is graded in the vertical direction so that the first cell near the ground is 10 times smaller than the cell at the top of the domain. The total domain size is $900 \times 300 \times 300$ m. Figure 3 shows an image of the mesh near the building for one of the simulation setups. A logarithmic atmospheric boundary layer profile is prescribed at the inlet, with a reference velocity of 5 m/s 10 m above ground. A rough wall with a roughness factor of 0.05 is used for the ground, while a normal no-slip condition is used for the buildings. Slip boundaries are used for the sides and top of the domain.

The discretization scheme used for the convective term is a second-order central-upwind scheme. For the remaining Laplacian, gradient and surface normal gradient terms the linear schemes are used, which apply the Gauss divergence theorem with explicit non-orthogonal correction. The SIMPLE algorithm is used for pressure-velocity coupling, and the iteration procedure is run for 2000 iterations. The maximum velocity value in the region of interest between the buildings is monitored to ensure a convergent solution.

To enable efficient simulation of a large number of cases, we use an automated procedure, as illustrated in Fig. 4. OpenFOAM describes a simulation case as a number of text input files. Starting with a simulation template with the main parameters and settings for the problem, we then automatically generate a single case realization using the parameterization of the problem from Table 1. For the present problem, the input geometry must be modified, requiring remeshing for each case. In other parameterizations, for instance, looking at different wind directions, only boundary conditions need to be changed for each simulation. The proposed methodology is highly flexible in terms of how the problem is parameterized.

After each case has been set up, the simulations are performed in parallel, either through a cloud service such as Amazon AWS or a local computer cluster. For the simulations presented here, a cluster with 648 nodes is used. Each simulation is run in serial on a single core, due to the large number of simulations this is more efficient than parallel computing as it avoids the overhead of communication between the cores.

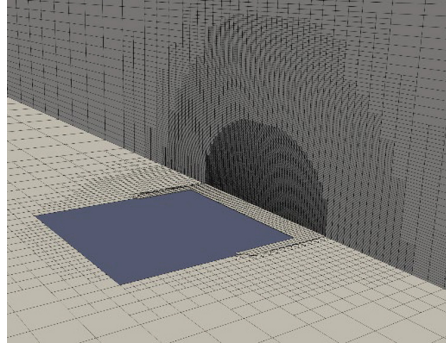


Fig. 3. Computational mesh showing refinement level near the building.

The final step of the methodology is to extract training data from the simulation. This procedure can be done either in the form of pictures, to be used in image-to-image translation algorithms, or in the form of tabular data. For the current problem, we extract the maximum velocity in the gap between the buildings for all cells up to 2 m above the ground (i.e. pedestrian level). This value is stored in a table along with the geometric parameters for the case.

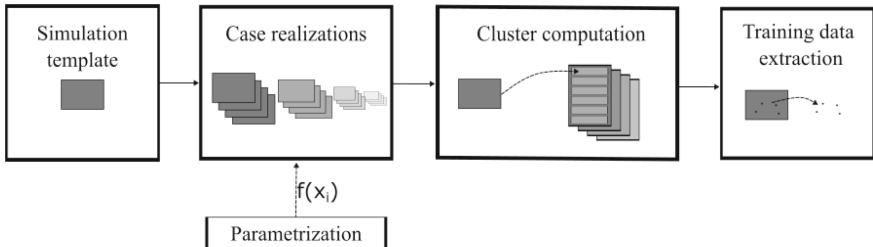


Fig. 4. Procedure to generate training data from simulation template and parameterization.

2.3 Regression Methods

Regression methods are a common technique used to investigate the relationship between variables and to build prediction models. Since this work focuses on the generation of training data, we use standard methods from the literature for the regression.

The most straightforward technique is linear regression, where linear approximations are used for predictions, and the least square distance to the training data is minimized. Other variations of this method are Bayesian ridge, elastic net, and support vector regression, which keep the assumption of linear relationships but use different techniques to calculate the weights.

A popular class of regression methods emerging in recent years is gradient-boosting regressors [14]. These methods work by building an ensemble of simple models, such

as decision trees, to make predictions. The idea is that combining the predictions of a series of weak models can make more accurate predictions. Gradient Boosting works by iteratively adding models to the ensemble and adjusting the weights of the models based on their performance. At each iteration, the model attempts to correct the errors made by the previous models by fitting a decision tree to the negative gradient of the loss function. The loss function is a measure of the difference between the predictions made by the current ensemble of models and the actual target values.

In this work, we use the gradient boosting method from the SKLearn library [15] as well as the XGBoost library [16], the LightGBM library [17] and the CatBoost library [18]. The models are trained on a randomly sampled subset of the simulation data, with the number of training samples set to 25% of the total simulated cases.

3 Results

3.1 Validation

An important part of the methodology is validation of the computational setup, to ensure the training data is of sufficient quality. Here, we compare simulations against experiments performed on buildings with side lengths of 20 m and a gap width of 6 m (at a wind tunnel scale of 1:400) [9]. The wind velocity was measured along the centerline in the gap 2 m above ground level. We used three different grid sizes for the comparison, where the base grid cell size was increased by a factor of 1.26 for each level, giving two times the number of total grid points. The medium mesh used for the remainder of the simulations contains approximately 150 000 grid points for the cases with a skyway. The simulation time was about 20 min per case.

Figure 5 compares simulated and experimental values. The simulated results follow the experimental values well for all grids. However, the difference between the coarse and medium grids is substantial, while the difference between the medium and fine grids is minor. Hence, we used the medium grid for the remainder of the simulations in this work.

3.2 Influence of Building Proximity on Wind Speed

To investigate the impact of a skyway, we first performed simulations on all geometries without a skyway added. Figure 6 shows streamlines for a case with buildings of side lengths 6 m and a 6 m gap between the buildings. We see how wind hitting the front of the building is pushed down and around the corner into the gap, leading to an acceleration of the velocity in the gap. The amplification factor is here defined as the ratio of the local cell velocity and the inlet velocity at 2 m above ground.

As an example illustrating all of the simulated data, Fig. 7 shows the amplification factor as a function of the square root of the frontal area for all 625 cases. Different colors and symbols are used for the gap width. At first sight, it appears to be a strong correlation between the frontal area of the building and the amplification of the wind. This is an intuitive observation, as a larger area will allow more wind to be collected and accelerated into the gap. However, for higher frontal areas, there is a sudden drop

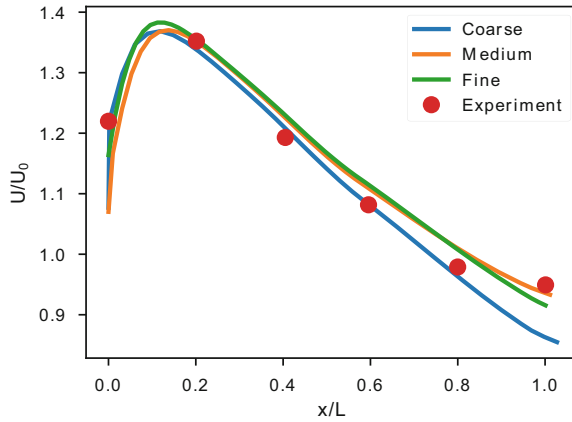


Fig. 5. Validation and mesh sensitivity. Comparison of velocity along a line through the center of the gap, at a height of $z = 2$ m, for different grid sizes and experimental results from [9].

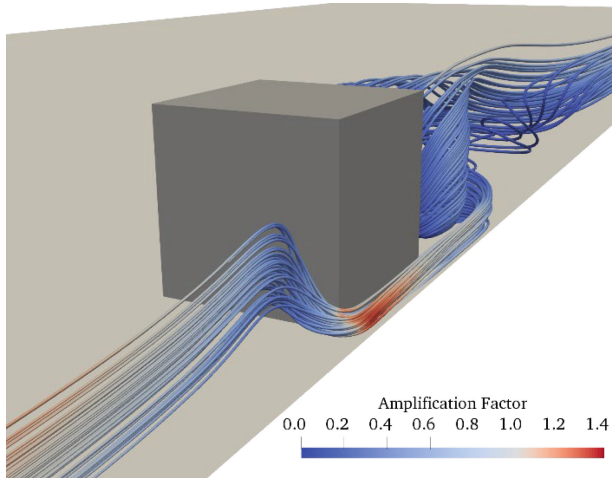


Fig. 6. Streamlines illustrating how wind is pulled down in front of the building and accelerated into the gap.

in amplification factor for the shortest gap width. For these narrow gaps relative to the building size, the flow resistance in the gap is high enough that more of the wind will instead pass around the building on the other sides, leading to less amplification inside the gap. This trend was also observed in other studies [9, 10].

3.3 Influence of Skyway on Wind Speed

Figure 8 shows the speedup compared to the case without skyway for all 2500 simulated cases. For this case, we also include the length of the building as a parameter for the reference length, as we found a greater influence of the length when a skyway was

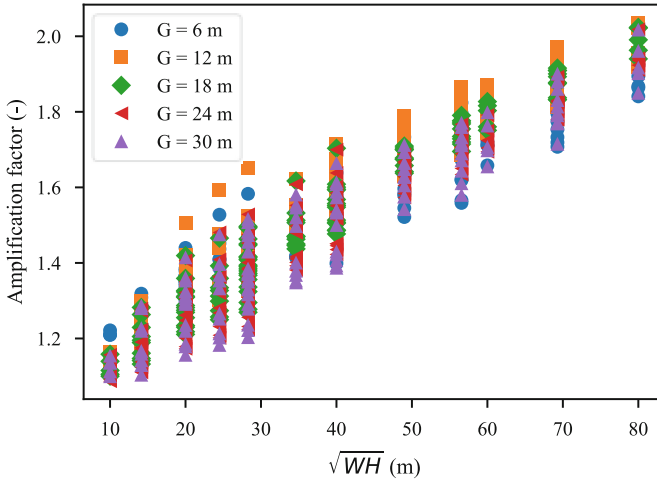


Fig. 7. Amplification factor as a function of squared frontal area for all 625 cases, colored by gap width.

included. Two important observations can be made from this figure. First, a skyway can *reduce* the maximum speed in the gap, not only increase it. Secondly, this effect is most prevalent for larger building sizes relative to the gap width.

For smaller buildings, the skyway primarily increases the wind velocity at ground level further, with speedups of up to 28%. Figure 9 shows one example of the flow with the skyway included. The plot shows velocity magnitude in a vertical slice through the center of the gap in the wind direction. For this case, the buildings dimensions are $10 \times 40 \times 12$ m and gap width the gap width is 30 m. For this large gap width and small building, there is almost no influence of the wind from the buildings and low flow resistance in the gap. Hence, the speedup becomes high as some of the wind hitting the skyway is forced underneath it toward the ground.

The opposite effect is shown in Fig. 10, where the skyway reduces the maximum wind velocity. Here, the gap width is small, and the building size is medium. The skyway acts as an additional flow resistance in the gap, pushing more of the wind above the gap.

3.4 Predictions Using Regression Methods

Finally, we illustrate one potential use of the training data generated from simulations. Since we have a parameterized geometry, and a single output in the form of maximum wind velocity in the gap, the problem lends itself to using regression methods for prediction. Figure 11 shows the predicted maximum velocity in the gap relative to the simulated velocity. A perfect match would lie on the solid black line, while the dotted lines indicate a $\pm 10\%$ deviation. The top row consists of linear regression methods, while the bottom row has the gradient boosting methods. There is a clear trend that the gradient-boosting methods perform better than linear methods. The best performance is found with the CatBoost algorithm, with a mean absolute error of 0.055 m/s and an R^2

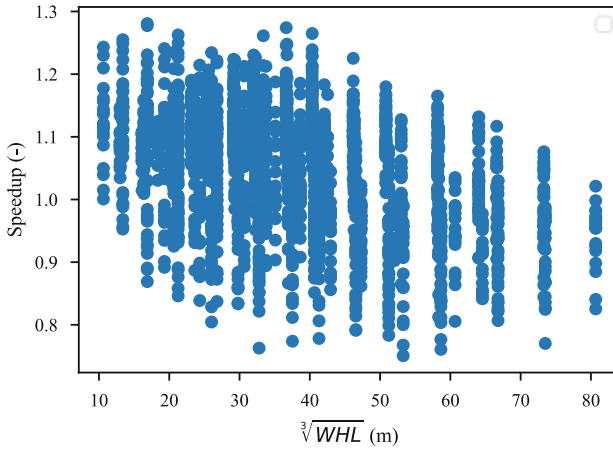


Fig. 8. Speedup of wind velocity in gap compared to the same case without skyway, for all 2500 simulated cases. Speedup is plotted as a function of the cube root of building volume.

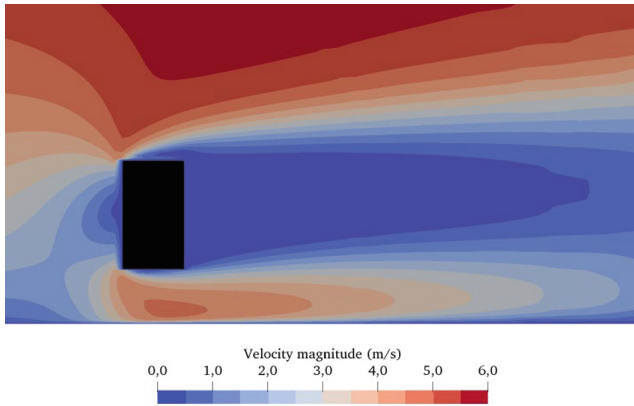


Fig. 9. Illustration of increase in velocity magnitude with skyway. Vertical contour plot of velocity magnitude in the center of the gap for building size $10 \times 40 \times 12$ m and gap width 30 m.

score of 0.991. However, it is important to be aware of potential pitfalls with the regression methods. Since the method was trained on just a few values for each parameter, for instance, just five discrete height values, it can behave in unexpected ways if used outside the training data. Figure 12 shows the amplification factor as a function of height with all other parameters kept constant. Instead of a gradual increase in amplification factor, as might be expected from a physical standpoint, the method predicts constant values around the heights it has seen in the training data with sharp jumps to the next level. This work aimed to demonstrate the methodology for generating training data from CFD simulations and highlight some of the influence of skyways on the pedestrian wind environment. If we want to design a predictive method that can produce smoother

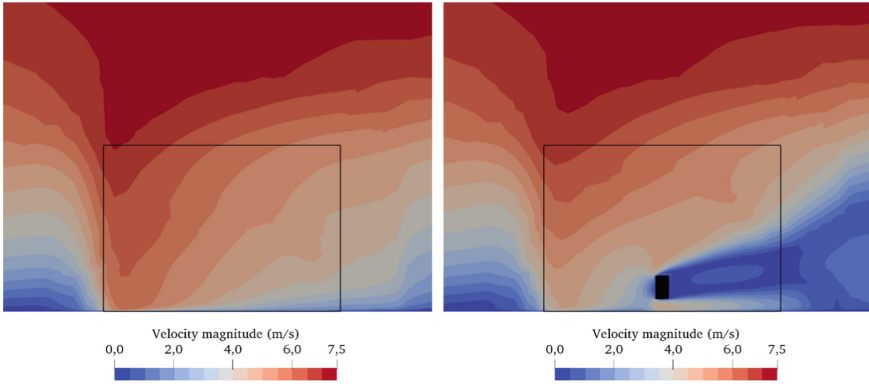


Fig. 10. Illustration of reduction in velocity magnitude with skyway. Vertical contour plot of velocity magnitude in the center of the gap for building size $60 \times 60 \times 42$ m and gap width 6 m, without skyway (left) and with skyway (right). The building location is indicated by black lines.

predictions, the input parameters should have a wide range of values, for instance, by random sampling instead of permutation of a fixed set of values (Table 2).

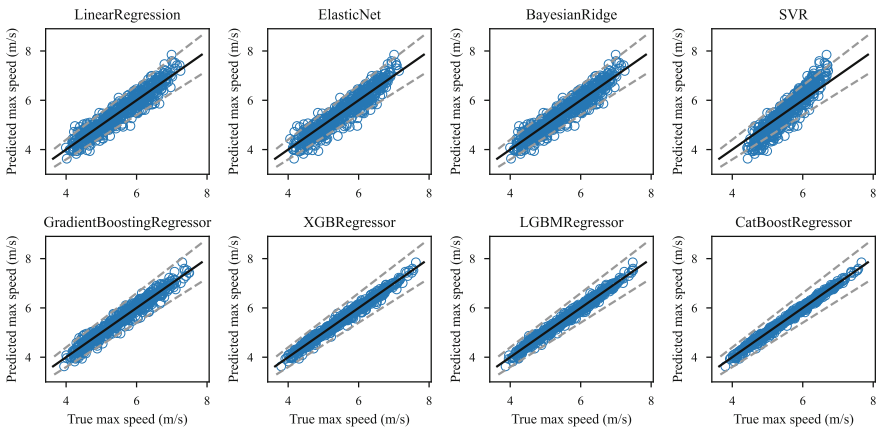
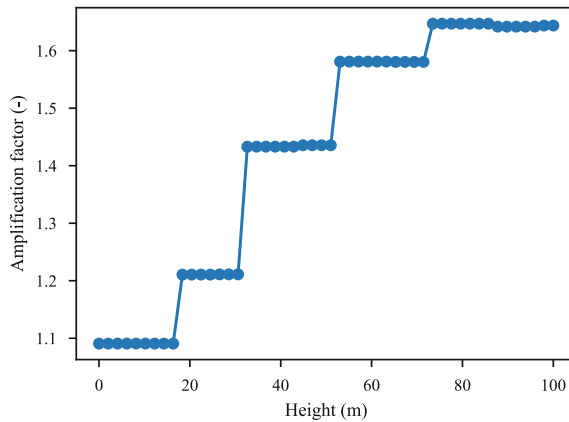


Fig. 11. Predicted max speed in gap between buildings compared to true max speed (from CFD simulations) for different regression methods for cases with skyway. The dotted lines indicate $\pm 10\%$ deviation from the true value.

Table 2. Errors for the different regression methods

Method	Mean absolute error (m/s)	R ² score (-)
Linear regression	0.201	0.898
Elastic net	0.210	0.887
Bayesian ridge	0.201	0.898
SVR	0.278	0.810
Gradient boosting	0.134	0.953
XGBoost	0.074	0.984
LGBM	0.078	0.983
CatBoost	0.055	0.991

**Fig. 12.** Predicted amplification factor as a function of height using CatBoost, for a building with width 12 m, length 80 m, gap width 12 m, skyway elevation 3 m and skyway height 3 m.

4 Conclusion

We presented a methodology for a data-driven approach to wind engineering, using an efficient approach for generating and running a large number of CFD simulations. The methodology was applied to analyze the speedup of wind between buildings of close proximity and the presence of a skyway between the buildings. We demonstrated some unexpected features in the data set, such as reduced wind amplification when the gap is small compared to the building size, and that a skyway can reduce the maximum wind speed. We also showed that regression methods could be built to give a fast prediction of maximum speed as a function of building parameters, but that careful setup of the training data is important to obtain a viable predictive model. The method is currently being extended to more complex data sets with multiple buildings and to predict more parameters than just the maximum wind speed.



References

1. Krish, S. A practical generative design method. *Computer-Aided Design*, 43(1), 88-100 (2011).
2. Huang, J. X., Li, Q. S., & Han, X. L. Recovery of missing field measured wind pressures on a supertall building based on correlation analysis and machine learning. *Journal of Wind Engineering and Industrial Aerodynamics*, 231 (2022).
3. Meddage, D. P. P., Ekanayake, I. U., Weerasuriya, A. U., Lewangamage, C. S., Tse, K. T., Miyanawala, T. P., & Ramanayaka, C. D. E. Explainable Machine Learning (XML) to predict external wind pressure of a low-rise building in urban-like settings. *Journal of Wind Engineering and Industrial Aerodynamics*, 226 (2022).
4. Lamberti, G., & Gorié, C. A multi-fidelity machine learning framework to predict wind loads on buildings. *Journal of Wind Engineering and Industrial Aerodynamics*, 214 (2021).
5. Diop, M., Dubois, P., Toubin, H., Planckaert, L., Le Roy, J. F., & Garnier, E. Reconstruction of flow around a high-rise building from wake measurements using Machine Learning techniques. *Journal of Wind Engineering and Industrial Aerodynamics*, 230 (2022).
6. Blocken, B., Stathopoulos, T., & Van Beeck, J. P. A. J. Pedestrian-level wind conditions around buildings: Review of wind-tunnel and CFD techniques and their accuracy for wind comfort assessment. *Building and Environment*, 100, 50-81. (2016).
7. Hågbo, T. O., Giljarhus, K. E. T., & Hjertager, B. H. Influence of geometry acquisition method on pedestrian wind simulations. *Journal of Wind Engineering and Industrial Aerodynamics*, 215 (2021).
8. Hågbo T. O. & Giljarhus K. E. T. Pedestrian Wind Comfort Assessment Using Computational Fluid Dynamics Simulations With Varying Number of Wind Directions. *Frontiers of the Built Environment*, 8 (2022).
9. Stathopoulos, T., & Storms, R. Wind environmental conditions in passages between buildings. *Journal of Wind Engineering and Industrial Aerodynamics*, 24(1), 19-31 (1986).
10. Blocken, B., Carmeliet, J., & Stathopoulos, T. CFD evaluation of wind speed conditions in passages between parallel buildings—effect of wall-function roughness modifications for the atmospheric boundary layer flow. *Journal of Wind Engineering and Industrial Aerodynamics*, 95(9-11), 941-962 (2007).
11. Shih, T. H., Liou, W. W., Shabbir, A., Yang, Z., & Zhu, J. A new $k-\epsilon$ eddy viscosity model for high Reynolds number turbulent flows. *Computers & fluids*, 24(3), 227-238 (1995).
12. Blocken, B., Roels, S., Carmeliet, J. Pedestrian wind conditions in passages through buildings—Part I. Numerical modeling, sensitivity analysis and experimental verification. Research report, Laboratory of Building Physics, Catholic University of Leuven (2003)
13. Weller, H. G., Tabor, G., Jasak, H., & Fureby, C. A tensorial approach to computational continuum mechanics using object-oriented techniques. *Computers in physics*, 12(6), 620-631 (1998).
14. Bentéjac, C., Csörgő, A., Martínez-Muñoz, G. A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review* 54, 1937-1967 (2021).
15. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
16. Chen, T., Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794. San Francisco, USA (2016).

17. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Qiwei, Y., Liu, T. Y. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems* 30 (2017).
18. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., Gulin, A. CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems* 31 (2018).



Anatomical Model Reconstruction (Solid Modeling) Using a Reverse Engineering Approach

Yosef W. Adugna^(✉) , Navaneethan S. Kurukkal, and Hirpa G. Lemu 

Department of Mechanical and Structural Engineering and Materials Science, University of Stavanger, Stavanger, Norway
yosef.w.adugna@uis.no

Abstract. In addition to the traditional applications in computer-aided systems such as machining and finite element analysis, solid modeling is nowadays utilized for various purposes in medical applications such as the design of implants and prostheses or quality control, verification, and analysis purposes. This is mainly because solid models provide properties such as mass and volume and can be used in mechanical assemblies. In the case of either designing or reconstructing medical applications, certain techniques are used to account for the missing information. This is not only due to insufficient scanned data but also due to missing information in cases where there are defective organs or bones, such as the anatomical reconstruction of defective skulls. Common techniques involved are mirroring, surface interpolation, deformed template, and slice-based reconstruction. In medical applications for designing and reconstruction purposes, the techniques that have been mentioned are surface fitting, contour skinning, volume polygonization, and implicit-function interpolation. In this study, the basis for solid modeling is an STL file that was obtained from 3D scanning as a reference. The Human Femur bone model is constructed from a DICOM file, which is converted to a 3D model, exported as a.stl file, and then printed. The model is 3D printed on Stratasys fused deposition modeling 3D Printer and scanned back using a handheld 3D Scanner. In this work, FreeCAD software was used for solid modeling. FreeCAD is commercially available open-source software that can be used for parametric 3D modeling, mesh manipulation, and repair, finite element analysis tools, robot simulation modules, and computer numerically controlled workbenches. The method proposed in this study is direct lofting of cross-sections of the mesh and taking advantage of the FreeCAD workbench to convert individual lofts to solids. Thereafter, a single solid model was created using Boolean functions.

Keywords: Solid modeling · freeCAD · Anatomical model

1 Introduction

Reverse Engineering through 3D Scanning is a technique for producing 3D data in the form of a triangle mesh or point cloud. As computer-aided design and manufacturing (CAD/CAM) technology advances, building geometric models of existent items plays an

essential role in reverse engineering, particularly when the prototype is built or modified on the shop floor, and the CAD model is lacking [1]. Additionally, reverse engineering (RE) is crucial when there is a need to recreate objects without CAD models, when parts already in production need to be updated by redesigning them to make up for any production-related flaws and to optimize the production process.

3D Scanning is highly capable of obtaining 3D measurements without physical contact and with high accuracy [2]. Although RE techniques vary from field to field, they are all developed with advances in computer and digital control throughout the previous century. In the past, manual measuring equipment has been used to measure diameters, depths, and other parameters, and manual design creation has also been practiced [3]. Currently, even though the choice of model reconstruction method is highly dependent on the specific application, the accuracy required for the final model, and the complexity of the object, several reconstruction methods can be used. Though there are various methods of solid modeling in RE, their working principle is the same i.e., using multiple sensors, such as cameras, lasers, or structured light, the scanner collects data about the object's surface. After processing and combining the data, a digital 3D model is created.

Reverse engineering has a broad application in many different industries, enabling the development of new designs, the improvement of existing ones, and the preservation of historical and cultural monuments. Among others, the range of the applications of RE includes aerospace and defense, automotive Industry, architecture, arts and entertainment, cultural heritage preservation, and medical industries [4, 5]. Langbein et al. [6] classified the applications of solid model reconstruction applications into three (3) categories: (i) reproduction; (ii) quality control and (iii) redesign and modification applications.

In the medical industry, the ability to generate 3D models from patient data using 3D Scanners allows the creation and design of preoperatively suitable customized applications like gloves, body shape, implants, prosthetics, and anatomical models' surface texture, restoring shape and color safely. It is also used to scan body structures for replication and characterization, estimation of bleeding in hemophilia, abdominal shape characteristics, and lung functions to help measurement of parameters like surface area, distances, and volume [7]. Some of the applications of reverse engineering in the medical field include use in dentistry [8] and dental implantology [9], prosthetic design [10, 11], organ fabrication [12], forensic anthropology [13], biomechanics [14, 15], neurosurgery [16], surgical planning [17], anatomical reconstruction of different organ parts [18–21] and others. This method provides a revolutionary advance in surgical planning and simulation based on patient-specific anatomy. The typical solid model reconstruction process from RE involves data acquisition, point cloud processing, surface modeling, solid modeling, and verification. When used in medical applications, these RE steps in general can be categorized as depicted in Fig. 1.

In medical applications, a segmentation is first reconstructed into a “watertight” surface model including surface area and volume information, and then translated to STL file format for further analysis. Several methods are utilized to design or reconstruct missing parts to account for the lack of information in medical field. This is not owing to a lack of scanned data, but rather to a lack of information regarding the presence of defective internal organs, bones, etc. One such study is mentioned in [22], where

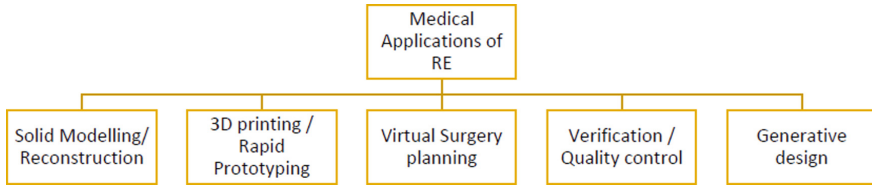


Fig. 1. Classification of medical applications of RE

current techniques are highlighted for the anatomical reconstruction of defective skulls. The methods described in this work include surface interpolation, deformed templates, mirroring, and sliced-based reconstruction. Surface fitting, contour skinning, volume polygonization, and implicit-function interpolation are generally the approaches that have been suggested for medical applications for designing and reconstructive purposes [23].

2 Solid Modeling Methods and Procedures

Solid modelling can be done in a variety of ways utilizing the STL reference file. It is possible to convert STL files to solid models [8] using commercial software, such as those from Autodesk (Inventor and Fusion360) or Geomagic. Another choice is to construct surfaces using the STL as a guide and then transform the surface model into a solid model using the same program. Alternately, one may generate sections and surfaces using part design tools based on the STL model, then turn that into a solid model. In this study, the development of a solid model of the femur using FreeCAD is suggested as a technique. The study done by Yang et al. [24] served as an inspiration for the procedure, who developed a surface model initially using the commercial program CATIA, which is then transformed into a solid model.

In this study, FreeCAD was used. It is an open-source software for parametric 3D modelling, mesh manipulation and repair, FEA tools, robot simulation modules, and CAM/CNC workbenches [25]. The approach involved directly lofting mesh cross-sections and using the FreeCAD workstation to turn each loft into a solid. Thereafter, a single solid model was created using Boolean functions following the procedure (steps) described below.

Step 1—Importing STL file into FreeCAD: FreeCAD provides the importation of STL files for manipulation, and it has an entire workbench for mesh design and repair as described in Fig. 2.

Step 2—Creating mesh cross sections: The cross-section tool, which generates a polyline drawing containing the intersection points of the model, is a useful feature in the mesh design workbench. The number of sections, their positions, and their tolerance for connectivity are the variables offered. Figure 3 shows that the sketch is composed of several points connected by straight lines. Hence, the sketch is composed entirely of simply edges and points. The several sections of the femur model can be visualized with this tool, as shown in Fig. 4.

Step 3—Creating sketches for lofting: The next step is to efficiently loft between these cross-sections using sketches that are created. It is possible to loft using the polyline

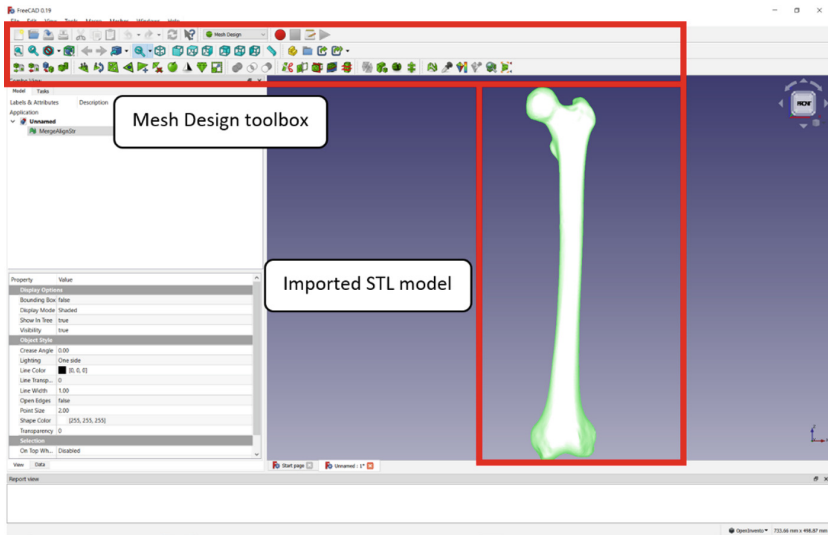


Fig. 2. Workbench for mesh design and imported model in FreeCAD

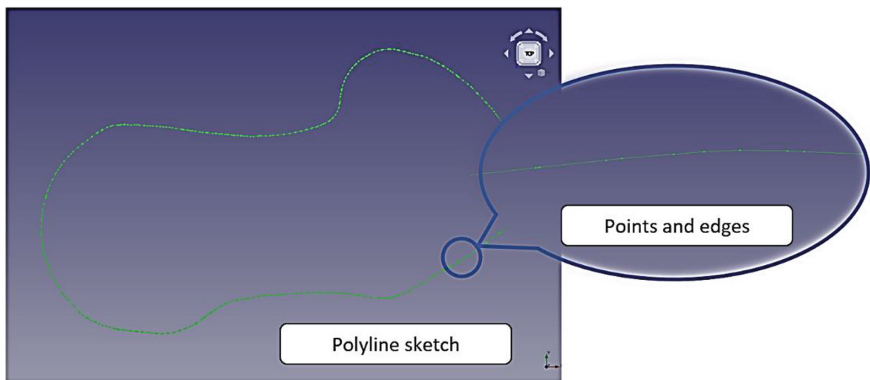


Fig. 3. Polyline sketch

sketch. Nevertheless, as was already indicated, this sketch has numerous edges, making it more difficult to compute the lofting. This method of conducting individual lofts of each part takes time as well. Figure 5 depicts lofts made using the polyline consisting of numerous surfaces, which can be seen in the wireframe model. It is observed that the interpolation between surfaces is incorrect, most likely due to the algorithm that tried to loft each edge to a corresponding edge in the next sketch. This choice is advantageous for simpler shapes if it is possible to reduce the curve points number without compromising geometric accuracy (in such cases, lofting requires computing fewer points).

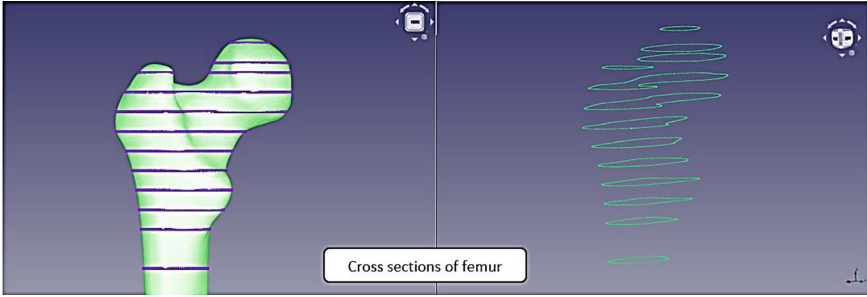


Fig. 4. Cross-sections of the femur

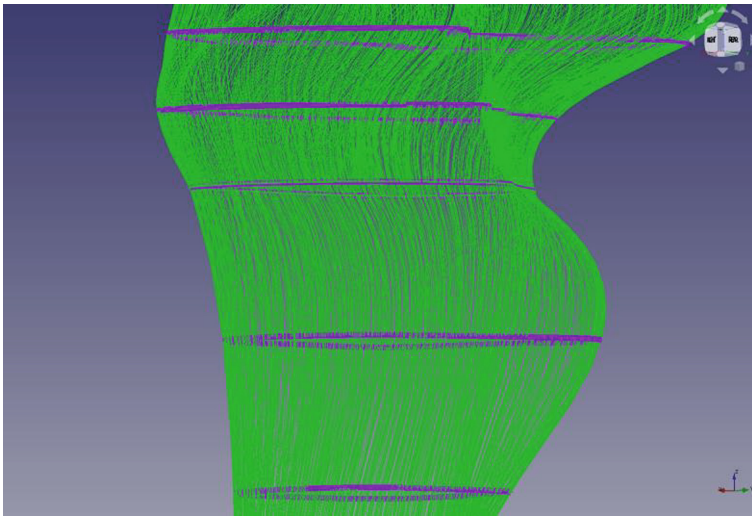


Fig. 5. Loft created using polyline (wireframe representation)

3 Results and Discussions

3.1 Results of Interpolating Curves

A viable solution to this issue is to change the polyline drawing into a curved closed sketch with a single edge, which will reduce the number of surfaces that need to be interpolated during lofting. To interpolate the existing points in these curves, many options including Bezier, Cubic Bezier, and NURBS curves will be examined. The drafts of the workbench model in FreeCAD contain tools for creating the abovementioned curves.

Bezier Curves: The vertices of a closed polygon are used as a reference to generate a Bezier curve [26]. The curve's degree is lower than the quantity of control points that were used to create it. Just the first and last points of the polygon are traversed by the curve, which otherwise follows the tangent of its other sides. Figure 6 displays the preliminary drawing produced using the Bezier curve. This was made by selecting points in combination with the sketch. The images show that the curve does not match to

the profile of the polyline sketch. However, a sketch with a closed loop was constructed. After constructing the closed loop, the Bezier curve's control points were adjusted to match the polyline sketch's profile, as illustrated in Fig. 7. Though Fig. 7 seems to be more accurate than Fig. 6, it does not match the polyline's profile.

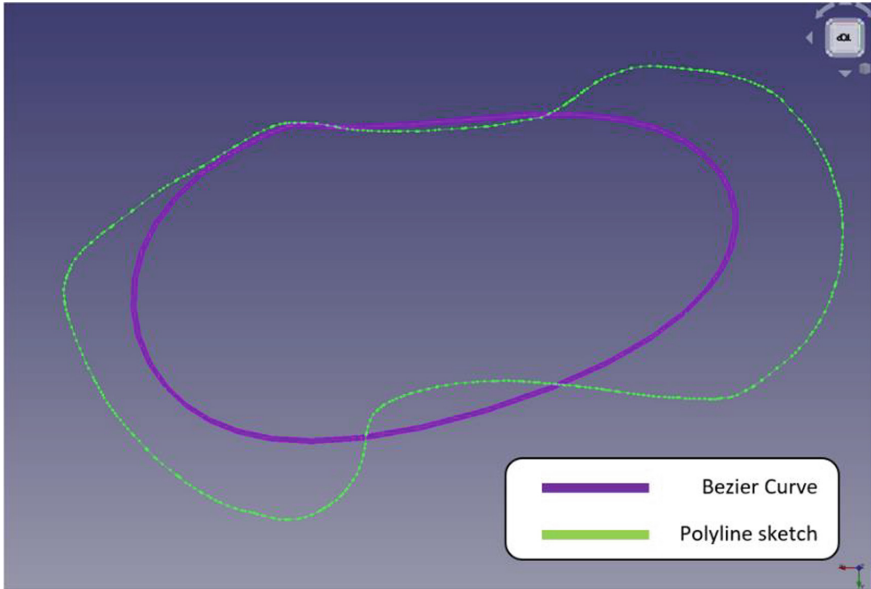


Fig. 6. Sketch using Bezier curve (before adjustments)

Cubic Bezier curves: Bezier curves of order three that have four control points are referred to as cubic Bezier curves. Making continuous splines out of several degrees, three Bezier curves are possible with the help of this FreeCAD tool. Figure 8 shows that while the cubic Bezier curve does provide a closer depiction of the design than the polyline sketch, the segments only produce a very rough outline. This was created by selecting points along with the sketch.

B-spline curves: B-spline curves, which use multiple Bezier curves connected end to end, are improvements of Bezier curves [12]. Local adjustments are possible with B-spline curves, and the degree of the curve is unaffected by the quantity of control points. Figure 9 illustrates how well the sketch is represented by the cubic B-spline curve and how well the segments match the polyline sketch's outline. This was done by choosing points and adding them to the sketch. This was created by selecting points along with the sketch.

Creating sketches using the method is a manual process. An efficient alternative will be to create the curves automatically using the points on the polyline sketch. An effective substitute would be to automatically generate the curves using the points on the polyline sketch. To do this, interpolation between the points on the polyline sketch was done using the points as a references. Tools for interpolating between points are included in FreeCAD's curves workbench. There are three possible approaches: (1) joining all

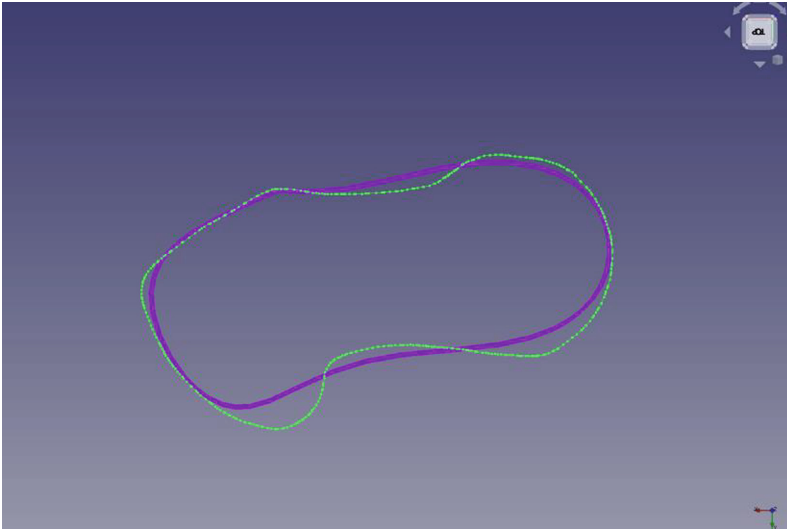


Fig. 7. Sketch using Bezier curve (after adjustments)

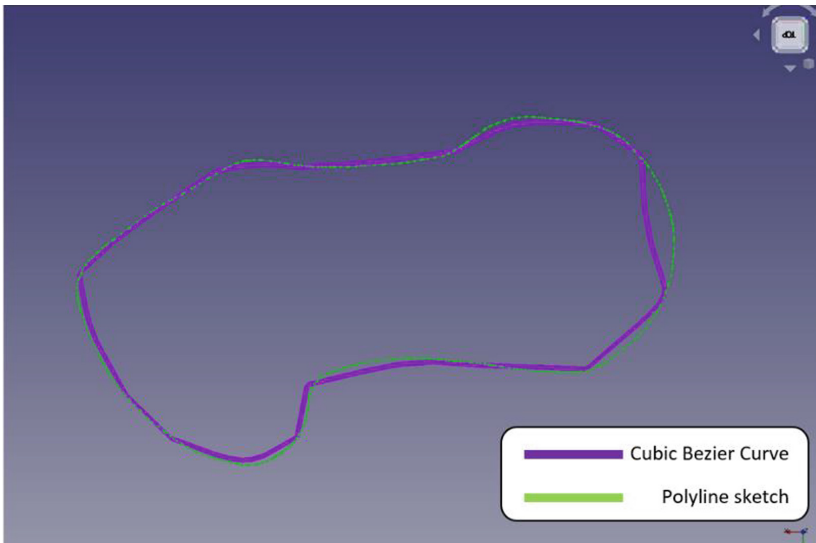


Fig. 8. Sketch using cubic Bezier curve

edges with B-Spline curves, (2) approximating points with a NURBS curve, and (3) interpolating between points using B-Spline curves.

NURBS curve: A non-uniform rational B-spline curve is like a non-uniform B-spline curve but uses more blending function parameters. A NURBS curve can be modified in more varied ways than a B-spline curve. Moreover, NURBS curves may accurately depict conic curves, but B-splines can only approximate them [12]. Yet, the FreeCAD

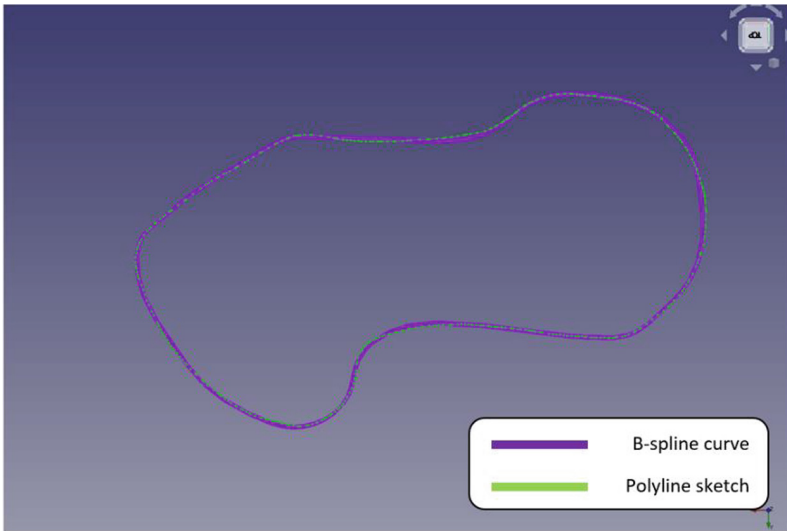


Fig. 9. Sketch using B-spline curve

program generates a roughly NURBS curve from the coordinates. This tool might not be entirely effective because it needs modifications, including closing the loop and altering the curve's maximum degree, to produce a smooth closed loop free of oscillations. Figure 10 displays the initial curve and the subsequent revisions.

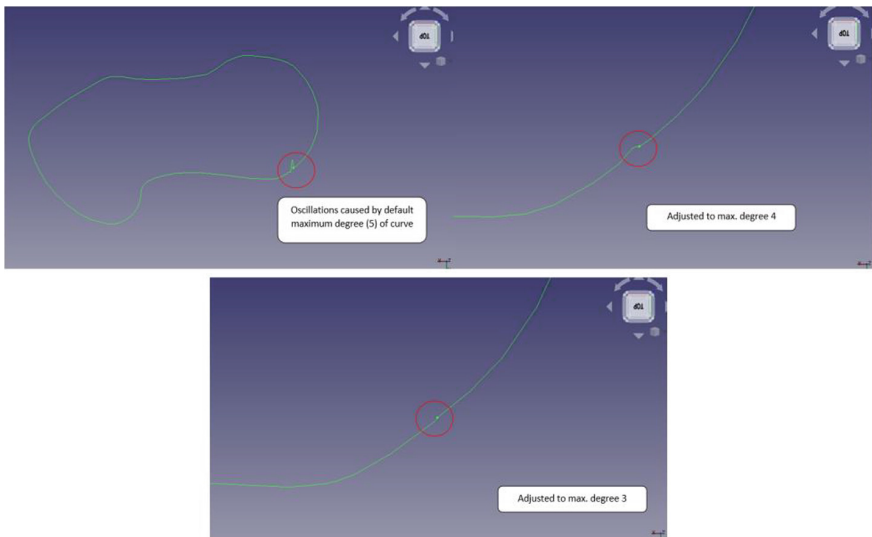


Fig. 10. Sketch created using NURBS curve

Interpolation using B-splines/Join using B-splines: There are only two differences between these two functions. There is no closed loop created by the interpolation function. A closed-loop sketch requires an additional step of setting the parameter “periodic” to “true.” The curve created by the join tool is not smooth. Furthermore, the tools can be used to generate a single closed-loop sketch from all the points in the polyline sketch as illustrated Fig. 11.

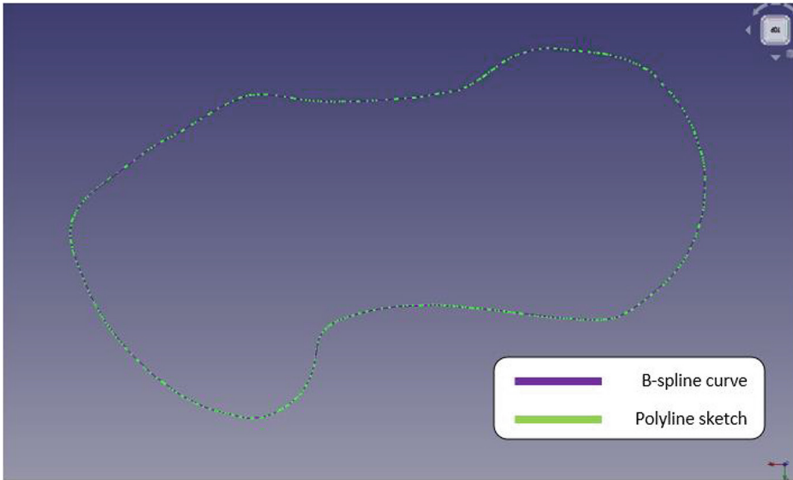


Fig. 11. Sketch created by joining points using B-Spline curves

After completing these three procedures, a reliable model of the femur can be produced. However, since the entire femur model need not be divided into sections in the same plane, splitting the mesh while modelling may be more practical. Three sections were separated from the femur mesh model (Fig. 12a). This can be done with the trim mesh tool, which splits a mesh model into polygons. After that, the previously specified first through third processes were carried out. Using the interpolation using points method, each mesh component was divided into smaller pieces and transformed into a closed-loop sketch (Fig. 12b). Each segment was lofted to construct a solid model once enough interpolation curves had been collected (Fig. 12c).

Among the sketches, B-spline surfaces are used in FreeCAD’s loft function. The number of sketches affects the interpolation. The interpolation is of degree three for more than ten drawings and one less than the number of sketches for less than ten sketches. As a result, several solid lofts were produced using these profiles and then combined using FreeCAD’s union tool. A more robust approach will be to apply Boolean functions to combine each loft. This is a simple process in FreeCAD. The first step is to change the solid feature from “false” to “true” in the data tab. Each loft can be combined using the Boolean function “union” in the part workbench.

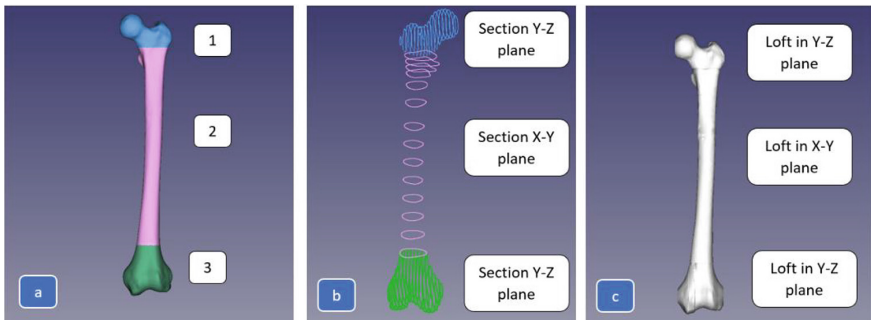


Fig. 12. The full procedure for split mesh model of the femur

3.2 Solid Modelling Results

This section compares the solid model produced by Autodesk Inventor using the automated conversion feature to the solid model produced by the method described above. The STL file from the Handyscan 700 3D Scanner was used to generate a solid model of the femur using the above-mentioned technique. Most of the geometry was not included in the original model that was produced. Thus, as shown in Fig. 13 (left), the femur model was divided into even smaller portions. This greatly facilitated the capture of the femur's characteristics. For solid modelling in the CAD program, the Handyscan model was selected. This was necessary to utilize as a benchmark for evaluating the final model built using the technique suggested in the study.

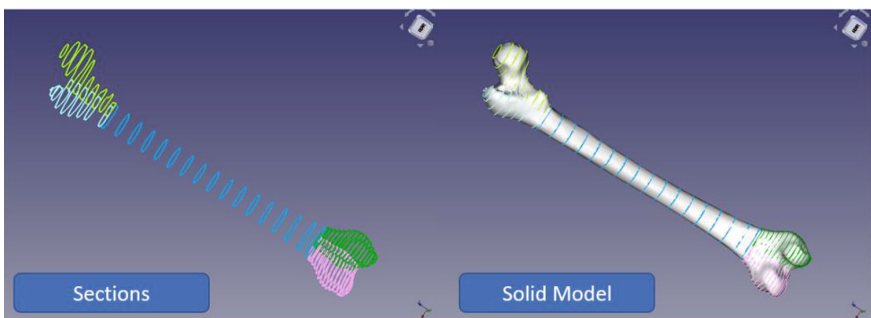


Fig. 13. Final solid model of the femur

Figure 14 shows how the distance from the reference mesh filter was applied to the solid model. There are a few deviations between the intended model and the reference STL file. Yet they are not important. The FreeCAD model was exported as a STEP file and loaded into Autodesk Inventor to ascertain the solid model's attributes. Both the developed (Fig. 15—Solid model using FreeCAD) and the converted models (Fig. 16—model from Handyscan 700 converted directly to solid in Autodesk Inventor) assigned Titanium material properties to compare the solid models' properties. The results are given below in Table 1, Figs. 15 and 16.

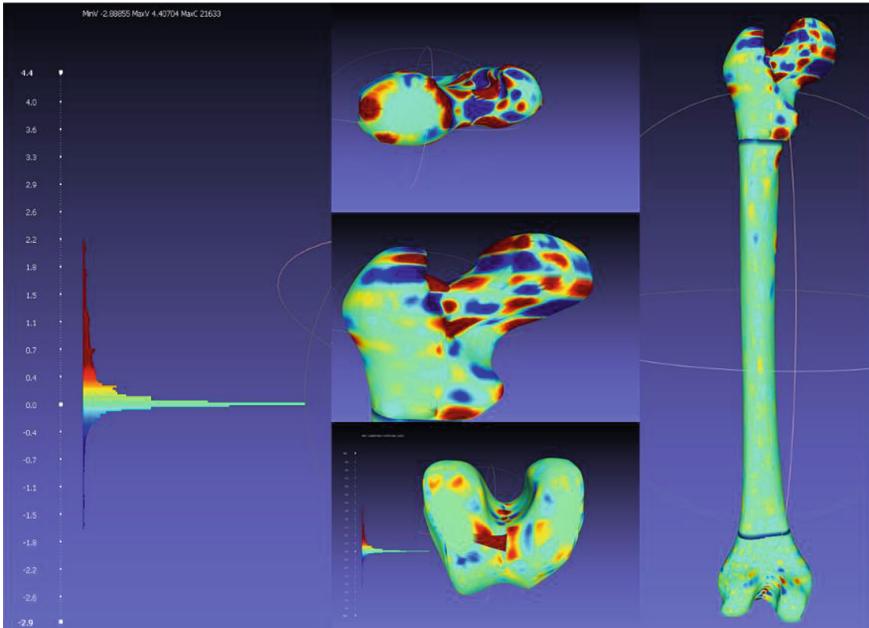


Fig. 14. Final solid model versus reference model

Table 1. FreeCAD versus Autodesk model

	FreeCAD model	Autodesk model
Weight	2.564 kg	2.566 kg
Volume	568485.596 mm ³	568861.768 mm ³
Area	66471.137 mm ²	66088.332 mm ²

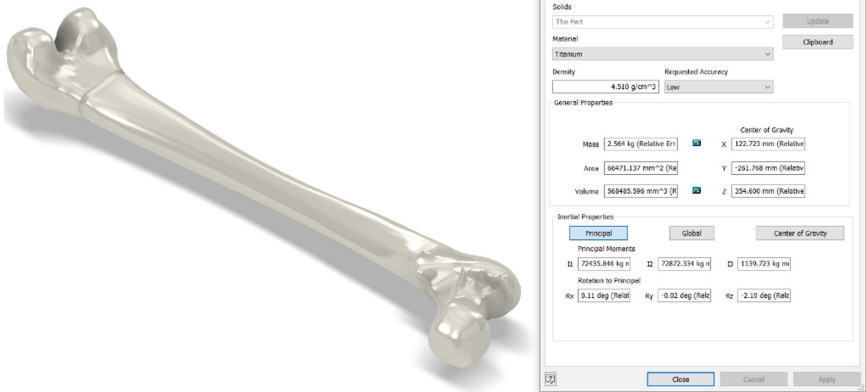


Fig. 15. Solid model using FreeCAD

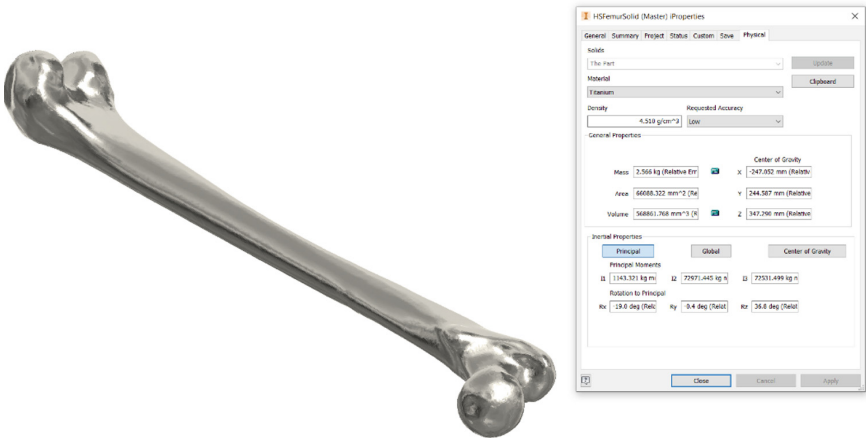


Fig. 16. Handscan 700 model converted directly to solid in Autodesk Inventor

4 Conclusion

According to the findings, the method that was provided for obtaining a solid femur model is beneficial since it has features that are comparable to those of the model that was directly converted. For some types of analyses, such as FEA, for example, this kind of model is necessary. Consequently, the proposed method can be a viable choice to use open-source software to generate solid models with an organic shape using a reference STL file. The solid modeling technique that was proposed using the open-source program FreeCAD proved to be quite effective as the results from this process matched the results from a file that had been directly converted to an STL format. The process of RE was effectively applied to the solid modeling of an anatomical structure (the femur), which was based on this method.


References

1. Wang, J., et al., A framework for 3D model reconstruction in reverse engineering. *Computers & Industrial Engineering* 63(4), 1189-1200 (2012).
2. Chromy, A. and L. Zalud, Robotic 3D scanner as an alternative to standard modalities of medical imaging. *SpringerPlus* 3, 1-10 (2014).
3. Helle, R.H. and H.G. Lemu, A case study on use of 3D scanning for reverse engineering and quality control. *Materials Today: Proceedings* 45, 5255-5262 (2021).
4. Geng, Z. and B. Bidanda, Review of reverse engineering systems—current state of the art. *Virtual and Physical Prototyping* 12(2), 161-172 (2017).
5. Wang, W., *Reverse engineering: Technology of reinvention*, Crc Press (2010).
6. Langbein, F.C., A.D. Marshall, and R.R. Martin, Choosing consistent constraints for beautification of reverse engineered geometric models. *Computer-Aided Design* 36(3), 261-278 (2004).
7. Chromy, A., Application of high-resolution 3D scanning in medical volumetry. *International Journal of Electronics and Telecommunications* 62(1), 23-31 (2016).
8. Ghafoor, H., Reverse engineering in orthodontics. *Turkish Journal of Orthodontics* 31(4), 139 (2018).
9. Kernen, F., et al., Accuracy of three-dimensional printed templates for guided implant placement based on matching a surface scan with CBCT. *Clinical implant dentistry and related research* 18(4), 762-768 (2016).
10. Rhyne, B.J., et al. Reverse engineering a transhumeral prosthetic design for additive manufacturing. In *International Solid Freeform Fabrication Symposium*, University of Texas at Austin (2017).
11. Babić, M., et al., Reverse engineering-based integrity assessment of a total hip prosthesis. *Procedia Structural Integrity* 13, 438-443 (2018).
12. Macko, M., et al. The method of artificial organs fabrication based on reverse engineering in medicine. In *Proceedings of the 13th International Scientific Conference, Computer Aided Engineering Springer* (2017).
13. Reynolds, M.S., et al., Standardized anthropological measurement of postcranial bones using three-dimensional models in CAD software. *Forensic science international* 278, 381-387 (2017).
14. Elsayyad, A.A., et al., Biomechanics of 3-implant-supported and 4-implant-supported mandibular screw-retained prostheses: A 3D finite element analysis study. *The Journal of Prosthetic Dentistry* 124(1), 68. e1-68. e10 (2020).
15. Ingrassia, T., et al., Biomechanical analysis of the humeral tray positioning in reverse shoulder arthroplasty design. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 12, 651-661 (2018).
16. Kinoshita, M., et al., Reverse Engineering Glioma Radiomics to Conventional Neuroimaging. *Neurologia medico-chirurgica* 61(9), 505 (2021).
17. Abouel Nasr, E., et al., A digital design methodology for surgical planning and fabrication of customized mandible implants. *Rapid Prototyping Journal* 23(1), 101-109 (2017).
18. Chantarapanich, N., et al., 3D CAD/reverse engineering technique for assessment of Thai morphology: Proximal femur and acetabulum. *Journal of Orthopaedic Science* 22(4), 703-709 (2017).
19. Carfagni, M., et al., Fast and low cost acquisition and reconstruction system for human hand-wrist-arm anatomy. *Procedia Manufacturing* 11, 1600-1608 (2017).
20. Furferi, R., et al. 3d acquisition of the ear anatomy: A low-cost set up suitable for the clinical practice. in *XV Mediterranean Conference on Medical and Biological Engineering and Computing—MEDICON 2019: Proceedings of MEDICON 2019 Coimbra, Portugal*. Springer (2020).

21. Zhao, G., et al., Reconstruction of the three-dimensional model of cervical vertebrae segments based on CT image and 3D printing. *Zhongguo yi Liao qi xie za zhi= Chinese Journal of Medical Instrumentation* 43(6), 451–453 (2019).
22. Buonamici, F., et al., Reverse engineering techniques for virtual reconstruction of defective skulls: an overview of existing approaches. *Computer-Aided Design and Applications* 16(1), (2018).
23. Mun, D. and B.C. Kim, Three-dimensional solid reconstruction of a human bone from CT images using interpolation with triangular Bézier patches. *Journal of Mechanical Science and Technology* 31, 3875-3886 (2017).
24. Yang, B., et al. Three-dimensional solid reconstruction of femoral CT images based on reverse engineering. In *Journal of Physics: Conference Series* IOP Publishing (2019).
25. FreeCAD: Your own 3D parametric modeler. [Online]. Available: <https://www.freecadweb.org/> (accessed: March. 26 2023).
26. Lee, K., *Principles of CAD*. Reading, Mass.: Addison-Wesley, (1999).



Damage Assessment of Crack Growth Through Numerical Modelling and Rain Flow Counting Method

Yohannes Regassa¹(✉) and Hirpa G. Lemu² 

¹ Addis Ababa Science and Technology University, Addis Ababa, Ethiopia
yohannes.regassa@aastu.edu.et

² University of Stavanger, N-4036 Stavanger, Norway

Abstract. The performance and behavior of nearly all in-service structures can be affected by the damage resulting from external conditions such as weather, abrasion, loading condition, wear, operator misuse, or safety negligence that lead to catastrophic failure of the structures. Therefore, fatigue failures that resulted from crack growth and have a potential to cause collapse of the structures have to be managed earlier through quantifying the crack growth, for instance using the stress intensity factor (SIF). The SIF approach can calculate the stress level at the crack tip by modeling crack phenomenon under simple loading at small discrete steps using an appropriate Finite Element method with compatible software tools. In the study reported in this article, the displacement extrapolation method has been employed to calculate the SIF and rain-flow counting method is used for the analysis of fatigue data to calculate the life cycle of modeled crack growth under varying stress. The calculated stress intensity factor is useful to predict the crack propagation of the structure by the global and local detection approaches using the cycle-by-cycle integration methods. ABAQUS software and MATLAB were applied to conduct a virtual experiment on mode-I crack modeling, and the obtained results showed that this methodology could predict the crack propagation path and fatigue life of the modeled 2D structural component.

Keywords: Crack growth · Fatigue life · Fatigue fracture · Stress intensity factor · Rain flow method

1 Introduction

Mechanics of fracture involves a study of the presence of cracks loaded mechanical and structural systems. The process of fracture failure initiated at local defects like micro-cracks and voids at the grain boundaries can lead to the formation of a crack due to rupture at the molecule level, atomic bonds, or dislocation slips [1]. The cracked body subjected to one of the three modes of loads (Fig. 1), each with its intensity factor. These are designated as mode I (tensile opening, K_I), mode II (in-plane sliding, K_{II}), and mode III (out-of-plane shear/tearing, K_{III}) [2].

The stress intensity factor (SIF) is one of the most critical parameters in fracture mechanics analysis. It defines the stress field close to the crack tip and provides necessary

information about how the initiated crack propagates. In this study, the displacement extrapolation method employed to calculate the SIF of a selected mechanical component. While fatigue failure is a predominant type of crash in many mechanical and structural systems, the degree of impact of the fatigue is dependent on the intensity and frequency of the applied stress, particularly the number of load cycles and the stress amplitude.

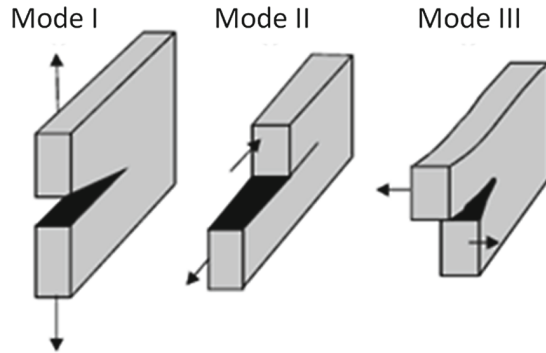


Fig. 1. Three modes of loading that can be applied to crack

The prediction of fatigue crack propagation under variable amplitude loading, even in intricate two dimensional (2D) structural components, is a difficult problem, mostly solved by sequentially using the so-called global and local analysis procedures. The crack path and associated stress-intensity factors (K_I and K_{II}) can be calculated by a finite element method (FEM) that involves global discretization of the part using appropriate crack tip elements, mesh regeneration schemes and crack increment criteria.

In this process, SIFs depend on the crack length, which can vary, and the geometry of the component. The crack is advanced a small size, and the number of cycles for the following crack increment and its direction evaluated and estimated. The challenge with the use of FEM can be for adequately prediction of the crack advancement. As a result of this problem, alternative computational methods of crack propagation are proposed and used, including the eXtended Finite Element Method (XFEM) [3, 4], and the Boundary Element Method (BEM) [5].

In this article, the growth of a mode-I crack in a selected component is modeled and its damage assessment is conducted. Following this introduction, Sect. 2 discusses the computation method of SIF to predict fatigue cracks using numerical modeling approach in ABAQUS. Due to the limitations of SIF based methods to predict failure of a dynamic system, the rain flow counting algorithm was developed in MATLAB, and its prediction of the cumulative damage assessment is presented using a randomly selected tension specimen in Sect. 3. Finally, Sect. 4 briefly presents the conclusions.

2 Computation of Stress Intensity Factor

Conducting numerical computation of stress intensity factor along the crack path by using two dimensional (2D) plane stresses, the finite element models requires three methods:

- (i). Displacement correlation technique (DCT) [6];
- (ii). Potential energy release rate computed through a modified crack-closure (MCC) integral method [7];
- (iii). J-integral calculated utilizing the equivalent domain integral (EDI) together with a mode decomposition scheme [8].

Since all these three methods can predict the SIF successfully, DCT has been chosen for this study because it is relatively simpler and adaptable to ABAQUS software, which is used for in this study. When applying DCT, the displacements obtained from the FE analysis at specific locations were correlated with the analytical solutions expressed in terms of the stress-intensity factors. For quarter-point singular elements near the crack tip, the crack opening displacement δr is given by [9]

$$\delta(r) = (4v_{j-1} - v_{j-2})\sqrt{\frac{r}{L}} = K_I \left(\frac{k + 1}{\mu} \right) \sqrt{\frac{r}{2\pi}} \tag{1}$$

where v_{j-1} and v_{j-2} are the relative displacements in the y-direction at the $j-1$ and $j-2$ nodes, L is the size of the element (illustrated in Fig. 2), and $k = 3-4\nu$ in-plane strain or $k = (3-\nu)/(1 + \nu)$ in the plane stress, ν is the poisson’s ratio and μ is the shear modulus.

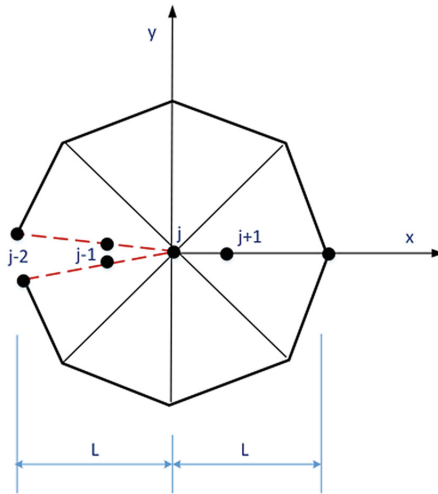


Fig. 2. Quarter-point elements at the crack tip

The mode-I and mode-II stress intensity factors can be evaluated from Eq. (2)

$$K_I = \left(\frac{\mu}{k + 1} \right) \sqrt{\frac{2\pi}{L}} (4v_{j-1} - v_{j-2}) \text{ and } K_{II} = \left(\frac{m}{k + 1} \right) \sqrt{\frac{2\pi}{L}} (4u_{j-1} - u_{j-2}) \tag{2}$$

where u_{j-1} and u_{j-2} are relative displacements in the x-direction at the $j-1$ and $j-2$ nodes. Crack initiation occurs if applied stress intensity is more significant than material toughness ($K > K_{IC}$) [10]. Fracture mechanics provides a quantitative way to predict the

onset of rapid fracture for a stressed part with a crack. As proposed by Irwin, the stress intensity factor (K) or the fracture toughness is given by [11]:

$$K = \sigma \sqrt{\pi a} [f(a/W)] \quad (3)$$

where a is the crack length, $f(a/W)$ = geometrical correction factor for the loading conditions, part geometry, and crack length.

2.1 Crack Growth Modeling Development

Fracture and failure modeling allows product designers to maximize the safe operating life of components and structures. Such crack phenomena can be modeled and analyzed by software packages like ABAQUS, which includes modeling and post-processing capabilities for fracture mechanics analysis. These features provide interactive access to the contour integral fracture mechanics technology. With these tools, any designer needs to create the part models to estimate the stress intensity factors, and crack propagation directions. For this study, standardized compact tension specimen was modeled, and the stress intensity factor results were compared with those generated from the applicable American Society for Testing and Materials (ASTM) standards [10].

The stress intensity factors are used in linear elastic fracture mechanics to measure the strength of the local crack-tip fields. Crack propagation analysis is allowed for five types of fracture criteria, i.e. (1) critical stress at a certain distance ahead of the crack tip, (2) significant crack opening displacement, (3) crack length versus time, (4) virtual crack closure technique (VCCT), and (5) low-cycle fatigue criterion based on the Paris law [12]. Displacement Extrapolation Method (DEM) is one of the known methods to evaluate the stress intensity factor [13]. Such an analysis method uses a fit of the nodal displacements in the vicinity of the crack and can be modeled by ABAQUS 6.13. As part of this study, this method was employed, and the workflow is illustrated in Figs. 3 and 4. The crack modeled in a 2D planar space (Fig. 3(a)) assumes an isotropic and deformable material (aluminum, $E = 72$ GPa, and poisson's ratio of 0.33) under plane stress condition (Fig. 3(b)).

2.2 Estimation of the SIF

From the conducted simulation for the modeled part, the numerical values of SIF (K_I and K_{II}) have been estimated at three nodes (Table 1) that are exposed to crack growth, and the crack growth confirmed along the path. As can be observed from the pictorial plot in Fig. 5(a), the deformed shape of the cracked model showed that there is clear evidence for crack propagation along the pre-designed crack area and von Mises failure criteria are satisfied for failure propagation. Besides, there is an indication of increasing SIF value along with the nodes, which may lead to failure propagation in the part (Fig. 5(b)). The deformation magnitude (Fig. 5(c)) indicates a tendency of higher values for the load application side and lower values for the fixed (lower) side of the part that is used for further structural optimization.

Though the SIF along the developed pre-cracked area may give a prediction of crack propagation, there is a limitation to assess the fatigue life of, particularly dynamic structures under crack growth. For this purpose, the rain flow counting method was employed and discussed in Sect. 3.

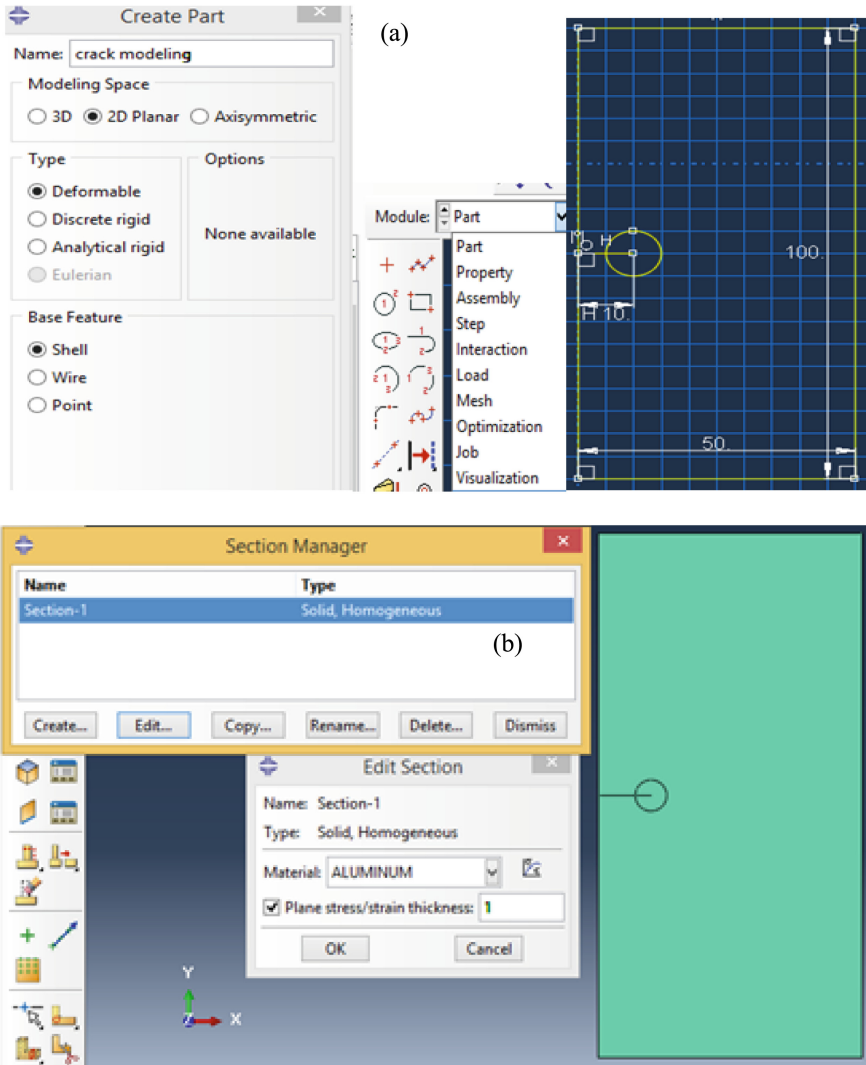


Fig. 3. Crack geometry modeling and assignment of the property of part module (a) 2D part modeling (b) material property assignment

3 Fatigue Crack Growth Analysis by Rain Flow Cycle Counting Method

Cycle counting methods are algorithms that identify fatigue cycles by combining and extrapolating information from extreme (maxima and minima) in a time series. These algorithms are used together with damage accumulation rules that can calculate the total damage as a summation of increments [14].

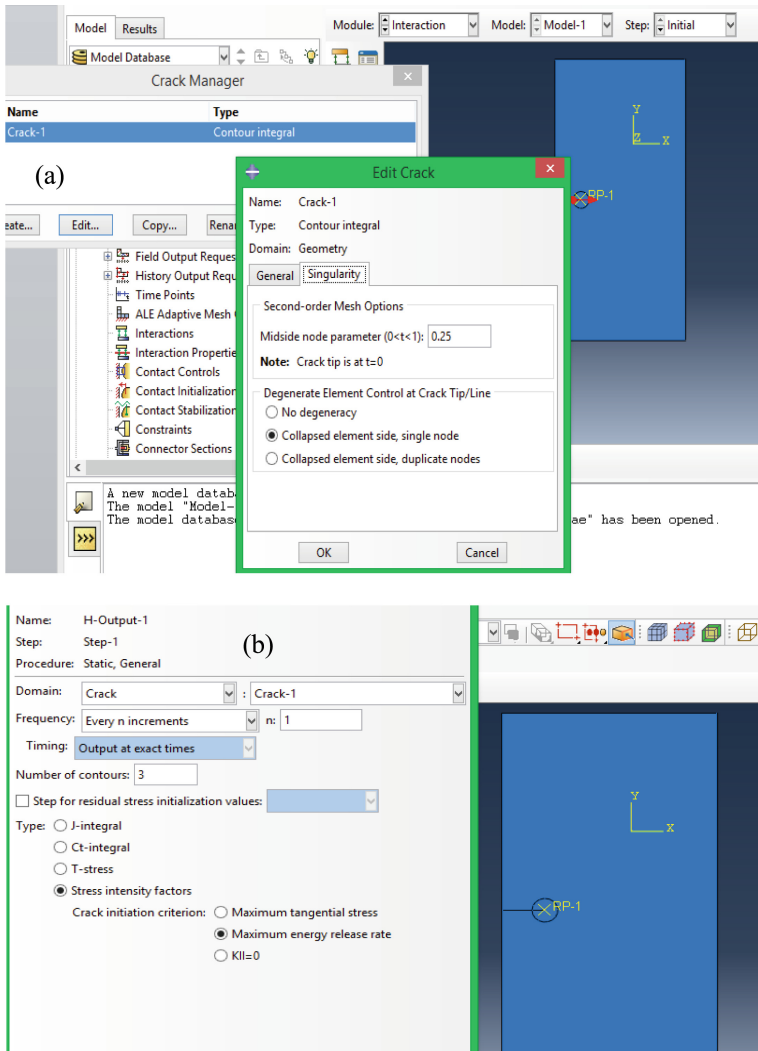


Fig. 4. Property assignments on crack model (a) definition of element control and (b) interaction module and stress intensity factor

There are several cycle counting methods that are used in the area including, rain flow counting (RFC) method, peak-valley counting (PVC), level-crossing counting (LCC), range counting (RC), and range-pairs counting (RPC) [15]. The most popular method among these counting methods is the so-called rain flow counting method, which is often employed jointly with the Palmgren-Miner's rule of linear damage accumulation to calculate the expected damage [16]. Though the Palmgren-Miner's rule is popular due to its simplicity, its application is based on the assumption that the load is constant and interaction and sequence effects that might have a significant contribution to the damage, such as random loading, are neglected [17].

Table 1. SIF values at nodes of crack growth.

Name	K factor estimate results		
	Node set		
	Node 1	Node 2	Node 3
K1	1220	2182	3088
K2	-56.34	286.2	397.5
J from Ks	2.0716E-05	6.729E-05	1.347E-04

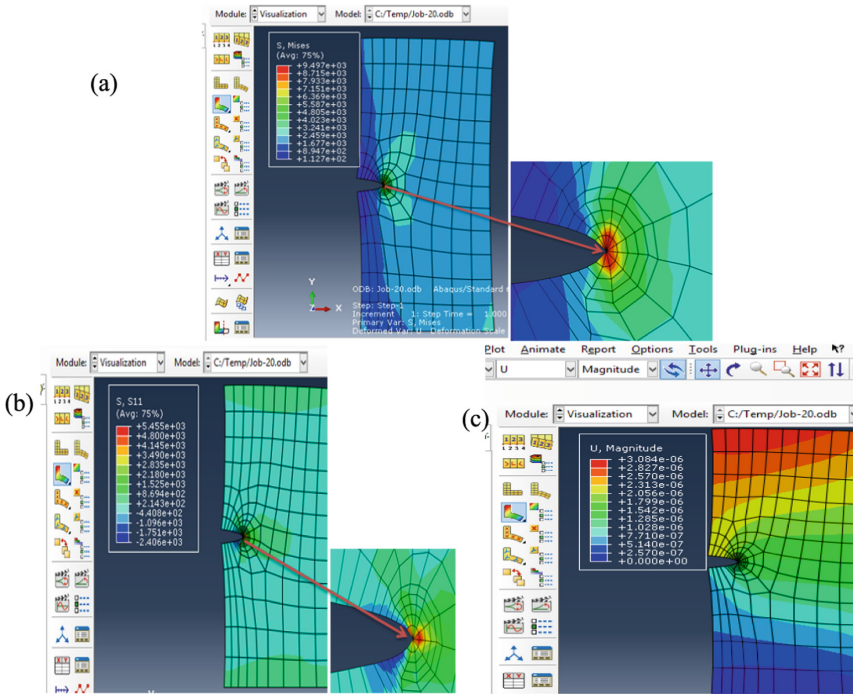


Fig. 5. Display of results in terms of (a) von Mises stress, (b) stress intensity factor and (c) deformation magnitude

The rain flow counting method is considered to have a complex sequential and nonlinear structure to decompose arbitrary sequences of loads into cycles, and its name comes from its analogy with roofs collecting rainwater to explain the algorithm, sometimes referred to as pagoda roof [18–20].

The working procedure of the RFC method (adapted from [16]) is depicted in Fig. 6. In this procedure, the algorithm usually extracts cycles from the stress or strain history obtained from measurement or simulation. By counting several full periods and half-cycles, different amplitudes, and mean values are obtained. Using the advantage provided

by fatigue damage accumulation hypotheses, like Miner’s rule, the algorithm gives the possibility to compute the expected fatigue life under random loading conditions [21].

Applying the rain flow counting method can result in the complicated random load function, Load X(t), which is converted into equivalent load cycles and gives amplitudes that serve as measures of identifying stress reversals, and Miner’s rule used to perform the damage summation. Many materials have an explicit relation between the number of cycles to failure and the cycle amplitude. This relation is often described by the S-N diagram or Wohler curves that are depicted in log-log or semi-log scales [22].

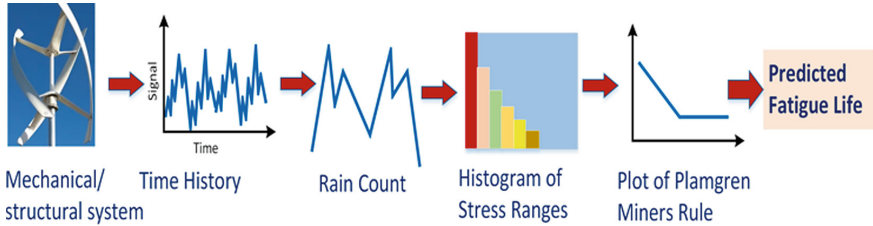


Fig. 6. Rain flow counting based damage estimation procedure.

For many materials there is an explicit relation between number of cycles to failure and cycle amplitude, which is known as S-N or Wohler curves and given as a line in a log-log scale as depicted by [10].

$$S^k N = K \tag{4}$$

where, k and K are material specific parameters and N is the number of cycles to failure at a given stress amplitude S. For a given time history, the total damage under the linear accumulation damage (Palmgren-Miner’s) rule is described as

$$D(T) = \sum_{i=1}^{N(T)} \Delta D_i = \sum_{i=1}^{N(T)} \frac{1}{N_i}, \tag{5}$$

where ΔD_i is the damage increment associated with each counted cycle, N_i is the number of cycles to failure associated with stress amplitude S_i and the number of all counted cycles $N(T)$. The S-N curve relationship in (4) is rewritten as:

$$D(T) = \sum_{i=1}^{N(T)} \frac{S_i^k}{K} \tag{6}$$

Implementation of the RFC method requires rules to be followed, including the following [23].

1. A rain flow starts at each peak and trough.
2. For a rain flow path that started at a trough and came to a tip of the roof, the flow stops if the opposite channel is more harmful than the one where the flow started from;

- 3. The stream started at a peak is stopped by a mountain, which is more favorable than the one at the flow start point.
- 4. If the rain flowing down a roof intercepts flow from an earlier path, the present way is stopped.

The rain flow counts was applied to the process in Fig. 7 and generated results in half cycles of trough and the peak stress ranges for compressive and tensile stresses, respectively. Miner’s rule is employed to assess the damage after tilting the generated signal at 90° and by preparing a rain flow sketch for the analysis, as presented in Fig. 8. In both figures, point G is the most recent signal, while point J is the previous signal point, and so on. The results of the counting cycles are given in Table 2.

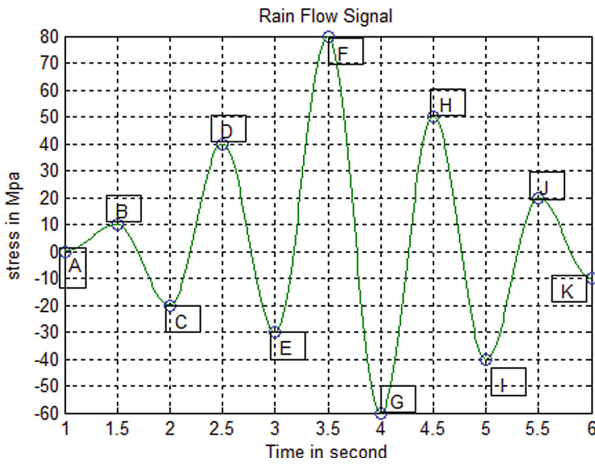


Fig. 7. Inverted rain flow counting

Based on the relation in (6), the cumulative damage *D* and the number of cycles to fracture *N* were computed using the histogram of the stress range of each sigma group and the applied cycles *n_i* by applying Miner’s principle, which is expressed as

$$D = \sum_i \frac{n_i}{N} \geq 1 \tag{7}$$

The relation in (7) implies, according to this method, that damage has occurred if the cumulative damage exceeds one (1). This method (Palmgren-Miner’s rule) is often referred to as a non-conservative prediction of fatigue failure [21]. As a result, this is the widely proposed method for the assessment of fatigue failure of mechanical and structural systems even under variable amplitude stresses.

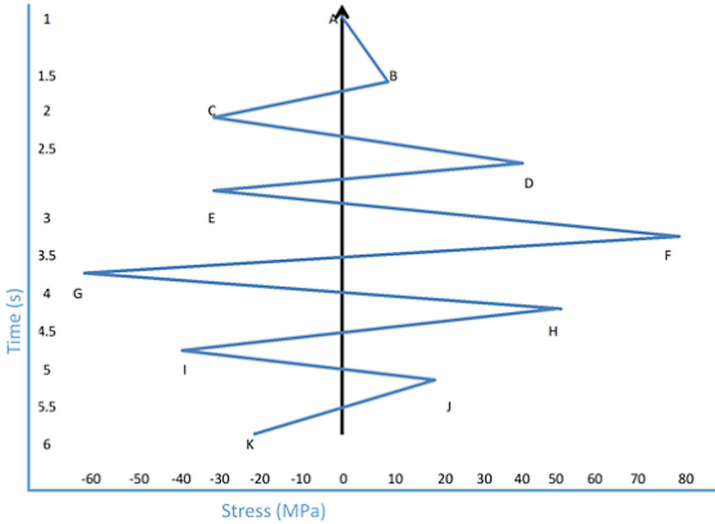


Fig. 8. Rain flow methods for damage assessment of model

Table 2. Rain flow counting cycles

Path	Stress range	Cycle (n)	Sigma, g(Sg)	Cycle, g(Ng)	Path
AB	10	0.5	150
BC	40	0.5	140	0.5	FG
CD	70	0.5	120
DG	100	0.5	110	1.5	EF, GH
EF	110	1	100	0.5	DG
FG	140	0.5	90	0.5	HI
GH	110	0.5	70	0.5	CD
HI	90	0.5	60	0.5	IJ
IJ	60	0.5	40	0.5	BC
JK	40	0.5	10	0.5	AB

Note Sg is stress of grouped path and Ng stands for sum of cycle for grouped path

4 Conclusions

This paper presents a method for modeling crack growth and damage assessment intended to demonstrate crack prediction in 2D models with random induced signals for fatigue testing purposes. The obtained simulation results using ABAQUS and MATLAB are in very close conformance with those reported in ASTM standards and other published experimental results. The assessment also indicated good prediction of crack growth with stress intensity factor.

The stress intensity factor method described in this paper provides powerful capabilities for crack growth prediction using FEM, which is capable of simulating complex real-world problems at acceptable accuracy. FEM models, as virtual experimental tools, can be validated with experimental results so that the results can effectively be used to explore without a need for any expensive preliminary testing. The problem of determining stress intensity factors for a crack emanating from 2D structural components in a tensile loaded infinite plate is prime importance in damage tolerance analysis. The method used in the study and detailed by the report is a helpful technique to calculate the stress intensity factor for many other loading cases and crack lengths. These provide essential information for subsequent studies, especially for fatigue loads, where stress intensity factor is necessary for the crack growth rate determination.





References

1. Farahmand, B.: Fracture mechanics of metals, composites, welds, and bolted joints. Springer Verlag Publishers (2000).
2. Anderson, A.T.: Fracture mechanics: Fundamentals and applications. CRC Press, Boca Raton, FL, US (2005).
3. Belytschko, T., Black, T.: Elastic crack growth in finite elements with minimal re-meshing. *International Journal for Numerical Methods in Engineering* 45(5), 601-620 (1999).
4. Rege, K., Lemu, H.G.: A review of fatigue crack propagation modeling techniques by using FEM and XFEM. In *Proceeding of COTech2017 Conference*, IOP Conference Series, Material Science and Engineering 276 (2017).
5. Mi, Y., Aliabadi, M.H.: Three-dimensional crack growth simulation using BEM. *Computers and Structures* 52(5), 871-878 (1994).
6. Leung, A.Y.T. Zhou, Z., Xu, X.. Determination of stress intensity factors by the finite element discretized symplectic method, *International. Journals of Solids Structures*, 51(5), 1115–1122 (2014).
7. Rokach, I.V., Stress intensity factor calculation by smoothed finite element method, *Solid State Phenomena*, 250, 163-168 (2016).
8. Zhu, X.K., Joyce, J.A., Review of fracture toughness (G, K, J, CTOD, CTOA) testing and standardization, *Engineering Fracture Mechanics.*, 85, 1 – 46 (2012).
9. Kurishita, H., Yamamoto, T., Nagasaka, T., Nishimura, A., Muroga, T., Jitsukawa, S., Fracture toughness of JLF-1 by miniaturized 3-point bend specimens with 3.3 – 7.0 mm thickness, *Materials Transactions*, 45 (3), 936–941 (2004).
10. ASTM E 399–90.: Standard test method for plane-strain fracture toughness of metallic materials, American Society for Testing and Materials; Philadelphia , 1997.
11. Irwin, G.R. Linear fracture mechanics, transition fracture, and fracture control, *Engineering Fracture Mechanics*, 1(2), 241-257 (1968).
12. Lim, I.L., Choi, S.K., Comparison between various displacement-based stress intensity factor computational techniques Nodal lettering, *International Journal of Fracture*. 58, 193–194, (1992).
13. Guinea, G.V., Planas, J. Elices, M., KI evaluation by the displacement extrapolation technique, *Engineering Fracture Mechanics*. 66(3), 243 – 255 (2000).
14. Marsh, G. Wignall, C. et al.: Review and application of Rainflow residue processing techniques for accurate fatigue damage estimation, *International Journal of Fatigue*, 82, 757–765 (2016).
15. Larsen, C.E., Irvine, T., A review of spectral methods for variable amplitude fatigue prediction and new results, *Procedia Engineering*, 101, 243–250 (2015).

16. Ding, J., Chen, X., Fatigue damage evaluation of broad-band Gaussian and non-Gaussian wind load effects by a spectral method, *Probabilistic Engineering Mechanics*, 41, 139–154 (2015).
17. Hensel, J., Nitschke-pagel T., Rebelo-kornmeier, J., Dilger, K., Experimental investigation of fatigue crack propagation in residual stress fields, *Procedia Engineering*, 133, 244–254 (2015).
18. Glinka, G., Kam, J.C.P. Rainflow counting algorithm for very long stress histories,” *International Journals of Fatigue*, 9(4), 223 – 228 (1987).
19. Lindgren, G., Rychlik, I., Rain flow cycle distributions for fatigue life prediction under Gaussian load processes, *Fatigue & Fracture of Engineering Materials & Structures*, 10(3), 251 - 260 (1987).
20. Olgan, M., Practical computation of statistical properties of rain flow counts, *International Journal of Fatigue*, 16(5), 306 – 314 (1994).
21. Hammerum, K., Brath, P., Poulsen, N.K., A fatigue approach to wind turbine control,” *Journal of Physics*, IOP Conf. erence Series, 75, 012–081 (2007).
22. Singh, A., The nature of initiation and propagation of S-N curves and below the fatigue limit, *Fatigue & Fracture of Engineering & Materials & Structures*, 25, 79 – 88 (2002).
23. Baek, S.H., Cho, S.S., Joo, W.S., Fatigue life prediction based on the rain flow counting method for the end beam of a freight care bogie, *International Journal of Automotive Technology*, 9(1), 95-101 (2008).



Influence of Viscosity on Density-Unstable Fluid-Fluid Displacement in Inclined Eccentric Annuli

Maryam Ghorbani¹  , Knut Erik Teigen Giljarhus² ,
and Hans Joakim Skadsem¹ 

¹ Department of Energy and Petroleum Technology, University of Stavanger, Stavanger, Norway
maryam.ghorbani@uis.no

² Department of Mechanical and Structural Engineering and Materials Science, University of Stavanger, Stavanger, Norway

Abstract. Construction of wells for oil production, geothermal energy recovery or geological storage of carbon dioxide is performed in stages by drilling the rock formation to a certain vertical depth, and then isolating the drilled section by running and cementing a casing string in the hole. The conventional cementing strategy relies on displacing the annular space behind the casing from the bottom and toward the surface. The conventional circulation direction therefore places the denser cementing fluids below the original annular fluid. An alternative cement placement strategy is to displace the annular space from the surface and downward by injecting cementing fluids directly into the annulus. Such reverse circulation operations have the benefit of lower circulation pressures and a reduced risk of fracturing the formation during placement, but also leads to density-unstable displacement conditions and increased risk of fluid contamination. To better understand how the design of cementing fluids and their placement rate affect reverse circulation displacements, we performed a series of computational simulations of density-unstable displacements using a realistic three-dimensional geometrical model of a wellbore annulus. We address the effects of wellbore inclination and inner pipe eccentricity, with a particular focus on impacts of the fluid viscosity hierarchy on the annular displacement efficiency. Our results show that increasing the displaced fluid viscosity will act to suppress the tendency for buoyant backflow, while it can worsen displacement of the narrow side of the eccentric annulus. Transverse secondary flows, which are stronger in cases where the displaced fluid is less viscous, contribute to the displacement of the narrow, low side of the annulus. We also observe that increasing the imposed axial velocity will tend to stabilize the annular displacement and suppress backflow, in agreement with previous work for buoyant pipe displacements. The present computational study is a step toward understanding how buoyant, inertial, and viscous stresses affect density-unstable displacement flows for reverse circulation cementing.

Keywords: Density-unstable fluids displacement · Viscosity ratio · Eccentric annuli

1 Introduction

Density-unstable fluid-fluid displacement, where one fluid is displaced by a denser fluid, exists in many geophysical [1–4] and industrial processes such as well cementing within the petroleum industry [5]. In primary cementing, the drilling fluid is displaced from the annular space between the casing and the formation and replaced by a cement slurry. The displacement can be performed by either conventional circulation or by the reverse circulation method. In conventional circulation primary cementing, the cement slurry is pumped down the well inside the casing and then up the annulus from the bottom of the well. Reverse circulation primary cementing is an alternative method for establishing annulus cement where the cementing fluids are injected down the well directly in the annulus. Figure 1 illustrates the flow path for conventional and reverse cementing methods.

Since the cement slurry is normally denser than the drilling mud, fluid displacement by the reverse circulation method generally leads to density unstable configuration in the annulus behind casing, and consequently higher risk of fluid contamination. Benefits of the reverse circulation cementing method include a reduced equivalent circulation density during placement and the possibility for shorter slurry thickening times, which makes this placement technique preferable compared to the conventional circulation method in certain cases [6–8]. For instance, the lower circulation pressure associated with reverse circulation cementing is an advantage for cementing of wells across weak or depleted formations [7, 8].

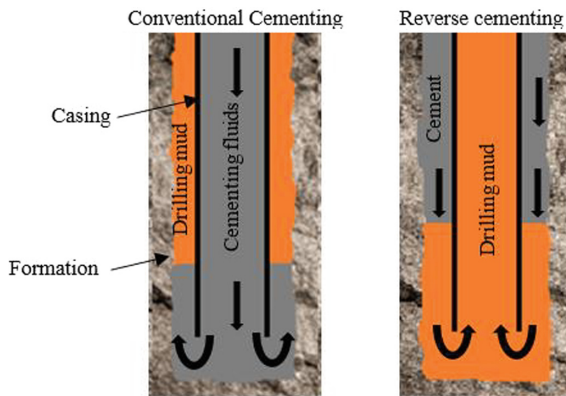


Fig. 1. Comparison of fluids displacement for a conventional and reverse circulation cementing. The cement is the displacing fluid, and the drilling fluid is the displaced fluid.

Once placed behind the casing, the cement slurry will harden to form an annular cement sheath that should seal the formation behind the casing and provide hydraulic isolation and mechanical support of the well. Isolation failure can lead to contamination of potable groundwater or leakage to the surface. Leakage can also compromise well productivity [5]. To successfully achieve hydraulic isolation outside the casing, the drilling fluid that occupies this annular space initially needs to be fully removed. Hence, investigation of the drilling fluid displacement has been motivation of several studies.

Annular fluid displacement is dependent on the geometry of the well and the annulus to be displaced. The annular space is often very narrow, and the casing is normally not perfectly centralized in the hole. Thus, casing eccentricity, which exists in most wells, will lead to non-uniform flow velocity over the annulus cross-section and promote an axial elongation of the interface between successive fluids during displacement. Drilling fluids, which are often yield stress fluids, can cease to flow altogether on the narrow side [9] or in parts of the annulus where the fluid velocity is low, such as inside a washed out section of the well [4, 10]. In an eccentric annulus, cement may favour the widest side and bypass slower-moving mud in the narrower side. The tendency of the cement to bypass the mud is a function of the geometry of the annulus, the density and viscosity of the mud and cement, and the imposed flow rate [11]. Tehrani et al. experimentally and theoretically studied the effect of eccentricity and flow rate and rig inclination on displacement efficiency for conventional circulation displacements. A positive effect on density stable displacement was found for increasing flow rates, while a negative effect was found for higher inclinations. However, the reduction of displacement efficiency can be compensated by careful selection of fluid rheology parameters [12, 13]. McLean et al. reported effects of buoyancy forces and rheological properties of mud and cement in the cementing process of an eccentric annulus [11]. Experimental studies of Malekmohammadi et al. in density stable displacement showed that small eccentricity, higher viscosity ratio, higher density ratio and slower flow rates all appeared to improve a steady displacement for Newtonian fluids and non-Newtonian fluids. However, the role of flow rate was less clear for non-Newtonian fluids.

The density unstable fluid configuration in reverse circulation cementing can result in Rayleigh-Taylor-like instabilities [14, 15] and increased intermixing between the fluids as a result. Séon et al. studied experimentally the buoyancy-driven exchange flow and the growth of the mixing region between two miscible fluids in a long pipe at different inclinations. The experiments showed that the mixing zone between the two fluids grows faster by increasing the inclination of the pipe up to an angle of 60° from the vertical direction. At higher inclinations, the greater transverse component of buoyancy leads to increased segregation and stratification of the fluids [16–18]. Etrati et al. performed a series of experiments to study the effect of viscosity ratio on the density-unstable fluid displacement in an inclined pipe. They found that increasing the viscosity of the displaced fluid destabilizes the flow and decreased the displacement efficiency [19]. Eslami et al. examined buoyant displacement flows in a vertical eccentric annulus experimentally. They demonstrated that there is higher level of instability and mixing between the fluids when the displaced fluid is more viscous than the displacing fluid. As a result, the fluid displacement was found less effective for these viscosity hierarchy configurations. A stronger effect of eccentricity on the fluid displacement was reported for viscoplastic fluids in comparison to Newtonian ones [20]. Skadsem and Kragset studied the buoyant displacement in a concentric vertical and inclined annulus using 3D numerical simulations [4]. The effect of imposed velocity was discussed and the importance of imposed flow rate and viscous stresses in suppressing the destabilizing effect of buoyancy were indicated.

Despite the aforementioned studies concerning density-unstable fluid displacement, reverse fluid displacement within the annular geometry is still relatively unexplored

from a computational point of view. In Skadsem and Kragset [4], the computational simulations were limited to iso-viscous fluids in a concentric annulus. In this paper we study the combined effect of eccentricity and viscosity ratio on the reverse fluid displacement in an annulus. Based on previous validation studies in buoyant exchange flows [21] and in reverse circulation displacement of iso-viscous miscible fluids [22], we limit the present study to computational simulations.

2 Geometry and Fluids

The geometry and fluids are chosen similar to Skadsem and Kragset [4], but now generalized to different eccentric orientations of the inner tube, and to Newtonian fluids with different viscosity ratios. The domain is a relatively narrow annulus with inner diameter (D_i) of 244.4 mm and the outer diameter (D_o) of 313.6 mm, giving a diameter ratio of 0.78. These dimensions correspond to the outer diameter of a 9 5/8-in production casing inside a typical 13 3/8-in intermediate casing, which are considered relevant for production wells in the North Sea [23].

In Skadsem and Kragset [4], the focus was on concentric and near-vertical annulus. Even if measures are taken in the field to centralize the casing within the wellbore, e.g. using bow-spring or rigid centralizers, the mechanical and hydraulic forces that act on the inner casing will normally de-centralize it in the wellbore [24, 25]. The annulus that is to be displaced to a cement slurry is therefore generally an eccentric annulus where fluids tend to move faster in the wider sector of the annulus. Further, with the exception of the very upper part of the wellbore, most of the deeper wellbore sections are drilled at an inclination from the vertical to optimize reservoir penetration and to enable production of remote areas far from existing infrastructure. To study the combined effects of eccentricity and inclination on the annular displacement, we consider inner casing eccentricities of 0.1, 0.3 and 0.7, and fix the inclination to 20° from vertical, with the narrow sector of the annulus to the low side. Regarding the properties of the fluids, we assume mass densities of 1820 kg/m^3 for the displaced fluid and 1900 kg/m^3 for the displacing fluid. These densities are realistic fluid densities based on industrial cementing properties [4, 23]. The corresponding Atwood number ($At = (\rho_H - \rho_L)/(\rho_H + \rho_L)$) is 0.022, where subscripts H and L denote heavy (displacing) and light (displaced) fluids, respectively. A constant viscosity of 0.1 Pa·s is assumed for the displacing fluid and values of 0.02, 0.1, 0.5 Pa·s are taken as the viscosity of the displaced fluid. The imposed bulk velocity of the fluid in the annulus is set to 0.44 m/s and 0.1 m/s. This corresponds to imposed flow rates of approximately 800 l/min and 182 l/min, respectively.

The problem outlined above is a density unstable displacement in an inclined eccentric annulus where the displacing fluid is injected from the top of the annulus at a constant imposed velocity. In the absence of interfacial tension between the fluids, and assuming smooth inner and outer casing surfaces (no wall surface roughness) the displacement problem may be characterized by the following 5 dimensionless numbers:

$$Re = \frac{\bar{\rho}V_0d}{\mu}, Fr = \frac{V_0}{\sqrt{Atgd}}, \beta, e = \frac{2\delta}{D_o - D_i}, m = \frac{\mu_L}{\mu_H}$$

Here V_0 is the imposed velocity, $d = (D_o - D_i)/2$ is the radial gap width of the annulus, $\bar{\rho}$ is the mean density of two fluids, g is the gravitational acceleration, β is the

inclination (from vertical reference line), δ and e are the casing deflection from the well centerline and the eccentricity of the annulus respectively. The mean viscosity of the fluids is denoted by $\bar{\mu}$, and μ_L and μ_H represent the viscosity of light and heavier fluid, respectively. The ratio of inertial to viscous forces is written as quantified by the Reynolds number, Re , and the ratio of mean-imposed velocity to an inertial-buoyant velocity scale, $V_I = \sqrt{Atgd}$, is the densimetric Froude number. The range of dimensionless numbers considered in this study are summarized in Table 1.

Table 1. Range of dimensionless numbers considered in the simulations.

Parameter	Value
Re	21.5, 64.5, 94, 107, 283, 471
Fr	1.15, 5.15
e	0.1, 0.3, 0.7
m	0.2, 1, 5
β	20°

3 Numerical Method

The fluid displacement in the annulus is simulated by OpenFOAM, version 2012, employing the twoLiquidMixingFoam solver, which is a solver for two-phase, incompressible, and miscible fluids. The concentration is evolved by a mixture fraction equation,

$$\frac{\partial \alpha}{\partial t} + \nabla \cdot (\alpha \mathbf{U}) = 0 \tag{1}$$

where α is the volumetric phase fraction of the displacing fluid. We assume molecular mixing between the fluids is negligible [26, 27].

The momentum and continuity equations are as follows:

$$\frac{\partial \rho \mathbf{U}}{\partial t} + \nabla \cdot (\rho \mathbf{U} \mathbf{U}) = -\nabla p + \nabla \cdot 2\mu \mathbf{S} + \rho \mathbf{g}, \tag{2}$$

$$\nabla \cdot \mathbf{U} = 0, \tag{3}$$

where $\mathbf{S} = [(\nabla \mathbf{U}) + (\nabla \mathbf{U})^T]/2$ is the rate of strain tensor, \mathbf{g} is the gravitational acceleration. The mixture fluid properties can be found from the phase fraction as

$$\rho = \alpha \rho_H + (1 - \alpha) \rho_L, \tag{4}$$

$$\mu = \alpha \mu_H + (1 - \alpha) \mu_L, \tag{5}$$

Here, μ is the mixture fluid viscosity and ρ is the mixture density.

Second order discretization schemes and a second order implicit scheme are employed for spatial terms and the time discretization, respectively. To stabilize the solution an adjustable time step with a maximum Courant number of 0.5 is used. For the boundary conditions, velocity Dirichlet conditions are prescribed at solid boundaries and at the inlet (no slip and uniform, and bulk inlet velocity, respectively) and a Neumann outflow condition at the outlet. For the pressure, Neumann inlet and Dirichlet outlet conditions are used.

A computational domain with 8 m length, equal to 116 hydraulic diameters, in the axial direction is used. The initial interface between the two fluids is located 2 m below the inlet at the top. The computational grid is a structured grid, stretched towards the walls and the eccentric side. The computational mesh consists of 160 grid cells in the azimuthal direction, 30 cells in the radial direction and 1024 cells in the axial direction. The computational grid for different eccentricities is shown in Fig. 2. A grid sensitivity study was performed to ensure the results were not significantly influenced by the chosen grid resolution, see Appendix 1.

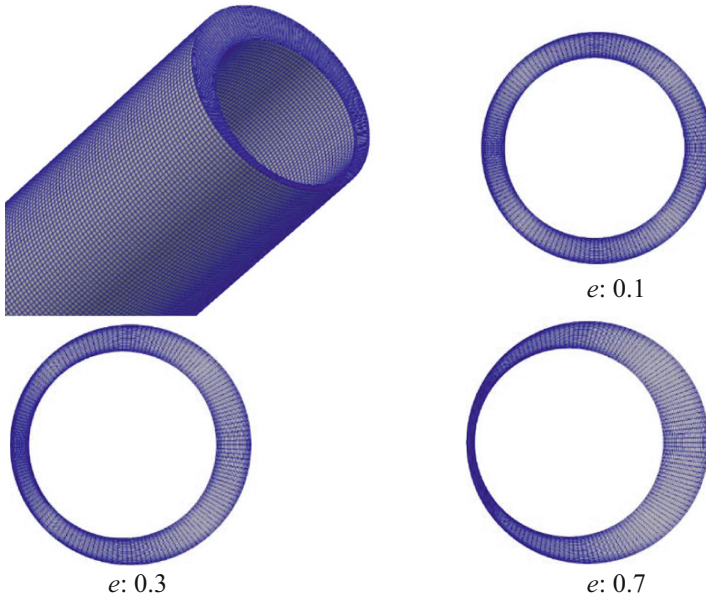


Fig. 2. The computational grid with different eccentricities

4 Results and Discussion

In the following sections, we present results from the 3D numerical simulations using different plotting techniques for studying and comparing how eccentricity, viscosity ratio and imposed flow rate affect the displacement. In addition to 3D renderings of

selected simulations, we also present results in the form of an “unwrapped” annulus where the fluid concentration is averaged across the radial gap, as shown in Fig. 3. For a given time in the simulation, this results in a 2D visualization of the displacement flow. Due to the inclination and the eccentricity, the fluid evolution is non-uniform over the annulus cross-section. Using the radially averaged concentration profiles, we plot the fluid concentration field in different azimuthal sides of the annulus, see Fig. 5 as an example. As indicated in Fig. 3, the letters “N”, “W” and “S” denote the narrow, wide, and lateral sides of the annulus, respectively and “Z” denotes the axial position. The outlet is taken at $Z = 0$ m, the inlet is at $Z = 8$ m, and the initial interface position is at $Z = 6$ m. Finally, we also present spatiotemporal diagrams, where the fluid concentration is averaged over the entire cross-section, and the resulting axial concentration profile is plotted as function of time since start of displacement. The spatiotemporal diagrams can be useful when comparing major features of the displacement cases and provide quantitative information about the extent and evolution of the fluid mixing zone.

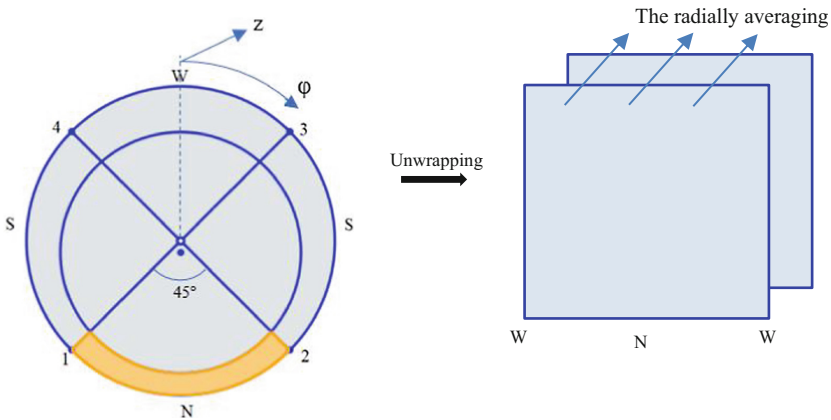


Fig. 3. Illustration of unwrapping and averaging procedure across the gap

4.1 Combined Effect of Imposed Velocity and Viscosity Ratio

Figure 4 compares the effect of the imposed velocity on the fluid displacement with different viscosity ratios using spatiotemporal diagrams. In the left column the displacing fluid is five times more viscous than the displaced fluid and vice versa in the right column. The imposed velocity in the plots in the second row is 4.4 times higher than plots in the first row. We define the mixing zone as the flow region where the concentration field, α , is between 0.01 and 0.99. In these plots the black dashed curves show the advancing front position during the displacement ($\alpha > 0.01$) while the light dashed curves show the rear, upper extent of the mixing zone ($\alpha < 0.99$). The local slope of the black dashed lines can be considered as the instantaneous front velocity, V_f .

As seen from the simulations with less viscous displaced fluid ($m = 0.2$), the tendency for backflow is prominent at the lowest imposed velocity, $Fr = 1.17$. As shown by the

white superimposed curve that tracks the rear edge of the mixing zone, some displaced fluid arrives above the initial interface position at $Z = 6$ m. This is referred to as *backflow* of displaced fluid. Backflow is associated with buoyant and unstable displacements, and result in rapidly growing mixing region. So, in this case ($Fr = 1.17$, $Re = 107$) the front velocity is higher than the case with viscosity ratio of 5 and the front position reaches the end of the annulus sooner at $t = 20$ s. We observe in Fig. 4 that backflow can be suppressed by either increasing the imposed velocity, and/or increasing the viscosity of the displaced fluid. At higher imposed velocity, $Fr = 5.15$, not only is there no backflow and slightly smaller front velocity, but there is also improved displacement of fluid at $m = 0.2$ compared to $m = 5$.

As reflected in the Reynolds number associated with these four spatiotemporal diagrams, increasing the viscosity of the displaced fluid (i.e., going from $m = 0.2$ to $m = 5$), results in a smaller Reynolds number. This is due to the Reynolds number being based on the *average* viscosity of the two fluids. Hence, by increasing the viscosity of the displaced fluid, leading to lower Re , the viscous forces suppress buoyancy, and there is no observable backflow at any of these flow rates. At the highest imposed flow rate ($Fr = 5.15$, $Re = 94.3$), the spatiotemporal diagram suggests reduced mixing between the fluids, and “stabilized” displacement where buoyancy is suppressed.

4.2 Combined Effect of Eccentricity and Viscosity Ratio

In a vertical and perfectly concentric annulus, there is no preferred direction for the dense, displacing fluid to flow since the driving forces (gravity and imposed flow) are parallel to the annulus axis, and the annular gap width is uniform. As mentioned in Sect. 2, the inner casing will in most practical cases be de-centralized inside the wellbore, and deeper well sections tend to be drilled at an inclination from the vertical direction. Inner-casing eccentricity will favor axial flow through the widest sector of the annulus, while gravity will promote flow of the denser fluid toward the low side of the annulus. In this section we will study the combined impact of inclination, eccentricity, and viscosity ratio on the fluid displacement for the case where the inner casing is offset toward the low side of the inclined wellbore.

Figure 5 presents gap-averaged concentration fields at initial stages of displacement, time = 6.6 s, for viscosity ratios of 0.2, 1 and 5 at eccentricity of 0.1, 0.3 and 0.7. By changing the viscosity of fluids, the preferred side of the annulus for the displacing fluid to flow in is also changed. Increasing the viscosity and the imposed velocity is seen to enhance the effect of eccentricity. In the cases with eccentricity of 0.1, at low imposed velocity, $Fr = 1.17$, the displacing fluid prefers the narrower side (Fig. 5a–c). For larger eccentricity, increasing the viscosity of the displaced fluid redirects the fluid to the wider regions; the lateral sides of the annulus (Fig. 5, e, h), or upper side (Fig. 5f, i). Increasing Fr strengthens the effect of eccentricity and the fluid flows in the wider side in most cases.

Figure 6 shows gap-averaged concentration fields at times that correspond to injection of equal volume of displacing fluid. The plots in the three columns on the left side correspond to the displacement with the velocity of 0.1 m/s and $Fr = 1.17$ at time of 20.6 s and the plots of the displacement with the higher velocity of 0.44 m/s and $Fr = 5.15$ at time of 5.6 s are shown on the right side. The highest level of instability, transverse

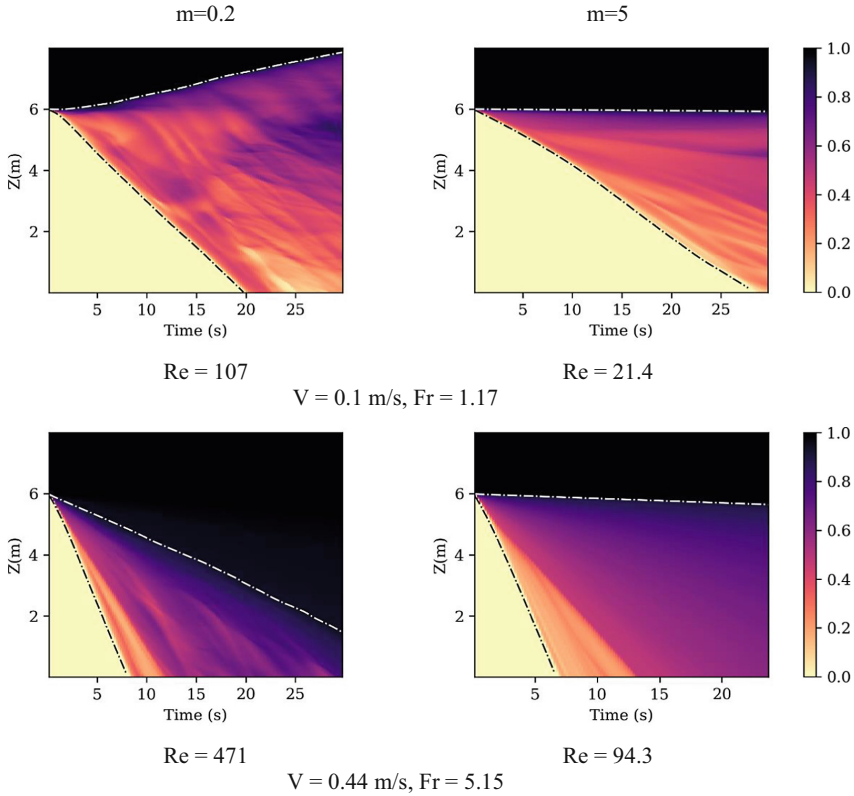


Fig. 4. Spatiotemporal diagrams for reverse circulation displacements in an inclined and eccentric annulus. The inclination is 20° and the eccentricity is 0.1

flow, and back flow of the lighter fluid are found for the cases with low imposed velocity and low viscous displaced fluid (Fig. 6 a, d, g). As discussed in Sect. 4.1, increasing the viscosity of the displaced fluid and imposed velocity tend to stabilize the displacement. It can be seen that there is less backflow in Fig. 6 (b, e, h) and no backflow for the viscous ratio of 5 in Fig. 6(c, f, i) and also more stable displacement is shown by imposing higher velocity Fig. 6 (right panels, $Fr = 5.15$).

To have a view on evolution of the mixing zone over the time of displacement the spatiotemporal diagrams from the simulations with viscosity ratios of 0.2 and 5 are presented in Fig. 7 at imposed velocity of 0.1 m/s and Fig. 8 at imposed velocity of 0.44 m/s. Cases with more viscous displaced fluid and higher imposed velocity correspond to more stable displacements, and spatiotemporal diagrams with smoother transitions across the mixing zone (Fig. 8-right column). However, this does not necessarily result in more effective displacement. According to Fig. 6, the displacement along the narrow side of the annulus is reduced as the displaced fluid viscosity increases. At $m = 5$, the resistance to flow along the narrower side of the annulus for the displaced fluid is stronger than it is at $m = 0.2$ and $m = 1$; this results in reduced displacement efficiency on the narrow side and a redirection of the flow toward the wider side.

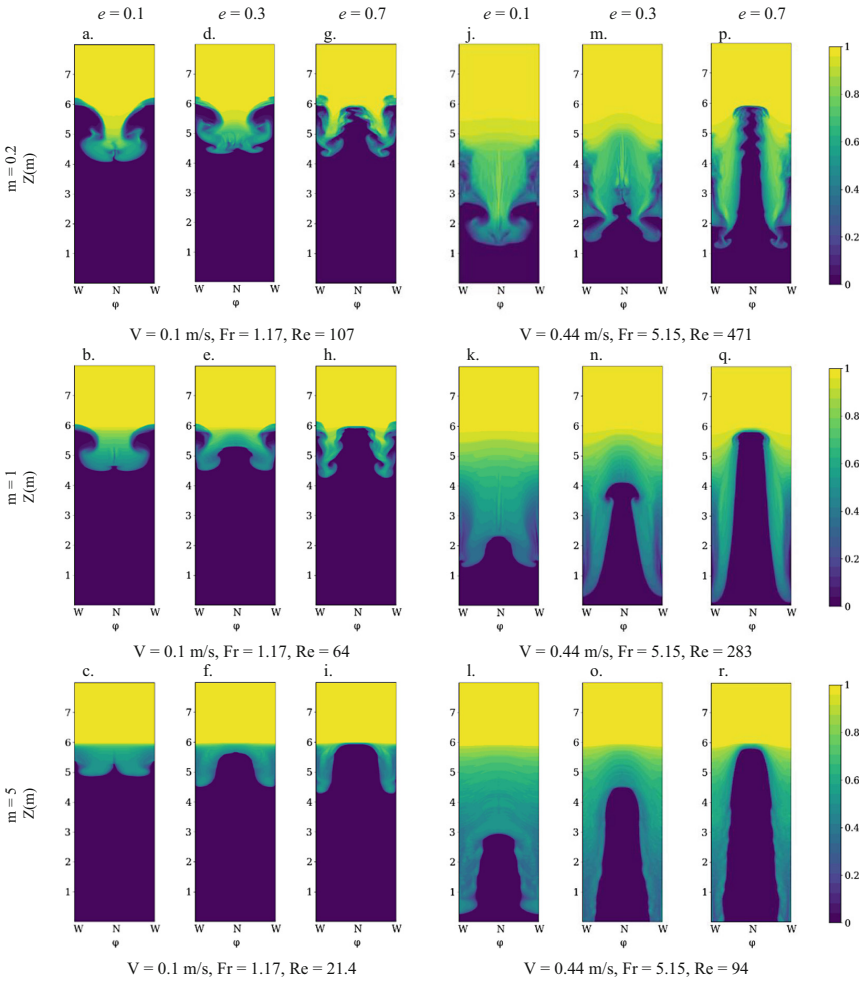


Fig. 5. The fluid concentrations averaged across the annular gap, at time 6.6 s

In Fig. 8, the dashed white line shows the front position in the *narrow* side of the annulus. By increasing the viscosity ratio, the slope of this line decreases meaning that the front velocity in this region decreases. The horizontal dashed white line in Fig. 8 at $e = 0.7$ and $m = 5$ indicates that there is no movement of the fluid-fluid interface at the narrow side of the annulus.

To do a quantitative study of the combined effect of eccentricity and viscosity ratio on the front velocity of the mixing region, V_f , the calculated front velocities are normalized by the imposed velocity, V_o , and plotted in Fig. 9 against eccentricity for all values of m and V_o . Figure 9 makes it clear that there is a different trend for the combined effect of eccentricity and viscosity ratio on the front velocity at $V_o = 0.1$ m/s ($Fr = 1.17$) and $V_o = 0.44$ m/s ($Fr = 5.15$). At $Fr = 5.15$ increasing the viscosity of the displaced fluid results in higher front velocity, while at $Fr = 1.17$, for $m = 0.2$ the front velocity is the highest.

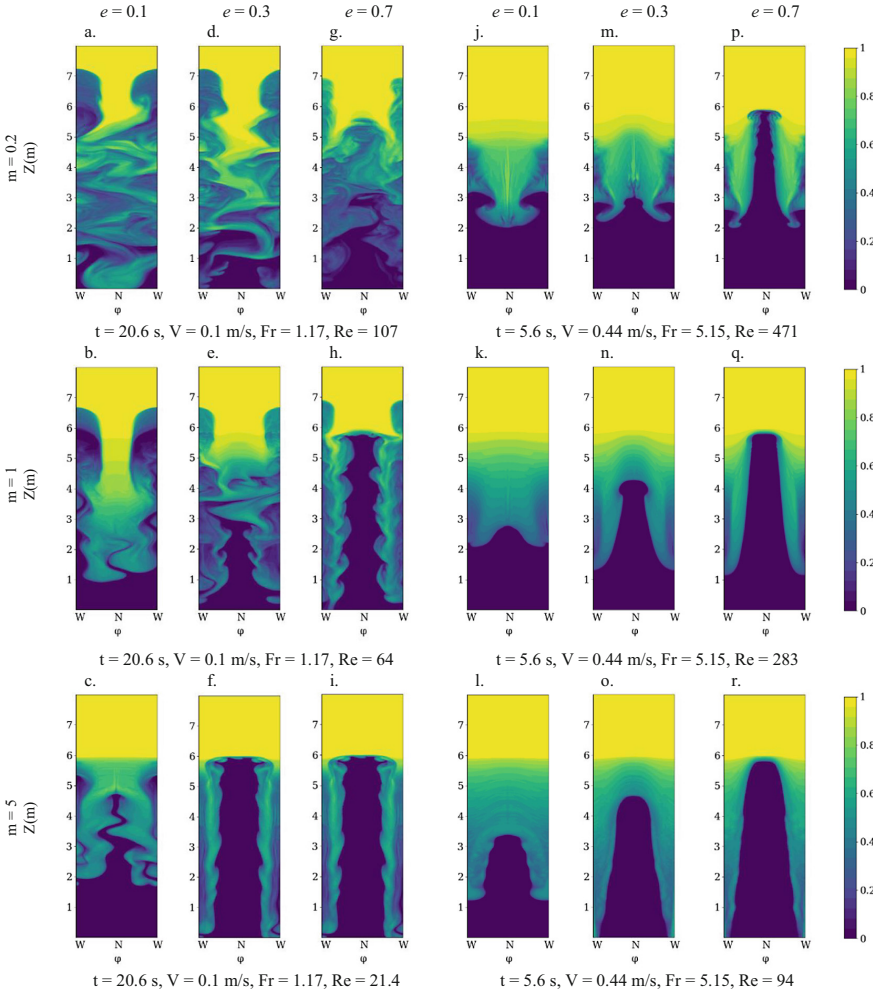


Fig. 6. The fluid concentrations averaged across the annular gap after imposing the same amount of displacing fluid.

The weaker impact of eccentricity for $m = 0.2$ and stronger effect of eccentricity for higher viscosity ratios can also be observed. The case that corresponds to the smallest normalized front velocity at the considered eccentricities is $m = 0.2$ and $V_0 = 0.44$ m/s. Since the displacements are subject to a constant imposed injection rate at the top, the smaller normalized front velocity suggests more effective displacement for this case ($m = 0.2$ and $V_0 = 0.44$ m/s) compared to the other cases considered.

4.3 Displacement in the Narrow Side

As discussed in the previous sections, the fluid prefers the wider side to flow in, so the displacement along the narrower side is often more challenging. Figure 10 compares

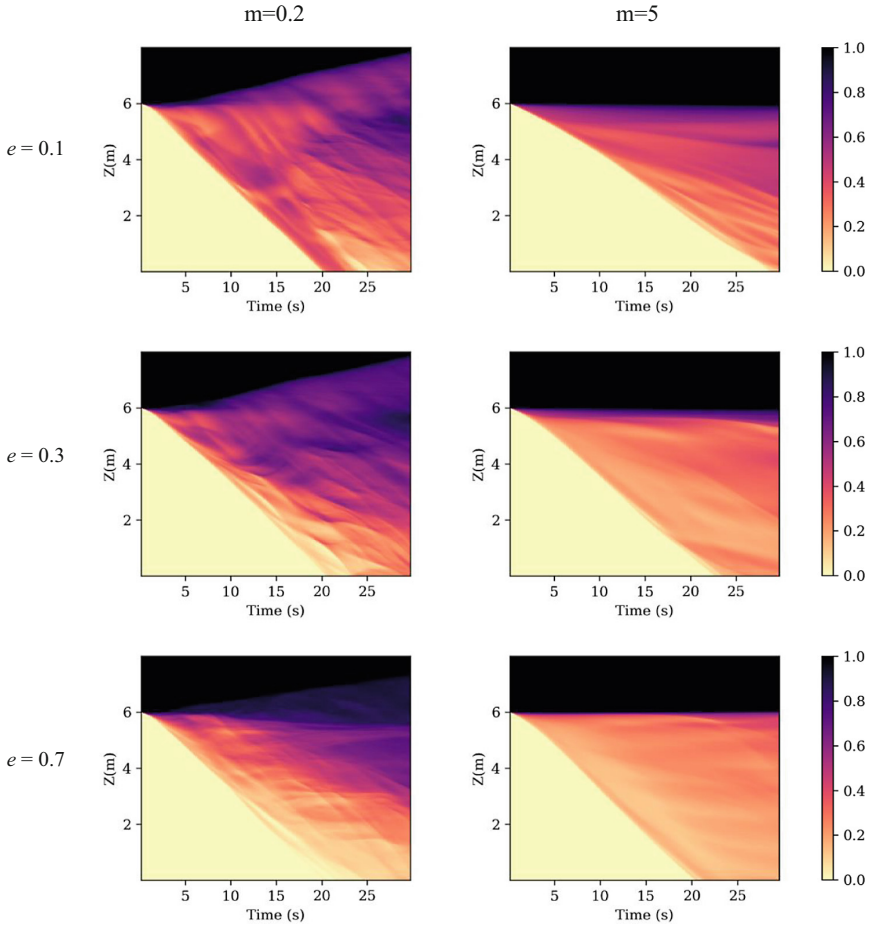


Fig. 7. Spatiotemporal diagrams for reverse circulation displacements in an inclined and eccentric annulus. The inclination is 20° and $V_o = 0.1$ m/s

the fluid concentration fields in the highly eccentric ($e = 0.7$) inclined annulus after imposing equal amount of displacing fluid at the two different velocities ($V_0 = 0.1$ m/s and time = 29.6 s in (a, b), $V_0 = 0.44$ m/s and time = 9.6 s in (c, d)). The left most diagrams (a, c) correspond to the cases with viscosity ratio of 0.2 and the other ones show the results of viscosity ratio of 5.

Further, Fig. 11 shows results from 3D simulation at different times for the highly eccentric and inclined annulus at the velocity of 0.44 m/s. The left panels correspond to $m = 0.2$ and the right panels to $m = 5$. As seen in Figs. 10 and 11, the greater viscosity in the displaced fluid ($m = 5$) results in higher restriction along the narrow/low side, and this negatively impacts the displacement here. In the cases with low viscosity displaced fluid, the displacement in the narrow side is facilitated by stronger transverse flow. In Fig. 11, we can see when the fluid cannot flow in the lower/narrower side at $m = 5$, the

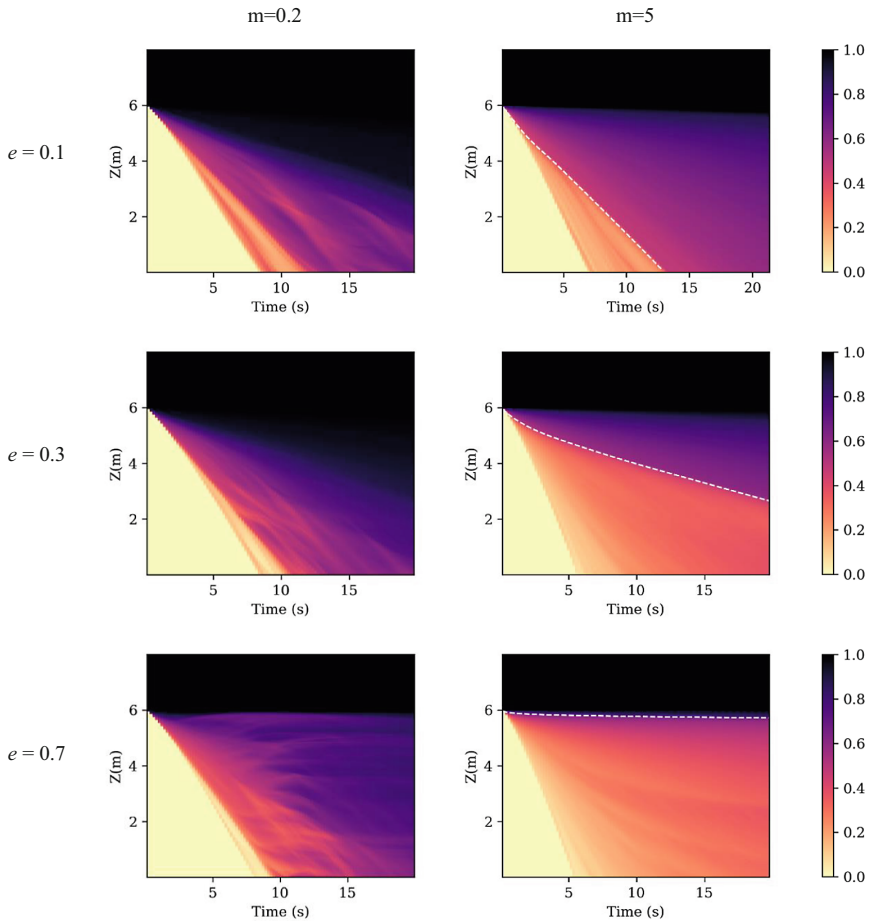


Fig. 8. Spatiotemporal diagrams for reverse circulation displacements in an inclined and eccentric annulus. The inclination is 20° and $V_o = 0.44$ m/s. The white dashed line represents the front position in the narrower side.

fluid flow in the upper/wider side with a higher velocity in comparison to the case with $m = 0.2$.

To have a better quantitative view on the combined effect of eccentricity and viscosity ratio on the displacement, the displacement efficiency in the narrow side of annulus (the orange section, N, in Fig. 3) is studied separately. In Fig. 12, the fluid displacement efficiency along this side is plotted versus time. The displacement efficiency at a given time corresponds to the volume fraction of displacing fluid along the narrow side. In these plots the results of the displacement with different viscosity ratios in the annulus with the eccentricities of 0.1, 0.3 and 0.7 are given. The imposed velocity for these cases is 0.44 m/s. It is shown that for each eccentricity, the most efficient displacement in the narrow side is when the displaced fluid is less viscous than the displacing fluid ($m = 0.2$). This observation agrees with the assessment of the normalized front velocity in

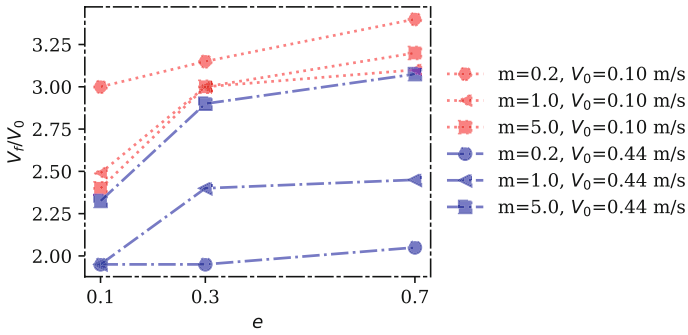


Fig. 9. Front velocity, V_f , normalized by imposed velocity, V_0 , for different eccentricity, e .

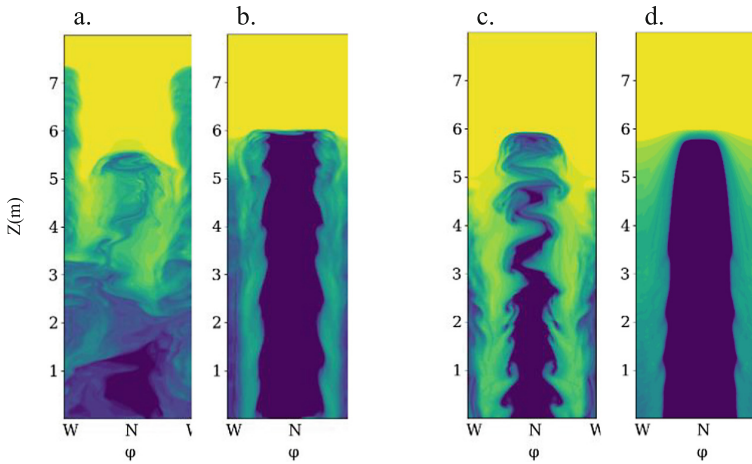


Fig. 10. The fluid concentrations averaged across the annular gap after imposing the same amount of displacing fluid, $e = 0.7$ and the imposed velocity for the two left plots is 0.1 m/s and $t = 29.6$ s, and for the plots on the right side $V_0 = 0.44$ m/s and $t = 9.6$ s.

Fig. 9. By increasing the eccentricity, the effect of viscosity ratio is strengthened, as by tracking the curves in the plots from left to right, it can be seen the difference between them increases.

5 Summary and Conclusion

Using 3D numerical simulations, we have investigated density-unstable displacement flows of Newtonian miscible fluids in an inclined annulus with different eccentricities. The study was motivated by so-called reverse circulation primary cementing of casing strings in wells but is also considered relevant for other industrial processes involving annular flow of two or more fluids. Our results have shown backflow and the highest level of instability at low imposed velocity for the case with less viscous displaced

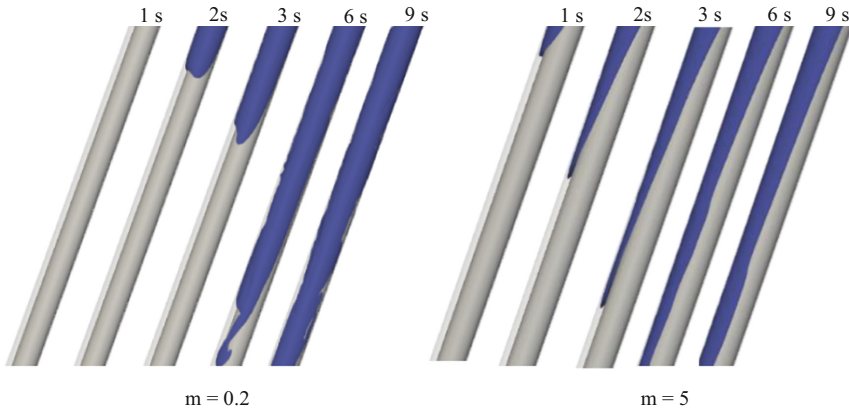


Fig. 11. 3D isocontour showing the displacing fluid for the highly eccentric and inclined annulus, $V_0 = 0.44$ m/s

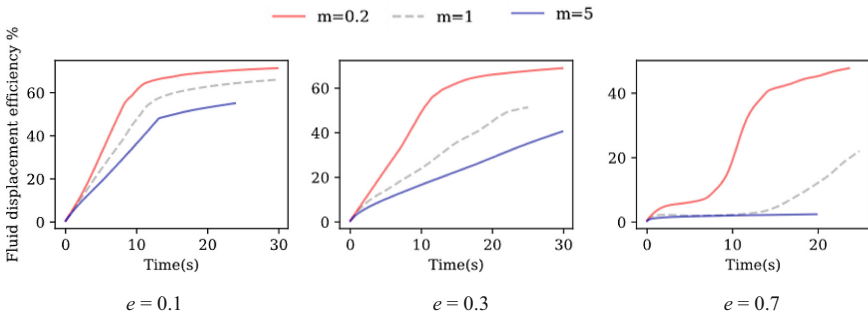


Fig. 12. The fluid displacement efficiency in the lower/narrower side of the annulus for imposed velocity $V_0 = 0.44$ m/s.

fluid. Increasing velocity can suppress this instability and results in more effective displacement. Increasing viscosity of the displaced fluid is also helpful for suppressing the backflow, but it does not necessarily lead to better fluid displacement.

The fluid evolution over the cross section of the annulus is uniform for a concentric vertical annulus. Eccentricity and/or inclination breaks this geometric symmetry and tend to promote displacement along either the wide or the narrow side of the annulus. The fluid prefers the wider side in an eccentric annulus, while gravity directs the denser fluid toward the lower side. Therefore, eccentricity and inclination can balance the effect of each other in cases where the narrow side of the annulus coincides with the low side. Our results have indicated that the fluid displacement in an eccentric annulus is not the same for all the viscosity ratios. When the displacing fluid is less viscous the fluid prefers the narrower/lower side at low imposed velocity and by increasing the eccentricity the fluid flows in the lateral sides. By increasing the viscosity of the displaced fluid, the effect of eccentricity becomes more significant, and the fluid is directed to the lateral

and wider sides at lower eccentricity. By imposing a higher imposed velocity, the effect of eccentricity becomes dominant, and the fluid prefer the wider/upper side to flow in.

We showed the effect of viscosity ratio on the front velocity is not the same at different Fr . At the low value of Fr , the highest front velocity is calculated for $m = 0.2$, while at the highest value of Fr the front velocity is higher at $m = 5$. The spatiotemporal plots have shown more stable pattern of the mixing zone at higher viscosity ratio. On the other hand, we discussed that for high eccentric annuli the instability exists at lower viscosity ratio is required, as these instabilities can facilitate the displacement on the narrower side.

In this study we studied the effect of viscosity ratio when the viscosity of the displaced fluid is constant which results in different value of mean viscosity. In future work, we will also investigate the effect of viscosity ratio at constant mean viscosity.

Appendix 1

We performed a grid sensitivity study to ensure the results were not significantly influenced by the chosen grid resolution. The simulation with parameters $Fr = 5.15$ and $Re = 283$ and an eccentricity of $e = 0.3$ was run with four different grid sizes, denoted coarse, medium, fine and extra fine. The total number of cells was doubled for each level, where fine corresponds to the grid resolution used in the remainder of the work.

Figure 13 shows the gap-averaged concentration field after 8 s. The overall evolution is almost identical for all resolutions, indicating that the fine resolution is satisfactory. This is further corroborated by the calculated front velocity and total displacement efficiency (ratio of displacing fluid volume and total annulus volume), which are shown in Fig. 14. Again, the results are not significantly influenced by the grid resolution.

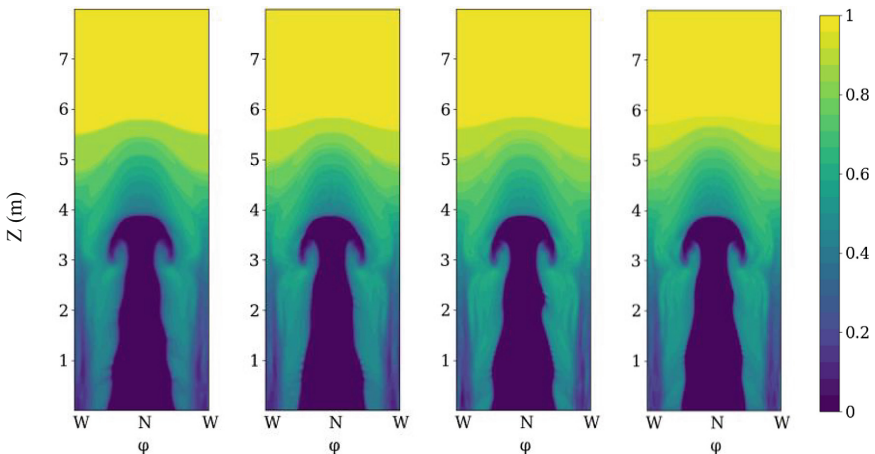


Fig. 13. Grid sensitivity study. Comparison of concentration field after 8 s. From left to right: Coarse, medium, fine, and extra fine.

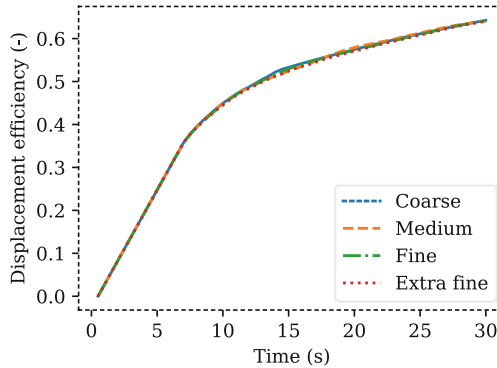


Fig. 14. Grid sensitivity study. Comparison of displacement efficiency for different grid sizes.

References

1. Benjamin, T.B., *Gravity currents and related phenomena*. Journal of Fluid Mechanics, 1968. **31**(2): p. 209-248.
2. Shin, J., S. Dalziel, and P. Linden, *Gravity currents produced by lock exchange*. Journal of Fluid Mechanics, 2004. **521**: p. 1-34.
3. Hallez, Y. and J. Magnaudet, *A numerical investigation of horizontal viscous gravity currents*. Journal of Fluid Mechanics, 2009. **630**: p. 71-91.
4. Skadsem, H. and S. Kragset, *A Numerical Study of Density-Unstable Reverse Circulation Displacement for Primary Cementing*. Journal of Energy Resources Technology, 2022. **144**.
5. Guillot, D. and E. Nelson, *Well cementing*. Schlumberger Educational Services, Sugar Land, TX, 2006.
6. McNerlin, B., et al. *Open Hole Fluid Displacement Analysis-Forward vs. Reverse Circulations*. in *SPE Annual Technical Conference and Exhibition*. 2013. OnePetro.
7. Macfarlan, K.H., et al., *A Comparative Hydraulic Analysis of Conventional-and Reverse-Circulation Primary Cementing in Offshore Wells*. SPE Drilling & Completion, 2017. **32**(01): p. 59-68.
8. Griffith, J., D. Nix, and G. Boe. *Reverse Circulation of Cement on Primary Jobs Increases Cement Column Height Across Weak Formations*. in *SPE Production Operations Symposium*. 1993. Society of Petroleum Engineers.
9. Walton, I. and S. Bittleston, *The axial flow of a Bingham plastic in a narrow eccentric annulus*. Journal of Fluid Mechanics, 1991. **222**: p. 39-60.
10. Roustaei, A. and I. Frigaard, *Residual drilling mud during conditioning of uneven boreholes in primary cementing. Part 2: Steady laminar inertial flows*. Journal of Non-Newtonian fluid mechanics, 2015. **226**: p. 1-15.
11. McLean, R., C. Manry, and W. Whitaker, *Displacement mechanics in primary cementing*. Journal of petroleum technology, 1967. **19**(02): p. 251-260.
12. Tehrani, A., J. Ferguson, and S. Bittleston. *Laminar displacement in annuli: a combined experimental and theoretical study*. in *SPE annual technical conference and exhibition*. 1992. OnePetro.
13. Tehrani, M., S. Bittleston, and P. Long, *Flow instabilities during annular displacement of one non-Newtonian fluid by another*. Experiments in Fluids, 1993. **14**(4): p. 246-256.
14. Drazin, P.G. and W.H. Reid, *Hydrodynamic stability*. 2004: Cambridge university press.
15. Sharp, D.H., *An overview of Rayleigh-Taylor instability*. Physica D: Nonlinear Phenomena, 1984. **12**(1-3): p. 3-18.

16. Séon, T., *Du mélange turbulent aux courants de gravité en géométrie confinée*. 2006, Université Pierre et Marie Curie-Paris VI.
17. Séon, T., et al., *Buoyancy driven miscible front dynamics in tilted tubes*. *Physics of fluids*, 2005. **17**(3): p. 031702.
18. Seon, T., et al., *Front dynamics and macroscopic diffusion in buoyant mixing in a tilted tube*. *Physics of Fluids*, 2007. **19**(12): p. 125105.
19. Etrati, A., K. Alba, and I.A. Frigaard, *Two-layer displacement flow of miscible fluids with viscosity ratio: Experiments*. *Physics of Fluids*, 2018. **30**(5): p. 052103.
20. Eslami, A., S. Akbari, and S. Taghavi, *An experimental study of displacement flows in stationary and moving annuli for reverse circulation cementing applications*. *Journal of Petroleum Science and Engineering*, 2022. **213**: p. 110321.
21. Ghorbani, M., et al. *Computational fluid dynamics simulation of buoyant mixing of miscible fluids in a tilted tube*. in *IOP Conference Series: Materials Science and Engineering*. 2021. IOP Publishing.
22. Ghorbani, M., A. Royaei, and H.J. Skadsem, *Reverse circulation displacement of miscible fluids for primary cementing*. *Journal of Energy Resources Technology*, 2023: p. 1–19.
23. Skadsem, H.J., et al., *Study of ultrasonic logs and seepage potential on sandwich sections retrieved from a North sea production well*. *SPE Drilling & Completion*, 2021. **36**(04): p. 976-990.
24. Gorokhova, L., A. Parry, and N. Flamant, *Comparing soft-string and stiff-string methods used to compute casing centralization*. *SPE Drilling & Completion*, 2014. **29**(01): p. 106-114.
25. McSpadden, A.R., O. Coker III, and G.C. Ruan, *Advanced casing design with finite-element model of effective dogleg severity, radial displacements, and bending loads*. *SPE Drilling & Completion*, 2012. **27**(03): p. 436-448.
26. Cook, A.W. and P.E. Dimotakis, *Transition stages of Rayleigh–Taylor instability between miscible fluids*. *Journal of Fluid Mechanics*, 2001. **443**: p. 69-99.
27. Hallez, Y. and J. Magnaudet, *Effects of channel geometry on buoyancy-driven mixing*. *Physics of fluids*, 2008. **20**(5): p. 053306.



Numerical Simulation of Stress in a Rotational Autofrettaged Thick Walled Tube Using the Actual Material Stress-Strain Curve

Xiaoping Huang¹(✉) and Dimitrios Pavlou²

¹ State Key Lab of Ocean Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
xphuang@sjtu.edu.cn

² Department of Mechanical and Structural Engineering and Materials Science, University of Stavanger, 4036 Stavanger, Norway

Abstract. Autofrettage is a practical method for increasing the elastic carrying capacity and its fatigue life of thick-walled tube such as cannon, high-pressure tubular reactor. Rotational autofrettage is a recently developed technology. The analytical model for the stress analysis of a rotational autofrettage is based on the elastic-perfectly plastic assumption. Because of the effect of Bauschinger and the strain-hardening, most actual materials do not satisfy the elastic-perfectly plastic assumption. In this paper, FEA has been used to simulate the stress distributions in the tube under rotation and the residual stresses in the tube after Rotational Autofrettage. The actual tensile-compressive curve of material, which includes the material strain-hardening and the effect of Bauschinger, was adopted as the material model. Finally, the residual stress distribution caused by Rotational Autofrettage and Hydraulic Autofrettage with the same plastic deformation were compared and analyzed.

Keywords: Rotational autofrettage · Hydraulic autofrettage · Thick walled tube · Stress-strain curve · FEA · Residual stress

1 Introduction

For increasing the elastic carrying capacity, fatigue strength of the thick-walled tube, there are some technique procedures called autofrettage (such as Hydraulic, Swage, Rotational) were invented and studied. The most essential step is to accurately estimate the residual stress distribution in the tube wall after autofrettaged. Residual stress distributions can be determined by experiment or calculation. The calculation procedures include theoretical analysis [1–8], finite element analysis [9, 10]. These procedures mainly involve making simplifying assumptions about material behavior which may limit their accuracies [1–4]. For hydraulic autofrettage, many papers have been published to discuss the significance of the yield criterion, the Bauschinger effect and strain-hardening on the predicted residual stress distribution [1–6]. Rotational autofrettage is a recently developed technology. It is more suitable for reautofrettage of a rifled barrel. The process of achieving autofrettage by means of rotation of the cylinder was

first conceptualized by Zare and Darijani [7] in 2016. Kamal et al. [8] and Kamal and Perl [9] modeled an A723 steel tube of wall ratio 2.0. Parker [11] reviews their work and points out that: a. Rotational Autofretting is modeled as EPS. b. Tresca's yield criterion is employed but is based upon the maximum and minimum principal stresses within specific zones. This assumption produces slope discontinuities at the interface between the zones. As a result, the zone boundaries cannot describe the precise location of the elastic-plastic interface at peak autofretting. c. Bauschinger effect is excluded. By comparing equivalent overstrains, Zare and Darijani conclude that rotational autofretting is superior to traditional hydraulic autofretting. Kamal and Perl report general equivalence between Rotational Autofretting (RA) and Hydraulic Autofretting (HA) near-bore residual stresses. These conclusions were examined by case study comparison in Ref. [11] and will be further examined in this paper.

In this paper, we selected two tubes and candidate steels 30CrNiMo8 and ASME A 723 from Refs. [4] and [9] respectively. Then the Hydraulic autofretting of tube 1 from Ref. [4] were calculated by FEA and the calculated residual stress distribution is compared with that of theoretical and test results. Based on this verification, the tube 1 and tube 2 under HA and RA were simulated by using FEA with elastic-perfectly plastic material model and the real stress-strain curve model. Finally, the comparison was made.

2 Material Stress-Strain Relationship

Autofretting outcomes depend highly significantly upon the elastic-plastic behavior of the material of the original tube. This material behavior assumption includes behavior during the loading and unloading phases, the latter being firmly linked to the Bauschinger effect [2]. The real stress-strain curve of material in loading is somewhat different from that of in unloading, see Fig. 1(a), for performing the simulation of HA or RA of tube by using FEA, the first step is to determine the stress-strain curves of material during loading and unloading, respectively.

2.1 The Stress-Strain Curve of Materials Under Loading and Unloading

A typical candidate steel for autofretting tube is steel 30CrNiMo8. Its stress-strain curve is illustrated in Fig. 1(a). This stress-strain curve can be modeled by following method and also be simplified as elastic-perfectly plastic model, showed in Fig. 1(b).

2.2 The Stress-Strain Curve Model of Materials

A general tensile-compressive stress-strain curve of materials is shown in Fig. 2. The curve can be divided into 4 segments, O-A, A-B for loading, B-D and DE for unloading. These 4 segments can be expressed by 4 Eqs. (1)–(4).

By curve fitting of the stress-strain curve of steel 30CrNiMo8, the parameters of this material are listed in Table 1, where Real for real stress-strain curve model, EPP for elastic-perfectly plastic model.

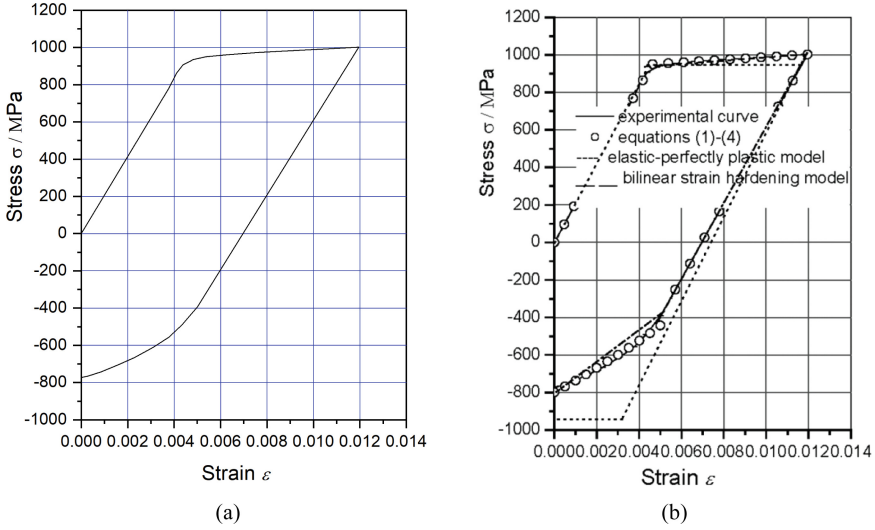


Fig. 1. Stress-strain curve of a candidate steel 30CrNiMo8 [4]

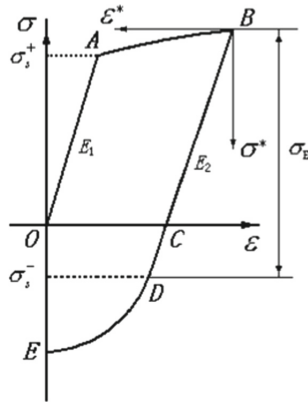


Fig. 2. General tensile-compressive curve model of material

(1) Loading

In the Cartesian coordinate system $\epsilon O\sigma$:

Elastic region OA:

$$\sigma = E_1 \epsilon \quad (\epsilon \leq \epsilon_s) \tag{1}$$

Strain-hardening region AB:

$$\sigma = A_1 + A_2 \epsilon^{B_1} \quad (\epsilon \geq \epsilon_s) \tag{2}$$

(2) Unloading

In the Cartesian coordinate system $\varepsilon^* B \sigma^*$:

Elastic region BD:

$$\sigma^* = E_2 \varepsilon^* \quad (\varepsilon^* \leq \varepsilon_D) \tag{3}$$

Strain hardening region DE:

$$\sigma^* = A_3 + A_4 (\varepsilon^*)^{B_2} \quad (\varepsilon^* \geq \varepsilon_D) \tag{4}$$

Table 1. Calculation parameters of 30CrNiMo8

σ_s (MPa)	E_1 (MPa)	A_1 (MPa)	A_2 (MPa)	B_1	σ_E (MPa)	E_2 (MPa)	A_3 (MPa)	A_4 (MPa)	B_2	REM
960.7	207,000	928.1	7026	1.0	1394	201,000	-0.54	10,850	0.405	Real
960.7	207,000				1921.4					EPP

3 Nonlinear Analysis of the Thick-Walled Tube

3.1 FEA of the Tube in HA Process and Result Verification

It is difficult to perform the FEA of the thick-walled tube under hydraulic autofretage process, because of the nonlinear behavior of the material and also the loading and unloading which obeys the different stress-strain relationship. In the FEA, two material models must be established, except for the elastic perfectly unloading. It is constantly in shortly description of the FEA procedure in papers which related to the numerical simulation of the autofretage tube by FEA [9, 10]. For thick-walled tubes, whether hydraulic autofretage or rotational autofretage is adopted, their geometry and external load or inertial load are symmetrical by the axis of the tube, which is a typical axisymmetric problem. Therefore, we can use plane axisymmetric element for modeling calculation.

FEA of the tube taken from Ref. [4] subjected to HA pressure (740 MPa) is carried out in ANSYS 2020 R1. In order to validate the FEA procedure, the FEM analysis was carried out for real elastic-perfectly plastic material with the incorporation of von Mises yield criterion, which is the same as that adopted in the analytical model [4]. Plane182 element was used to generate the mesh in the FEA model. For eliminating the effect of the end boundary, half the length of the tube is taken as 200 mm. The radii of the tube and autofretage pressure are listed in Table 2. The cylinder is free to expand in the radial and axial directions, the displacement component in axial direction $u_y = 0$ on one end (symmetry), on the other end, the displacement remains free. The mesh size is 1×1 mm. A typical geometry of the tube with displacement boundary conditions and loading condition and the calculated circumferential stress distribution along wall thickness is shown in Fig. 3(a) and (b) respectively. The comparison of the FEM solution with the analytical solution and the test data is shown in Fig. 4. The comparison shows that the FEA results are in good agreement with the analytical solution and the test data.

Table 2. Radii of the tube and the autofrettage pressure

r_i (mm)	r_o (mm)	p_a (MPa)	Rotspeed (rad/s)
19.3	43.7	740/848	10,600/11,000

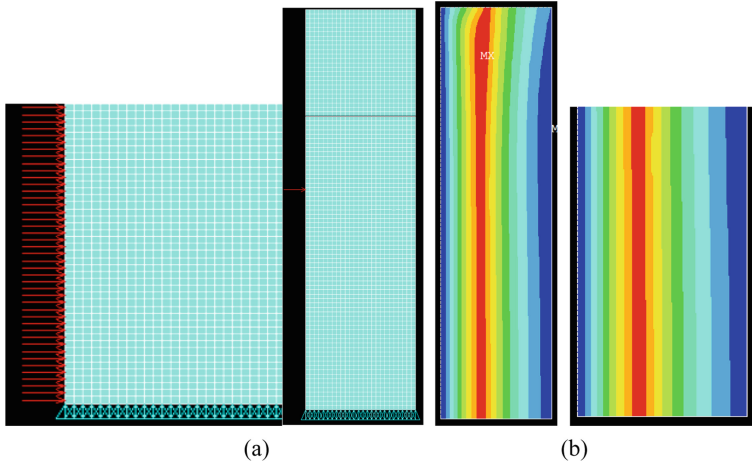


Fig. 3. (a) Finite element model with boundary condition and internal pressure, (b) distribution of circumferential stress along wall thickness

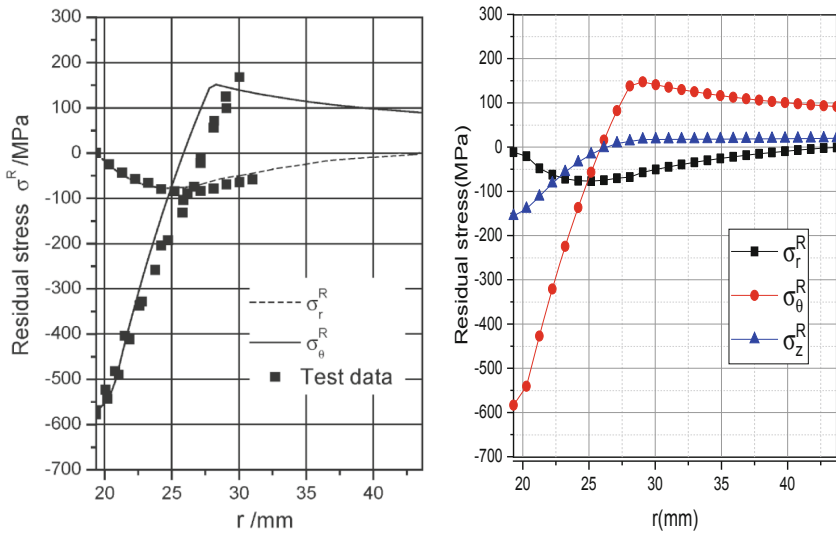


Fig. 4. Comparison the residual stress distribution by FEA with those by theoretical analysis and the test

3.2 FEA of the Thick-Walled Tubes Under RA

3.2.1 Example 1

The FE model of above tube (named tube 1) is taken as for rotational autofrettage analysis. The material model, geometry, meshed model, boundary conditions are the same, only the internal pressure was replaced by rotation speed. The overstrain level is set to 35% and 70% which corresponds to rotation speed 10,600 rad/s and 11,000 rad/s, respectively. It can be seen that the peak residual stress by elastic-perfectly plastic model is higher than that by real stress-strain curve model of the material and the effect of the material model on residual stress distribution is not significant for RA tube (Fig. 5).

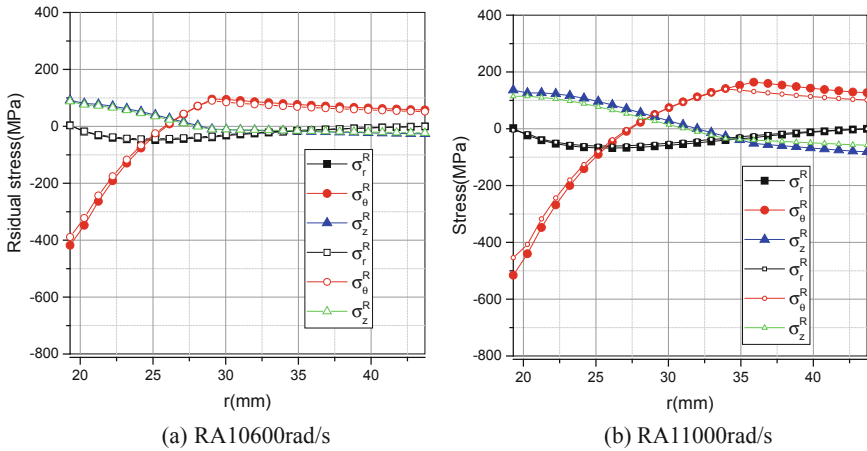


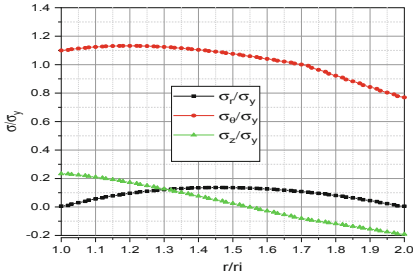
Fig. 5. Comparison of the residual stress distribution obtained by using elastic-perfectly plastic model and the real stress-strain curve of material 30CrNiMo8 (tube 1, the overstrain level is about 35 and 70%, the solid symbols indicate the results by elastic-perfectly plastic model, hollow symbols for the real stress-strain curve)

3.2.2 Example 2

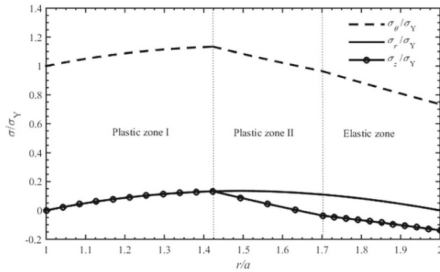
The tube is taken from the Ref. [9], The steel A723 tube with inner radius $r_1 = 60$ mm, outer radius $r_o = 120$ mm was subjected to plastically deforming rotational velocities of 3995 rad/s. Because the real stress-strain curve of steel A723 is not available in Ref. [9], here the Bauschinger effect coefficient is assumed to be 0.35. The material model parameters of A723 are listed in Table 3. The geometry of the tube with displacement boundary conditions and loading condition is like that shown in Fig. 3(a), but the internal pressure equals zero. The comparison of results by present work and by Kamal and Perl [9] with the same calculation parameters shown in Fig. 6. The difference between the corresponding curve is not significant, but the axial stress distribution. Figure 7 shows the comparison of the residual stress distribution in the thick wall tube with the elastic-perfectly plastic model and the simplified stress-strain curve of steel A723 under rotational speed 3995 rad/s.

Table 3. Material parameters of A723

σ_s (MPa)	E_1 (MPa)	Bef	σ_E (MPa)	REM
1050	200,000	0.35	1420	Real
1050	200,000		2100	EPP



(a) Results by present FEA



(b) Results taken from reference [9]

Fig. 6. Comparison of the stress distributions in the thick wall tube with the elastic-perfectly plastic model of steel A723 under rotational speed 3995 rad/s (overstrain level of 70%)

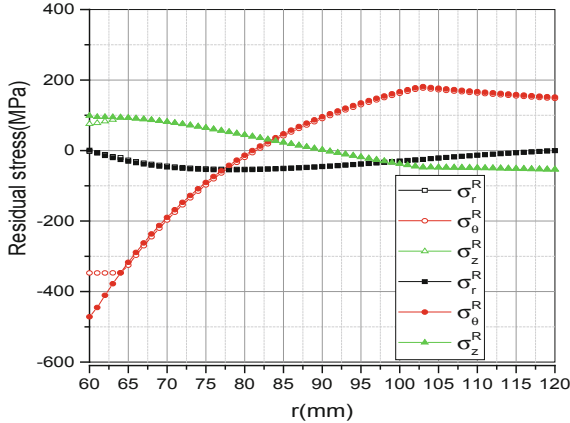


Fig. 7. Comparison of residual stress distribution in the thick wall tube with the elastic-perfectly plastic model (hollow symbols) and the simplified stress-strain curve model (solid symbols) of steel A723 under rotational speed 3995 rad/s

4 Discussion of the Effect of the Material Model on the Residual Stress in the Tube After HA and RA Treatment

With the same FE model of tube 1, we further examine the effect of the material stress-strain model on the residual stress in the tube after HA and RA treatment with the same overstrain level. The overstrain level is set to 35% ($r_c = 28$ mm, 740 MPa/10,600 rad/s)

and 70% ($r_c = 36$ mm, 848 MPa/11,000 rad/s) corresponding to the elastic-perfectly plastic model, respectively.

material model to be elastic-perfectly plastic. The FEA results are shown in Figs. 8 and 9. Comparing the residual stress distribution and the peak value of the residual stress in tube by HA and by RA treatment, the residual stress obtained from the ideal elastic-plastic model is larger than that from the actual tension-compressive stress-strain model, and the effect on HA is greater than that from the RA. The residual stress generated by HA is higher than that by RA with the same material and the same overstrain level. The conclusions obtained in Ref. [11] are also made from this study (Fig. 10).

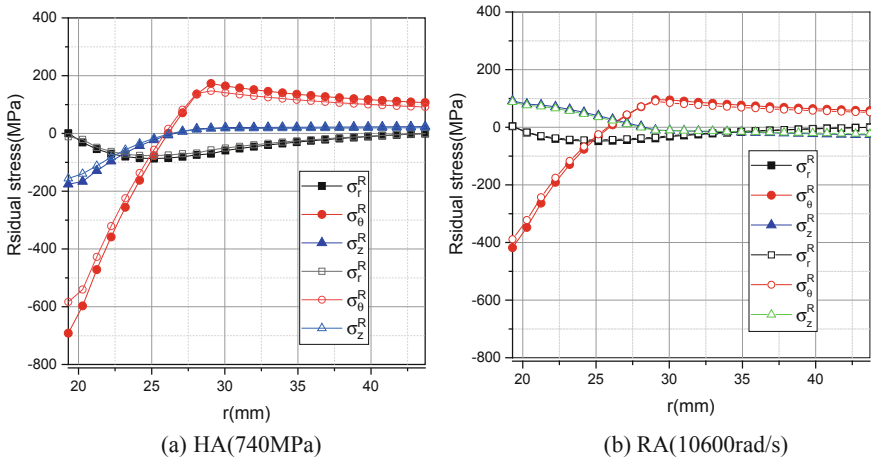


Fig. 8. Comparison of the residual stress distribution obtained by using elastic-perfectly plastic model and the real stress-strain curve of material 30CrNiMo8 (the overstrain level is about 35%, the solid symbols indicate the results by elastic-perfectly plastic model, hollow symbols for the real stress-strain curve model)

5 Summary

The effectiveness of the finite element method is verified by the finite element analysis and comparison between the theoretical calculation and the experimental results. On this basis, the finite element analysis of residual stresses in HA and RA tubes is carried out by using different simplified stress-strain curve models and the results are compared. The following conclusions are reached:

- (1) The residual stress obtained from the ideal elastic-plastic model is larger than that from the actual tension-compressive stress-strain model, and the effect on the HA is greater than that from the RA.
- (2) With the same material and the same overstrain level, the residual stress generated by HA is higher than that by RA.
- (3) The effect of the material stress-strain model on the residual stress distribution in the tube by RA can be neglected, but with a little difference in peak value.

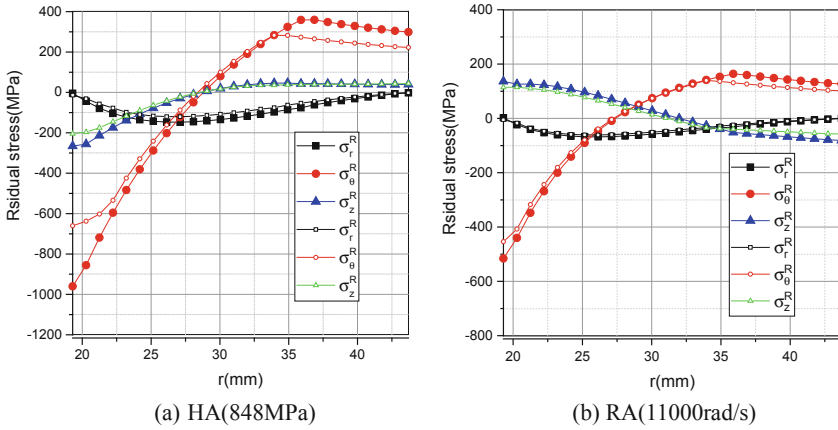


Fig. 9. Comparison of the residual stress distribution obtained by using elastic-perfectly plastic model and the real stress-strain curve of material 30CrNiMo8 (tube 1, the overstrain level is about 70%, the solid symbols indicate the results by elastic-perfectly plastic model, hollow symbols for the real stress-strain curve model)

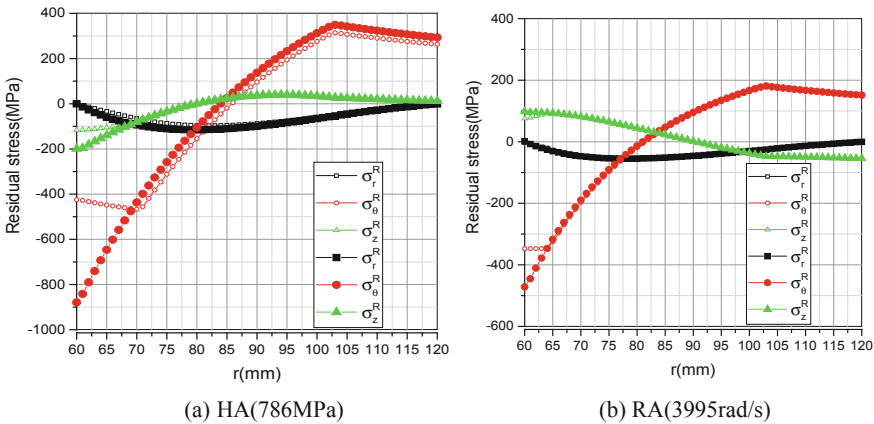


Fig. 10. Comparison of the residual stress distribution obtained by using elastic-perfectly plastic model and the real stress-strain curve of material A723 (tube 2, the overstrain level is about 70%, the solid symbols indicate the results by elastic-perfectly plastic model, hollow symbols for the real stress-strain curve model)

References

1. Chen P.C.T. The Bauschinger and hardening effects on residual stresses in an autofretted thick-walled cylinder. *Trans. of ASME*, Vol.125(2):174–182, 1986
2. Parker A.P. Autofretting of open end tubes – pressures, stresses, strains and code comparisons, *J. Press. Vessel Technol.*, 123: 271–281, 2001
3. Stacey A. Webster G.A. Determination of residual stress distributions in autofretted tubing. *Int. J. Pres. Ves. & Piping* Vol.31, 1988:205–220

4. Huang X.P. A general autofrettage model of a thick-walled cylinder based on tensile compressive stress strain curve of a material, *J. Strain Analysis for Eng. Design* Vol.40(5), 2005:599–607
5. Huang XP, Moan T. Residual stress in an autofrettaged tube taking Bauschinger effect as a function of the prior plastic strain, *J. of Pressure vessel Technology, ASME*, 2009, Vol.131:2071–7
6. Huang XP, Tian Xie, Modeling Tube-to-Tubesheet Hydraulically Expansion Based on General Stress-strain Curves of the Tube and Tubesheet Materials, *J. of Pressure vessel Technology, ASME*, 2010, Vol.132
7. Zare, H. R. and Darijani, H., “Strengthening and design of the linear hardening thick-walled cylinders using the new method of rotational autofrettage”, *International Journal of Mechanical Sciences* Volumes 124–125: 1–8, 2017.
8. Kamal, S. M., Perl, M. and Bharali, D., “Generalized Plane Strain Study of Rotational Autofrettage of Thick-walled Cylinders. Part I–Theoretical Analysis, *ASME Journal of Pressure Vessel Technol.* 141(5): 051201, 2019.
9. Kamal, S. M. and Perl, M., “Generalized Plane Strain Study of Rotational Autofrettage of Thick-walled Cylinders. Part II–Numerical Evaluation”, *ASME Journal of Pressure Vessel Technology*, 141(5): 051202, 2019.
10. Hu, Z., Gibson, M.C., Parker, A. P., Swage autofrettage FEA incorporating a user function to model actual Bauschinger effect, *International Journal of Pressure Vessels and Piping*, 191:104372, 2021.
11. Parker, A. P., Rotational Autofrettage with Bauschinger Effect: VMP Formulation, Residual Stresses and Other Features, *Proceedings of the ASME 2021, PVP2021-66396*

Diagnostic-Prognostic Methods and Structural Integrity



Identification of Unbalance in a Rotating System Using Artificial Neural Networks

Ioannis Tselios and Pantelis Nikolakopoulos^(✉)

Machine Design Laboratory, Mechanical and Aeronautical Engineering, University of Patras,
Rio Patras, 26500 Patras, Greece
itselios@ac.upatras.gr, pnikolakop@upatras.gr

Abstract. Rotating systems are one of the most common systems that transfer power in several industrial sectors. Unbalance is a severe fault that contributes to machine downtime and unscheduled maintenance actions and can damage crucial rotary systems such as turbines and compressors. Estimation of unbalance in rotor bearing systems is crucial for safe and efficient operation of the machine. The current work introduces a method that can identify the mass, angle, and location of unbalance using Artificial Neural Networks (ANNs). The inputs of the neural networks are some signal features derived from the bearing displacement signals. To generate data for the ANN training, a finite element model of a multidisc rotor bearing system supported by two journal bearings is used to simulate different unbalances (mass and angle) in the disks along the shaft. Regarding the calculation of the dynamic coefficients (stiffness and damping) of the Fluid Film Bearing, the Finite Difference Method is used for the solution of the Reynolds equation. The training data of the ANN consist of several unbalance masses and angles at two disks for 1500 rpm rotor speed. Feature selection methods used to provide the best signal features for the creation of the ANN. The results show that the ANN can accurately estimate the unbalance mass and angle of rotor. The ANN models of the current work can be part of an online condition monitoring system for fault diagnosis of rotating systems.

Keywords: Journal bearing · Artificial neural network · Fault diagnosis · Rotating machines · Rotor unbalance

1 Introduction

1.1 Unbalance as a Fault in Rotating Machines

Rotating machines are one of the most fundamental components of engineering systems and are widely used in the industry. These machines should be carefully designed to avoid catastrophic failures during operation. The vibration problems that can occur in the operational conditions may cause damage to the rotating parts of the engine or even failure of the engine completely [1]. Rotating machinery such as turbines and compressors are key equipments in oil refineries, power plants, and chemical engineering plants. Defects and malfunctions of these machines will result in economic loss. Therefore, these machines must be under constant surveillance. When a possible fault is detected, diagnosis is carried out to pinpoint the fault [2].

Rotor unbalance is the main cause of machine vibration that reduces the life of the rotating machines. The presence of rotating machines in almost all industrial equipment indicates the importance of the supervision of its normal operation. Fault diagnosis techniques allow for the detection and identification of any deviation of the operating parameters of these machines from its normal values. An even small unbalance can cause a severe problem if the operating speed of the rotor is high. Rotor unbalance is a condition in which the center of mass of a shaft and its fixed components, is not coincident with the center of rotation. In practice, rotors can never be perfectly balanced, imperfections will always be present in rotors due to manufacturing errors, despite high tolerances and sophisticated machinery used. Therefore, rotors often possess mass imbalance due to these imperfections and a mass imbalance results in an unbalanced excitation force which rotates with the rotor, i.e., at a frequency equal to the rotor spin speed [3, 4]. An effective maintenance program should provide incipient fault detection of rotors which reduces the unscheduled and excessive maintenance actions. By applying the right online condition monitoring machines, defects can be detected at an early stage [5].

1.2 Fault Detection in Rotor Bearing Systems

It is important to implement effective fault detection that can help avoid abnormal event progression, reduce offline time, forecast residual life, reduce productivity loss, and finally, avoid major system breakdowns and catastrophes [6].

Modern systems for monitoring and diagnostics of rotating machinery use, basically, classical spectral analysis of the signal of vibration displacement sensors, monitoring of other parameters of the turbogenerator (partial discharges, current, voltage, gas content, pressure, temperature) is carried out according to the nominal signal of the sensors and its comparison with the limiting characteristics [7–10].

To diagnose the state of a machine, usually signal-based monitoring systems are usually used as a good tool. After continuous study and development, signal processing technologies in Fault Diagnosis have developed into a series of methods and tools, including wavelet-based methods, empirical mode decomposition (EMD) methods autoregressive methods, cyclostationary fault detection methods and spectral kurtosis methods (SK) [11, 12]. These approaches to machinery diagnostics are generic rather than machine-specific, and data interpretation is based on qualitative rather than quantitative information [13].

The literature presents good efforts on mathematical simulation of the unbalance response of rotor systems with high accuracy [14]. Krodkiwski et al. [15] presented a nonlinear mathematical model to simulate the unbalance response of a rotor considering the nonlinear effect of its support system. Unbalance is a common and severe fault in rotating machinery, and it has been widely investigated using signal processing techniques, such as spectrum analysis, and time–frequency analysis [16]. It can be found that the amplitudes of spectrum energy at rotating speed are elevated under the different unbalance fault conditions. However, it is still difficult to interpret the severity of unbalance fault of rotor system [17]. In industry 4.0 era it is possible to have as accurate and fast diagnosing as possible, utilizing machine learning techniques and model-based monitoring systems. Model based monitoring systems give more accurate and faster information than conventional signal-based systems, since a-priori information about the rotor system is systematically included in the identification process. Therefore, the

type, position and severity of a fault can be estimated with more reliability and in most cases during operation without stopping the machine. Wang et al. [18] estimate the unbalance in rotor bearing systems (severity and position) using the equivalent load concept combined with optimization algorithms. This method provides efficient online identification of malfunctions.

1.3 Online Monitoring Using ANN

In modern times, industries consider the field of maintenance to be one of the main activities and it is necessary to develop methods for better maintenance planning. In the industry 4.0 era, traditional methods for fault detection have been sidelined, and focus is given to online monitoring system, which gives the ability to estimate faults in real time. In the current paper this perspective utilizing Artificial Neural Networks is examined. These smart systems have the ability to learn characteristics of complex systems and give results faster than traditional methods.

Condition monitoring of rotating machinery in general, involves maintenance philosophies applied for the purpose of reducing the operating and maintenance costs while, ensuring maximum operating time and achieving the highest possible production rate. An online unbalance monitoring process that would be based on artificial neural networks would offer higher accuracy.

The artificial neural network is a machine learning method that has already given sufficient results not only for fault identification but also for the prediction of the remaining useful life. Saridakis et al. [19] introduces a fault diagnosis model that uses artificial neural networks in order to identify the increase of wear depth and the increment of the misalignment angle. The proposed model uses eccentricity, attitude angle and minimum film thickness and feeds with their values an ANN that is trained to provide reliable identification of the variation of each defect.

Bouboulas et al. [20] investigate the use of ANN to predict crack characteristics of a cracked cantilever beam. In more detail, the first three natural frequencies of the beam are the inputs of the ANN and the crack characteristics, depth, and position, the outputs. The results showed that the network predicts crack characteristics with high accuracy.

Polyakov et al. [13] gives the basic principles of training a neural network for recognizing various types of failures and construct an Artificial Neural Network for the identification of a specific defect (uneven tightening of the mounting bolts to the base).

A system of neural networks will be presented that comprises one classification network and three regression networks.

2 Analysis of the Rotating System

2.1 Journal Bearing Analysis – Dynamic Coefficients

Hydrodynamic journal bearings are widely used in industry because of their simplicity, efficiency, and low cost. The geometry of the journal bearing is shown in Fig. 1, where O_b is the bearing centre, O_j is the journal center, R_b is the bearing radius, R_j is the journal radius, e is the bearing eccentricity, and L is the bearing length. The eccentricity

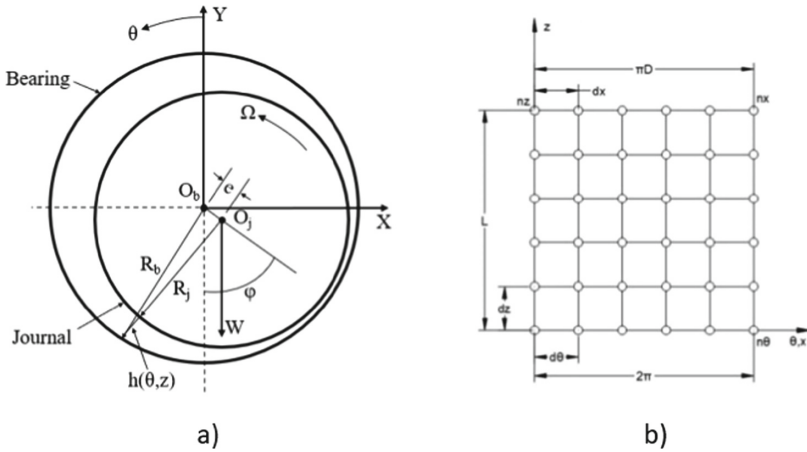


Fig. 1. a) Geometry of the journal bearing, b) grid of finite difference method

ratio is defined as $\varepsilon = \frac{e}{c}$, where c is the radial clearance. A vertical resultant load is very common (due to rotor weight). In order to calculate the dynamic coefficients of the journal bearing, it is necessary to calculate the hydrodynamic forces and the pressure distribution developed in the bearing for specific operational parameters [21]. The governing equation for the pressure distribution is the Reynold equation that can be derived under certain assumptions from the Navier-Stokes equations [22]. The Reynolds equation for a hydrodynamic bearing considering isoviscous lubricant with steady film thickness and unidirectional velocity approximations is as follows:

$$\frac{\partial}{\partial x} \left(h^3 \frac{\partial P}{\partial x} \right) + \frac{\partial}{\partial z} \left(h^3 \frac{\partial P}{\partial z} \right) = 6U\eta \frac{\partial h}{\partial x} \tag{1}$$

The equation that describes the oil film thickness is:

$$h = c(1 + \varepsilon \cos\theta) \tag{2}$$

The Reynolds equation is solved using a central difference scheme of finite difference method [23], which divides the bearing surface into a grid of small elements and discretizes the Reynolds equation into a set of algebraic equations. The hydrodynamic forces developed in the bearing can be evaluated by integrating the fluid pressure over the entire bearing area.

$$F_x = \int_{-L/2}^{L/2} \int_{\theta_1}^{\theta_2} P(\theta, z) R \cos\theta \, d\theta \, dz \tag{3}$$

$$F_y = \int_{-L/2}^{L/2} \int_{\theta_1}^{\theta_2} P(\theta, z) R \sin\theta \, d\theta \, dz \tag{4}$$

In the equilibrium state, the hydrodynamic forces must satisfy the following conditions:

$$F_x = 0 \quad (5)$$

$$F_y = F_{ext} \quad (6)$$

For a given external loading condition F_{ext} the static equilibrium position of the journal can be found from the above equations. For the determination of the equilibrium position, the numerical method Brent–Dekker [24] is used which is a hybrid algorithm that combines the bisection method, the secant method and inverse quadratic interpolation methods. Since the equilibrium position (ε, φ) and the hydrodynamic forces have been estimated, it is now possible to determine the dynamic coefficients of the bearing. The force-displacement relationship in the bearing is nonlinear but, provided that the amplitude of the displacements is small, a linear force-displacement relationship can be used, and a linear perturbation approach can be evaluated for the calculation of the stiffness and damping coefficients.

The stiffness matrix is:

$$K_b = \begin{bmatrix} K_{xx} & K_{xy} \\ K_{yx} & K_{yy} \end{bmatrix} \text{ (N/m)} \quad (7)$$

The damping matrix is:

$$C_b = \begin{bmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{bmatrix} \text{ (Ns/m)} \quad (8)$$

For stiffness coefficients, a small displacement of the journal is applied in the x, y direction and the hydrodynamic forces F_x, F_y estimated. Similarly, for damping coefficients, a small speed of the journal is applied in the x, y direction, and the hydrodynamic forces F_x, F_y estimated. Dynamic coefficients calculated in the static equilibrium position for specific operational conditions. Linear stiffness coefficients are derived from the following equations considering that the quantity $0.001c$ is quite small for the linearized behavior:

$$K_{xx} = \frac{\Delta F_x}{\Delta x} = \frac{F_{xnew} - F_x}{0.001c} \quad (9)$$

$$K_{xy} = \frac{\Delta F_x}{\Delta y} = \frac{F_{xnew} - F_x}{0.001c} \quad (10)$$

$$K_{yx} = \frac{\Delta F_y}{\Delta x} = \frac{F_{ynew} - F_y}{0.001c} \quad (11)$$

$$K_{yy} = \frac{\Delta F_y}{\Delta y} = \frac{F_{ynew} - F_y}{0.001c} \quad (12)$$

The linear damping coefficients are derived from the following equations:

$$C_{xx} = \frac{\Delta F_x}{\Delta \dot{x}} = \frac{F_{xnew} - F_x}{0.001 \times e \times \omega} \quad (13)$$

$$C_{xy} = \frac{\Delta F_x}{\Delta \dot{y}} = \frac{F_{xnew} - F_x}{0.001 \times e \times \omega} \quad (14)$$

$$C_{yx} = \frac{\Delta F_y}{\Delta \dot{x}} = \frac{F_{ynew} - F_y}{0.001 \times e \times \omega} \quad (15)$$

$$C_{yy} = \frac{\Delta F_y}{\Delta \dot{y}} = \frac{F_{ynew} - F_y}{0.001 \times e \times \omega} \quad (16)$$

The rotor is supported by two identical hydrodynamic journal bearings. Each journal bearing supports a vertical load of 7 N, the clearance is $c = 50 \mu\text{m}$ and the viscosity is $\eta = 0.03$ Pas. The length of the bearing is $L = 10$ mm, the diameter of the journal is $D = 10$ mm and the rotation speed is $\omega = 1500$ rpm. The eccentricity ratio for these parameters is $\varepsilon = 0.497$ and the attitude angle is $\varphi = 53.8^\circ$ (Tables 1 and 2).

Table 1. Stiffness coefficients of journal bearing

K_{xx} (N/m)	K_{xy} (N/m)	K_{yx} (N/m)	K_{yy} (N/m)
3.097×10^5	1.22×10^5	-5.56×10^5	4.06×10^5

Table 2. Damping coefficients of journal bearing.

C_{xx} (Ns/m)	C_{xy} (Ns/m)	C_{yx} (Ns/m)	C_{yy} (Ns/m)
2.73×10^3	-2×10^3	-2×10^3	5.91×10^3

2.2 Vibration Analysis of the Unbalanced Rotor

In a fully balanced rotor-bearing system and at moderately low rotational speeds, the only mode of motion is the rotor spinning at its designed operating speed. If the rotor bearing system is undisturbed, the system remains at its steady-state operating point. A small unbalance force will push the shaft to orbit around this equilibrium point [25]. The unbalance force is a centrifugal force: $F_u = m \times r \times \omega^2 \times e^{j(\omega t + \varphi)}$, where m the unbalance mass, r the unbalance radius, ω the rotation speed and φ the angle of unbalance mass.

The equation of motion for a rotor bearing system is presented in the following equation:

$$[M]\{\ddot{u}\} + ([C] + \Omega[G])\{\dot{u}\} + ([K] + [H])\{u\} = \{f\} \quad (17)$$

where $[M]$ is the mass matrix, $[C]$ is the damping matrix, $[G]$ the gyroscopic matrix, $[K]$ is the stiffness matrix, $[H]$ is the circulatory matrix, $\{f\}$ is the unbalance force vector, $\{u\} = \{x \ y \ \theta \ \psi\}$ is the displacement vector and Ω is the rotation speed of the rotor.

The full matrices of the system consist of the assembled matrices of the rotor and the bearings [26]. Analysis shows that the symmetric component of the stiffness matrix

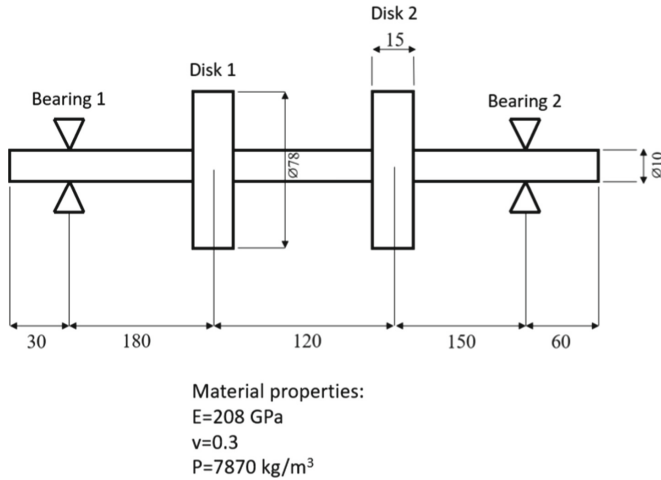


Fig. 2. Dimensions and properties of the rotor-disk system

produces conservative restoring forces, and its elements are properly called stiffnesses terms. The skew-symmetric component produces non-conservative tangential forces that can insert power into the rotor during each revolution and hence it can be destabilizing. It is clearly inappropriate to call the skew-symmetric matrix a stiffness matrix, instead, it is conventional to call it the circulatory matrix to reflect its circulatory action [27]. The rotating system of this study is depicted in Fig. 2. The rotor is simulated as 17 identical Timoshenko beams, each beam has two nodes, and each node has 4 dofs, translation and rotation in X and Y directions, respectively. Disks and bearings are modeled as concentrated parameters. The coefficients of the journal bearings are estimated in previous section. It is understandable that, if the training samples are not correct the results will not be correct either. For this reason, it is necessary to validate the model. The critical speed of the above model is 2420 rpm which is almost the same as Yao et al. [28] presented in their paper for a similar rotating system.

The orbit of the shaft on the left bearing (bearing 1) is depicted in Fig. 3 for two unbalances (0.4 g, 1.2 g) at disks 1, 2. As expected the vibration amplitude on the bearing is increasing for larger unbalance and decreasing as the unbalance force is moving away (from disk 1 to disk 2).

The main objective of this paper is the use of ANNs for unbalance identification in a rotating system. To develop such data-driven systems, a large database of vibration data is necessary. Data sets generated using the FEM model by simulating different unbalances in the two disks for a constant rotation speed of 1500 rpm. The unbalance mass takes values from 0.4 to 1.5 g and the angle of unbalance mass takes values from 1 to 360° in a radius: $r = 30 \text{ mm}$. The FEM model outputs one displacement vibration signal in horizontal direction for each bearing. Several statistical features and synchronous components will be extracted from the two signals and evaluated using feature selection methods such as Lasso, the F-test and Chi-square methods. These methods aim to efficiently select those features, for which a deep learning model can achieve better results [29]. It is well known that unbalance causes synchronous vibration.

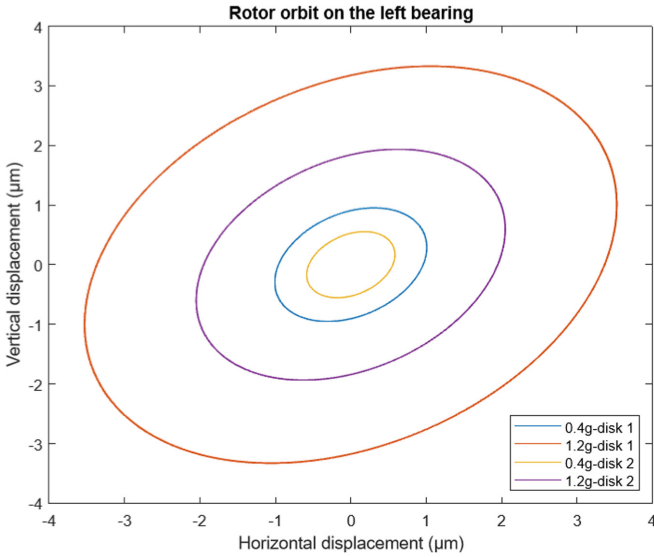


Fig. 3. Rotor orbit on the left bearing for different cases of unbalance

Therefore, for each simulation of the unbalanced state, two 1X vibration values will be obtained using the Fast Fourier Transform algorithm. These values of the algorithm are complex numbers that represent the 1X vibration magnitude and phase of rotor at each bearing.

In the Fig. 4 the vibration displacement signal is depicted in horizontal direction in time and frequency domain. As expected, due to unbalance, in the frequency domain there is a peak at the operational speed value.

2.3 Unbalance Detection Using ANN

Reddy and Sekhar [30] propose a method to identify unbalance and looseness in rotor bearing system utilizing artificial neural networks using two different methods: one by statistical features and the second by amplitude in the frequency domain. Experiments are conducted on the rotor bearing system and vibration data are collected by simulating different unbalance and looseness conditions in the rotor. ANNs are used to classify the unbalance severity and looseness.

Peyrano et al. [31] propose a method that utilizes a Multibody Simulation dynamic model of the rotor bearing system to generate vibration data from a series of representative unbalance conditions. Two machine learning techniques, an ANN and an SVR (Support Vector Regression), were trained using these data, so that the unbalance condition of the machine being modeled can be rapidly calculated from the synchronous vibration measured at its bearings. The two techniques were designed in such a way that they can receive data extracted from rotor bearing vibrations as inputs and generate a prediction of the magnitude and angle of unbalance along the shaft that gave origin to those vibrations.

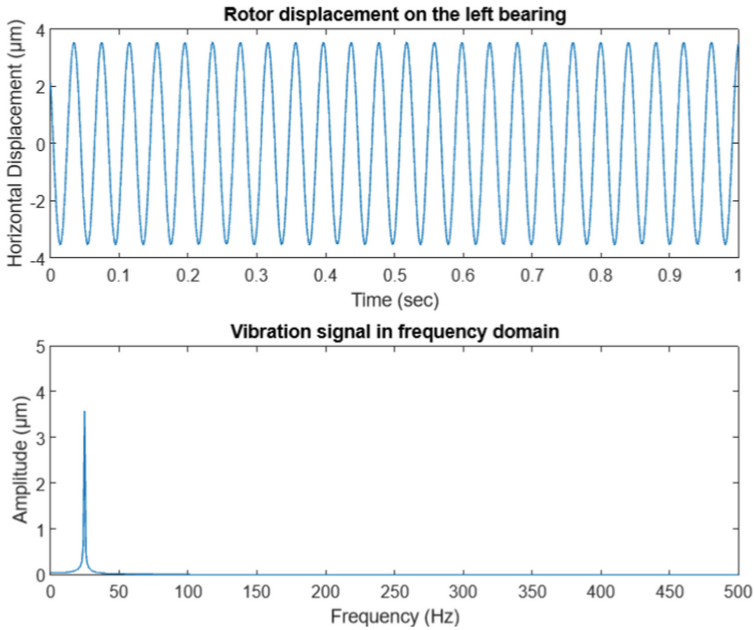


Fig. 4. Displacement signal in the left bearing in time and frequency domain

There are many different ANN types with different topologies. The one that will be used in this study is a feed-forward multilayer perceptron as shown in Fig. 5. The choice of the feed-forward multilayer perceptron is due to its simple nature and the accuracy of the results. The system of ANNs in the current work consist of four neural networks (one classification NN and three regression NN). The neural network in Fig. 5 is the first regression neural network of the system and will be explained in the following section in more detail.

For unbalance identification, a system of several neural networks will be created for localization of unbalance (disk 1, disk 2, disk 1 + 2) and estimation of unbalance mass and angle in each disk. Several records of values for input and output parameters are generated from the FEM model, as explained in the previous section. In general, an ANN is submitted to supervised training, in which the learning rule is provided with a set of examples of desired network behavior, $\{p_1, t_1\}, \{p_2, t_2\}, \dots, \{p_q, t_q\}$ with p_q to be the input to the network and t_q to be the corresponding target output. As each input is applied to the network, the network output is compared to the target. The error is calculated as the difference between the target output and the network output. The algorithm tends to the average of the sum of these errors by adjusting the weights and biases of the neural network.

In Fig. 6, the procedure for unbalance identification utilizing ANNs is depicted. At first a classification NN localizes the unbalance in the system. In the second step, and for each class a different regression neural network is trained to estimate the unbalance mass and angle at each case.

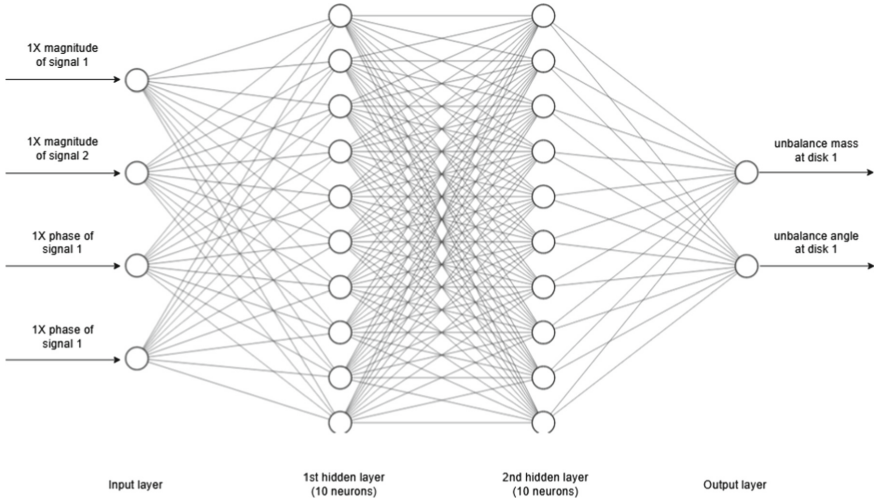


Fig. 5. Typical topology of a multilayer perceptron

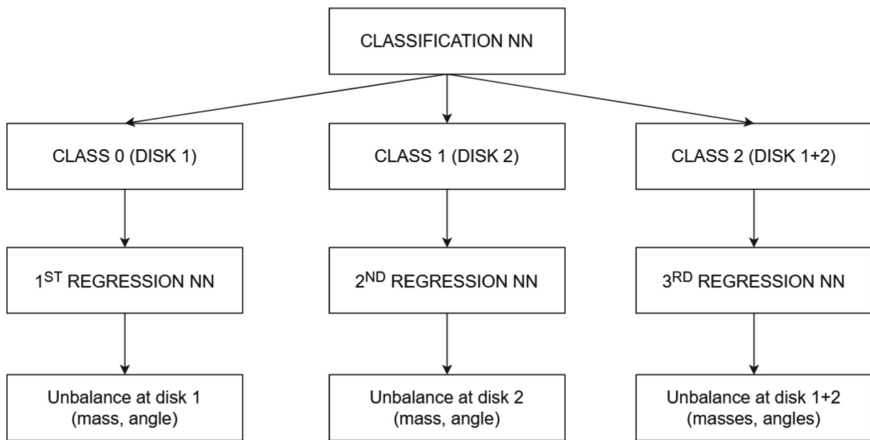


Fig. 6. System of neural networks for unbalance identification

3 Results and Discussion

In this section, the neural networks will be analyzed and the use of this system for fault identification will be evaluated. The performance of the networks will be evaluated on the test data set (data that the networks have not seen before).

The classification neural network consists of four inputs and three outputs. The inputs are the mean of the vibration signal from each bearing and the 1X vibration magnitude (from FFT) from each bearing. The loss function is the ‘sparse categorical cross-entropy’, the optimization algorithm is the ‘Adam’, the activation function for hidden layer is ReLU, and for the output layer is the softmax. The network has one

hidden layer with 50 neurons. The data set created for different unbalances on the two disks and the total number of samples is 13,500. Of these data, 70% have been used for training, 15% for validation, and 15% for testing. In Fig. 7, the confusion matrix is shown for the classification NN, and it is obvious that the network achieves great accuracy.

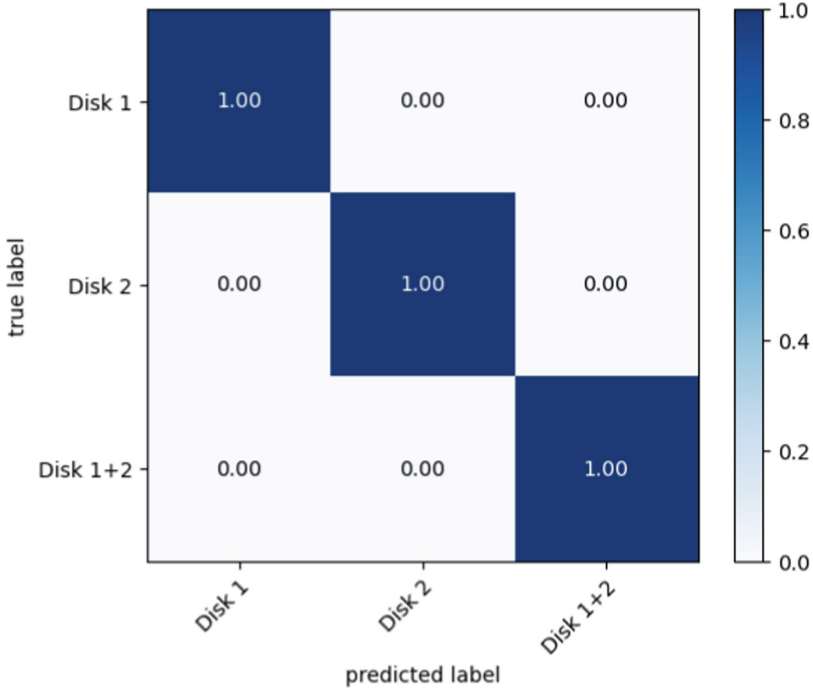


Fig. 7. Confusion matrix (accuracy) for the classification neural network

To evaluate the performance of regression neural networks, coefficient of determination R^2 and Mean Absolute Error (MAE) will be used. Coefficient of determination R^2 shows information about the goodness of fit [32]. The coefficient values are between zero and one ($0 \leq R^2 \leq 1$) and the largest value is the best value. The coefficient of determination is calculated from the following equation:

$$R^2 = 1 - \frac{\sum_{i=1}^m (X_i - Y_i)^2}{\sum_{i=1}^m (\bar{Y} - Y_i)^2} \tag{18}$$

where X_i the actual values of the dataset, Y_i the predicted values from the ANN and m the total number of training samples and \bar{Y} the mean of Y_i values.

Mean Absolute Error (MAE) is computed as follows:

$$MAE = \frac{1}{m} \sum_{i=1}^m |Y_i - X_i| \tag{19}$$

where X_i the actual values of the dataset, Y_i the predicted values from the ANN and m the total number of training samples.

The first and second regression neural networks consist of four inputs and two outputs. The inputs are the 1X vibration magnitude and phase (from FFT) of the two signals. The loss function is the ‘mean squared error’, the optimization algorithm is the ‘Adam’, the activation function for hidden layers is ReLU and for the output layer is the ‘linear’. Each network has two hidden layers with 10 and 10 neurons respectively. The dataset for the first NN created for different unbalances (mass, angle) on disk 1 and the total number of samples is 4320. Of these certain data, 70% have been used for training, 15% for validation, and the last 15% for testing. Similarly for the second NN the dataset created for different unbalances in disk 2 and the total number of samples is 4320. In Tables 3 and 4, the performance of the networks is demonstrated using the R^2 and MAE metrics. The R^2 is close to one and the MAE is very small, indicating that the neural networks have high performance.

Table 3. Performance metrics for the first regression neural network

1st regression neural network		
Metrics	Output 1 (m_1)	Output 2 (φ_1)
MAE	1.162×10^{-5} g	0.001 rad
R^2	0.999	0.999

Table 4. Performance metrics for the second regression neural network

2nd regression neural network		
Metrics	Output 1 (m_2)	Output 2 (φ_2)
MAE	8.744×10^{-4} g	0.001 rad
R^2	0.999	0.999

The third regression neural network consists of four inputs and four outputs. The inputs are the 1X vibration magnitude and phase of the two signals. The loss function is the ‘mean squared error’, the optimization algorithm is the ‘Adam’, the activation function for hidden layers is ReLU and for the output layer is the ‘linear’. The network has three hidden layers with 30, 60, 60 neurons respectively. The dataset for the NN created for different unbalances in disk 1 and disk 2 simultaneously and the total number of samples is 5200. Of these data, 70% have been used for training, 15% for validation, and the last 15% for testing. The four outputs of the network are the unbalance masses and angles at disk 1 and disk 2. In Table 5, the performance of the network is verified using the R^2 and MAE metrics.

Overfitting is a common problem in machine learning. Overfitting means that the model is over trained on the training data set, resulting in weights adjusted to the noise in

Table 5. Performance metrics for the third regression neural network

3rd regression neural network				
Metrics	Output 1 (m_1)	Output 2 (φ_1)	Output 3 (m_2)	Output 4 (φ_2)
MAE	0.008 g	0.054 rad	0.011 g	0.052 rad
R^2	0.998	0.985	0.998	0.995

the training data instead of learning “a general predicting rule” [33]. In order to achieve better performance of the networks and avoid overfitting several methods applied during the training such as early stopping, feature scaling and decaying learning rate of the optimization algorithm [34]. Specifically, early stopping method stops the training after some number of optimization steps are not leading to decreased validation loss. Feature scaling can be done for the whole dataset during the preprocessing stage, and it means that every feature in the dataset is scaled independently in the range 0–1.

In order to select an efficient architecture for neural networks, several different architectures were tested following a trial-and-error process. The proposed hyperparameters (number of layers, number of neurons, activation functions) proved to be efficient for the specific problem but probably it may not be the optimal ones, considering all aspects (convergence speed, accuracy, simplicity, computational cost). In any case, the results achieved in this study can be characterized as efficient for validating the proposed method as a powerful online fault diagnosis tool. The results show that, based on the vibration signals of the two journal bearings, the unbalance mass and angle in a rotor disk system can be fully estimated with great accuracy. The simultaneous calculation of the mass and angle of unbalance using the ANN methodology is the key point of this work.

Gohari and Kord [35] investigate a model that can identify the mass, radius, and location of an unbalanced rotating system based on the artificial neural network (ANN). It seems that the accuracy of the proposed model demonstrates good accuracy and can be developed for industrial purposes. Compared to the models of Gohari [35] and Reddy [30] the current study can not only estimate the disk (localization) and unbalance mass, but also the unbalance angle on the disk, which is a critical parameter in the balancing procedure.

4 Conclusions

The online unbalance diagnosis of a rotating system was addressed in this paper. The proposed method contributes to the identification of unbalance in a rotor bearing system with two disks. This work resulted in some important conclusions regarding the application of ANNs to unbalance estimation and condition monitoring. More specifically, the following are observed:

1. The utilization of Artificial Neural Networks provides a very effective tool for online monitoring of unbalance due to fast and accurate calculation.
2. The proposed method utilizes a FEM model of the rotor bearing system to generate vibration data from a series of representative unbalance conditions.

3. A system of ANNs is trained using these data, so that the unbalance condition of the machine being modeled can be rapidly calculated from the vibration measured in its bearings.
4. The system of ANNs consist of one classification and three regression neural network and it is able to localize and fully estimate the unbalance.

It seems that the proposed model demonstrates good accuracy and can be developed for industrial purposes. The estimation of unbalance in a rotating system gives the opportunity to change the maintenance schedule according to the real condition of the machine. Future work will deal with the creation of ANNs that can identify more defects, such as wear of journal bearing, bearing misalignment and cracks in a rotating system.

References

1. Jalali, M.H., Ghayour, M., Ziaei-Rad., S., Shahriari, B.: "Dynamic analysis of a high-speed rotor-bearing system". *Journal of the International Measurement Confederation* 53, 1–9 (2014).
2. Chen, J., Lin, C., Peng, D., Ge, H.: "Fault Diagnosis of Rotating Machinery: A Review and Bibliometric Analysis". *IEEE Access* 8, 224985–225003 (2020).
3. Tang, J., Guo, G., Wang, J., Xu, W.: "Fault injection and diagnosis of the centrifugal fan". *Journal of Physics: Conference Series* 2366(1), (2022).
4. Zhang, X., Zhang, Z., Xu, J., Si, P., Liu, S.: "A new multi-disk shaft balancing method for rotating machinery using strain gauges". *Journal of Physics: Conference Series* 2383(1), (2022).
5. Rahman, M.M., Uddin, M.N.: "Online Unbalanced Rotor Fault Detection of an IM Drive Based on Both Time and Frequency Domain Analyses". *IEEE Transactions on Industry Applications* 53(4), 4087–4096 (2017).
6. Chen, X., Wang, S., Qiao, B., Chen, Q.: "Basic research on machinery fault diagnostics: Past, present, and future trends". *Frontiers of Mechanical Engineering* 13(2), 264–291 (2018).
7. Bing, H., Zhao, Y., Pang, L., Zhao, M.: "Research on Fault Diagnosis Model of Rotating Machinery Vibration Based on Information Entropy and Improved SVM". *E3S Web of Conferences*, (2019).
8. Rai, A., Upadhyay, S.H.: "A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings". *Tribology International* 96, 289–306 (2016).
9. Odgaard, P.F., Stoustrup, J.: "Gear-box fault detection using time-frequency based methods". *Annual Reviews in Control* 40, 50–58 (2015).
10. Polyakov, R., Majorov, S., Kudryavcev, I., Krupenin, N.: "Predictive analysis of rotor machines fluid-film bearings operability". *Vibroengineering Procedia* 30, 61–67 (2020).
11. Wang, T., Han, Q., Chu, F., Feng, Z.: "Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: A review". *Mechanical Systems and Signal Processing* 126, 662–685 (2019).
12. Wang, Y., Xiang, J., Markert, R., Liang, M.: "Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications". *Mechanical Systems and Signal Processing* 66–67, 679–698 (2016).
13. Polyakov, R., Paholkin, E., Kudryavcev, I., Krupenin, N.: "Improving the safety of power plants by developing a digital twin and an expert system for adaptive-predictive analysis of the operability of gas turbine units". *Proceedings of the ASME Turbo Expo* (2020).

14. Alsaleh, A., Sedighi, H.M., Ouakad, H.M.: "Experimental and theoretical investigations of the lateral vibrations of an unbalanced Jeffcott rotor". *Frontiers of Structural and Civil Engineering* 14(4), 1024–1032 (2020).
15. Krodkiwski, J.M., Ding, J., Zhang, N.: "Identification of unbalance change using a non-linear mathematical model for multi-bearing rotor systems". *Journal of Sound and Vibration* 169(5), 685–698 (1994).
16. Yan, R., Gao, R.X., Chen, X.: "Wavelets for fault diagnosis of rotary machines: A review with applications". *Signal Processing* 96(1), 1–15 (2014).
17. Feng, Z., Liang, M., Chu, F.: "Recent advances in time-frequency analysis methods for machinery fault diagnosis: A review with application examples". *Mechanical Systems and Signal Processing* 38(1), 165–205 (2013).
18. Wang, J., Ye, L., Gao, R.X., Li, C., Zhang, L.: "Digital Twin for rotating machinery fault diagnosis in smart manufacturing". *International Journal of Production Research* 57(12), 3920–3934 (2019).
19. Saridakis, K.M., Nikolakopoulos, P.G., Papadopoulos, C.A., Dentsoras, A.J.: "Fault diagnosis of journal bearings based on artificial neural networks and measurements of bearing performance characteristics". *Civil-Comp Proceedings* 88, (2008).
20. Bouboulas, A., Nikolakopoulos, P., Anifantis, N., 2022, "Prediction of crack depth and position in vibrating beams using artificial neural networks". *Diagnostyka* 23(3), (2022).
21. Merelli, C.E., Barila, D.O., Vignolo, G.G., Quinzani, L.M.: "Dynamic Coefficients of Finite Length Journal Bearing. Evaluation Using a Regular Perturbation Method". *International Journal of Mechanical Sciences* 151, 251–262 (2019).
22. Stachowiak, G.W., Batchelor, A.W.: "Engineering Tribology", 4th Edition, Butterworth Heinemann, 2013
23. Bonneau, D., Fatu, A. and Souchet, D.: "Hydrodynamic Bearings", 1st edition, Wiley, 2014.
24. Brent, R.P.: "Algorithms for Minimization Without Derivatives". *IEEE Transactions on Automatic Control*, 19(5), 632–633 (1974).
25. Miraskari, M., Hemmati, F., Gadala, M.S.: "Nonlinear dynamics of flexible rotors supported on journal bearings - Part I: Analytical bearing model". *Journal of Tribology* 140(2), (2018).
26. Friswell, M.I., Penny, J.E.T., Garvey, S.D., Lees, A.W.: "Dynamics of rotating machines". 1st edition, Cambridge University Press, 2010
27. Nelson, F.C.: "Rotor dynamics without equations". *International Journal of COMADDEM* 10(3), 2–10 (2007).
28. Yao, J., Liu, L., Yang, F., Scapra, F., Gao, J.: "Identification and optimization of unbalance parameters in rotor-bearing systems". *Journal of Sound and Vibration* 431, 54–69 (2018).
29. Chandrasekhar, G., Sahin, F.: "A survey on feature selection methods". *Computers and Electrical Engineering* 40(1), 16–28 (2014).
30. Reddy, M.C.S., Sekhar, A.S.: "Identification of unbalance and looseness in rotor bearing systems using neural networks", 15th National Conference on Machines and Mechanisms, (2011).
31. Peyrano, O.G., Vignolo J., Mayer R., Marticorena M.: "Online unbalance detection and Diagnosis on Large flexible rotors by SVR and ANN trained by Dynamic Multibody Simulations". *Journal of Dynamics, Monitoring and Diagnostics* 1, 139–147 (2022)
32. Ikumi, T., Galeote, E., Pujadas, P., De la Fuente, A., López-Carreño, R.D.: "Neural network-aided prediction of post-cracking tensile strength of fibre-reinforced concrete". *Computers and Structures* 256, (2021).

33. Dietterich, T.: “Overfitting and Undercomputing in Machine Learning”. *ACM Computing Surveys (CSUR)* 27(3), 326–327 (1995).
34. Pang, B., Nijkamp, E., Wu, Y.N.: “Deep Learning with TensorFlow: A Review”. *Journal of Educational and Behavioral Statistics* 45(2), 227–248 (2020).
35. Gohari, M., Kord, A.: “Unbalance rotor parameters detection based on artificial neural network”. *International Journal of Acoustics and Vibrations* 24(1), 113–118 (2019).



A Finite Element Model of a Steel Truss Bridge Validated with Controlled Load Test Data

Kris Hectors^{1,2} , Lien Saelens² , Jona Bracke², Hans De Backer³ ,
and Wim De Waele³ 

¹ SIM Vzw, Technologiepark 48, 9052 Zwijnaarde, Belgium

Kris.Hectors@ugent.be

² Faculty of Engineering and Architecture, EMSME, Soete Laboratory, Ghent University, Technologiepark 46, 9052 Zwijnaarde, Belgium

³ Faculty of Engineering and Architecture, Department of Civil Engineering, Soete Laboratory, Ghent University, Technologiepark 60, 9052 Zwijnaarde, Belgium

Abstract. This work presents the development of a finite element model of a welded railway bridge constructed in 1955. This model will be used as the main tool for a structural health monitoring system of this ageing infrastructure. Such a system requires load monitoring for which the bridge was instrumented with fiber optic Bragg grating sensors spread over two longitudinal and two transverse girders. After the instrumentation, a series of passages using a locomotive with known axle loads and geometry was performed, during which strains were continuously recorded. A comparison between the experimental and numerical results showed excellent agreement for all sensors located remote from the bridge bearings. The comparison also showed that the influence of the bridge bearings on local deformations differs from the hypothesized design, especially at high load. It was found that deformation of structural elements is strongly constrained to the loaded sectors. The truss design efficiently transfers the loads on the longitudinal girders through the closest transverse girders onto the main trusses.

Keywords: Steel railway bridge · Fiber optic Bragg grating sensors · Structural health monitoring · Finite element model

1 Introduction

During the 1950s and 1960s, a surge took place in the construction of civil infrastructure [1]. Many of these structures are still in operation and are reaching their intended design life. A survey conducted within the scope of the European project ‘Sustainable Bridges’, revealed that 75% of Europe’s steel railway bridges were over 50 years old and almost 35% of them were over 100 years old [2]. Over time, allowable axle loads, traffic capacity and train speeds have increased to make railway transportation more competitive [3]. The combination of an ageing infrastructure and increasing performance requirements leads to an increase of structural components with serious fatigue damage. In 2008, a report by the Joint Research Centre of the European commission [4] attributed 38.3% of all failures of metallic bridges to fatigue failure, which was expected to increase even more.

A study by Dinas et al. [5] has shown that between maintenance and total replacement of old bridges, the former is undeniably the most sustainable in an environmental, social and economic context. To develop an effective maintenance plan, adequate knowledge of the structural behavior, daily traffic loads and the current condition of infrastructure are required.

In engineering practice, governing stresses are mostly calculated using simplified methods, combined with a standardized load spectrum and code-specified fatigue strength to determine the fatigue life. For ageing infrastructure this often results in a negative remaining useful life. An alternative method to determine load induced stresses, is through field measurement of strains at identified fatigue critical members using strain gauges [6]. For many bridges it is however not feasible to instrument and monitor all potentially fatigue critical members. The combination of a finite element model with in-situ strain measurements allows for accurate simulation of the global structural behavior, with a limited number of local measurements. An experimental study was performed by Spyarakos et al. [7] and Ermopoulos and Spyarakos [8] to develop a validated analytical model. Malm and Andersson [9] measured dynamic effects on a tied arch railway bridge and developed a 3D finite element model including dynamic vehicle–bridge interaction. Caglayan et al. [10] presented a case study for a steel railway bridge where dynamic and quasi-static measurements were used to calibrate a finite element model and propose a strengthening scheme. Leander et al. [11] undertook an extensive measurement campaign on the Söderström railway bridge in Sweden, comprising non-destructive testing and field measurements. A preceding nominal stress based analysis showed that the stringers were critical details with only a few years of remaining fatigue life. A thorough non-destructive testing inspection did not show any fatigue related damage however. Nonetheless, an analysis based on stresses obtained from the strain gauge measurements resulted in the same conclusions. Leander et al. suggested to re-analyze the results using a non-linear damage accumulation approach due to the shortcomings of the linear fatigue damage accumulation rule, i.e. Miner's rule (see [12, 13]).

Fatigue assessment of steel (railway) bridges has been studied extensively in literature, and as a result a number of comprehensive review articles have been published. Zhang et al. [14] published a review article on fatigue problems in orthotropic bridge decks, Aygül et al. [15] on fatigue failure assessments of welded bridge details and Alencar et al. [16] on fatigue cracking of welded railway bridges. A more general review on fatigue life assessment of steel bridges was given by Ye et al. [17]. A common conclusion is that the stress-life method is the most widely used for fatigue related design and assessment of steel bridges, but that the hot spot stress method (see [18]) has been proven more accurate and effective. An important constraint of the hot spot stress method is that it requires a two or three dimensional finite element model to be constructed. Li et al. [19] were amongst the first to develop a multi-scale finite element model where shell substructures are coupled with beam elements. This allows efficient analysis of the overall structure and the ability to analyze the coupled substructure joints with the hot spot stress analysis. Similar studies were performed by Wang et al. [20, 21].

Although the fatigue assessment of steel (railway) bridges has clearly evolved towards a multi-scale modelling approach, the accuracy of a finite element model strongly relies on experimental validation [22]. Validation is generally achieved through

quasi-static measurements, dynamic measurements or a combination thereof [8, 23, 24]. Dynamic measurements are generally used to calibrate the stiffness and mass matrices of a finite element model. Although this is an effective method to obtain a calibrated finite element model, the resulting model matrices have little physical meaning [25] and thus do not contribute to the general understanding of the influence of specific joints to the overall behavior. Quasi-static experiments, mostly based on strain measurements, provide a major advantage in this regard. Advances in fiber optic sensing have led fiber optic strain sensors to become a strong alternative to classical strain gauges for monitoring of civil infrastructure as they offer many advantages [26–28]. These include ease of installation, non-corrosive, immunity to electromagnetic interference and lower maintenance cost. Possibly the most well-known example of fiber optic based structural health monitoring of a bridge is the Tsing Ma bridge [29]. It has been demonstrated that fiber optic sensors can be used to detect the dynamic strain response induced by passing trains. Comparison with resistance strain gauge measurements has shown excellent agreement. Another example using FBGs is the work of Barbosa et al. [30] who used weldable FBGs for load tests and in-service monitoring of a pedestrian bridge in Aveiro, Portugal. Matta et al. [31] presented the use of distributed fiber optic sensors (DFOS) for strain measurements on the steel girders of a multispan highway bridge in Missouri, U.S., which was subjected to a diagnostic load test. Enckell et al. [32] used DFOSs to measure the strain profile and detect cracks along the length of the continuous steel girder Götaälv bridge in Sweden. Costa and Figueiras [33] installed electric and fiber optic strain sensors on the steel Trezói railway bridge in Portugal to characterize its structural behavior. Xu et al. [34] presented a full scale field monitoring campaign of a 1108 m suspension bridge using DFOSs to measure the bridge behavior during load tests with multiple trucks, demonstrating the feasibility for large scale infrastructure. Van Der Kooi et al. [35] instrumented a steel railway bridge, constructed in 1902, with a DFOS based system and performed three different loading scenarios using a train with known dimensions and weight. These are just a few examples showing the effectiveness of fiber optic sensor based systems. For a more exhaustive review, reference is made to [36, 37].

In 2018, the Flemish SafeLife project kicked off, which focused on large scale welded steel structures subjected to fatigue loading. The goal of this project was the development of load and health monitoring procedures and numerical tools that enable a more accurate lifetime prediction of welded steel structures. One of the structures that was studied in the project is the steel truss railway bridge over the river Scheldt in Temse, Belgium. Two major goals of the project addressed in this paper are the installation of an FBG based load monitoring system and the development of a validated finite element model of the bridge based on the output of that system. The combination of both should become one of the main building blocks in a structural health monitoring system for the ageing bridge. The bridge was instrumented with fiber optic Bragg grating sensors distributed over two longitudinal and two transverse girders. After the instrumentation, controlled load tests were performed using a single locomotive during which strains were continuously recorded. The results of these measurements could be used for the validation of the finite element model.

This paper is structured as follows. In the first section the studied railway bridge is described. Next the developed finite element model is discussed. An overview of the measurement campaign is then provided. Finally the results are discussed and an overview of the conclusions is provided.

2 The Temse Bridge

The Temse bridge is a steel through-truss bridge that was built in 1955. It has a total span of 365 m, subdivided in intermediate spans ranging from 36.85 to 81.7 m. The bridge is subdivided into a part for railway traffic and a part for road traffic. Until 2008, the bridge had the largest overall span of all bridges in Belgium. Because the bridge crosses the river Scheldt, a part of the bridge can be opened to let large vessels pass through. The movable part of the structure is a bascule type bridge and initially had a span of 33.2 m. In 1961, the span of the movable bridge section was increased to 54.4 m in order to facilitate the waterway traffic. While the original structure was constructed with riveted connections, the new movable section was constructed with welded connections. In 1997, additional plates were welded on the top flange (TF) of the longitudinal girders (LG) for reinforcement, as a mitigation to fatigue problems that had started to arise. In 2005, a new bridge was built next to the original one, as the old Temse Bridge was no longer able to fulfil functional demands. Four car lanes, two in each direction, had to merge into two lanes on the bridge, leading to daily traffic jams. The newly constructed bridge is 9 m longer than the original bridge, thereby taking the title of longest span bridge in Belgium. Figure 1 shows an aerial view of the railway bridge, the old road bridge and the new road bridge.



Fig. 1. Aerial view of the bridge, showing (from left to right) the studied railway bridge, the old road bridge and the newly constructed road bridge.

This work focuses on the welded (movable) section of the bridge, as indicated with a red dashed box in Fig. 2. This movable bridge section can be subdivided into four groups of structural components: (1) two main trusses, which are the main load bearing elements; (2) longitudinal girders, which transfer the loads from the tracks to the transverse girders (TG); (3) transverse girders, which transfer the loads from the longitudinal

girders towards the main trusses; (4) upper wind braces between the main trusses, and bottom wind braces between the longitudinal and transverse girders. The height of the main trusses varies from 7.6 m at its highest to 3.1 m at its lowest point. The total span of the structure is 54.4 m, the distance in between the transverse girders is 5.4 m, and the main trusses are placed 5.5 m apart.

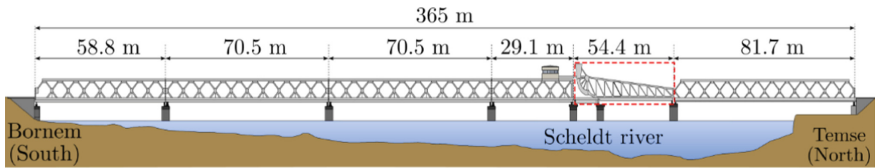


Fig. 2. Schematic lateral view of the Temse railway bridge. The studied section is indicated with the red box.

In modern construction in Belgium, the importance of labor costs generally exceeds that of the material and production costs. Contemporary structures will therefore show an increased degree of uniformity to ensure a quick and correct assembly on site. The movable part of the bridge however, was assembled in 1961 when labor costs were not at the same level as in present days. This shows in the composition of the main truss elements as fourteen different cross-sections have been used for its construction, adding to the complexity of the structure. A mix of built-up members and I-beams have been used for the main trusses. An overview of the built-up members in the main trusses is shown in Fig. 3; the diagonal cross-members comprise different types of DIN I-beams. At the south side, the main truss ends in a quarter-circular sector that is primarily designed to open and support the bridge whilst opened for passing ships. Because of the importance of the trusses as the main load-bearing elements, the circular sector will also influence the structural behavior when the bridge is closed.

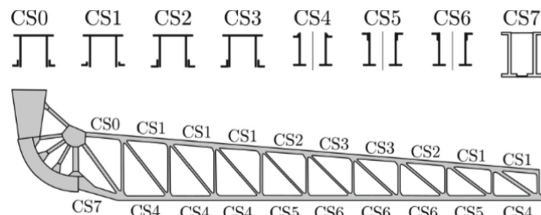


Fig. 3. Schematic representation of the different cross-sections used in the main trusses.

A top view of the bottom structure that carries the rails and sleepers is shown in Fig. 4. The two longitudinal girders are continuous I-girders that span the length of the movable section. The top flange of the longitudinal girder is fully continuous, passing through cut-outs in the transverse girders, whereas the web and bottom flange (BF) of the girder are welded to the transverse girders. The top flange of the longitudinal girders has been reinforced with welded cover plates on both the top and bottom where it passes through

the cut-outs in the transverse girder. The transverse girders have a variable cross-section, widening significantly towards the girder ends. Both the top and bottom flanges of the transverse girders are reinforced by welded cover plates. Figure 5 shows the transverse girder and a detailed view of the cut-outs for the top flange of the longitudinal girder. The studied bridge section is fully comprised of welded connections.

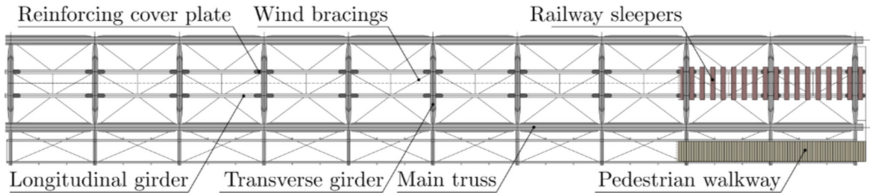


Fig. 4. Schematic top view of the bottom structure showing the longitudinal girders, transverse girders and centerlines of wind bracing elements.

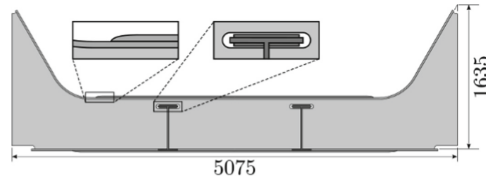


Fig. 5. Illustration of the transverse girder geometry with a detailed view of the cutouts where the top flange of the longitudinal girder passes through.

3 Finite Element Model

3.1 Structural Elements

A three dimensional finite element model of the studied bridge has been developed in Abaqus CAE (v2020). The structure contains a large number of non-standard structural joints each with a different geometry due to the large variation of structural elements. The use of shell elements allows for an accurate representation of these joints, which is required to assess the magnitude of stress concentrations occurring in these regions. The shell model is based on the mid-planes of the actual structure. Cover plates are included in the model as ‘skins’. Cope holes at the connections between the longitudinal and transverse girders are also included in the global model. Weld fillets were omitted as this study focusses on the global structural behavior, for the same reason, the sleepers and rails are not included in the model.

In total, 95 unique parts were modelled and assembled to construct the model. After assembly of all parts, the assembly was merged to create a single entity. In this way mesh continuity across the model is ensured. The assembled model is shown in Fig. 6.

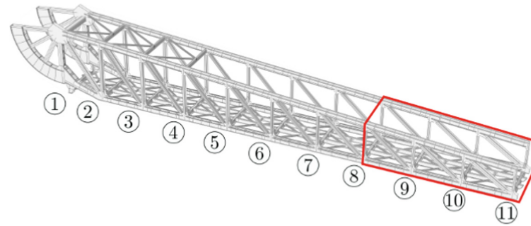


Fig. 6. The global finite element model of the movable section of the Temse railway bridge

3.2 Material

The steel type defined in the technical drawings is A. 52 H.S. Grade II; a steel type based on an outdated standard which best corresponds to the present day S355 construction steel. Conventional elastic properties are assumed. Specifically, a Young's modulus of 210 GPa and a Poisson's ratio of 0.3.

3.3 Boundary Conditions

No technical drawings remain of the bridge bearings, but a number of photographs were available from on-site inspections. When the bridge is opened, it is rolled on the quarter-circular section by means of a gear rack. The quarter-circular section then supports the whole structure. When closed, the bridge is supported at both ends by three bearings positioned below the main trusses and at the center of the outermost transverse girders (numbers 1 and 11 on Fig. 6). These bearings transfer the forces to the concrete foundation pillars. Figure 7a shows the northeastern bridge bearing and Fig. 7b shows the northwestern bridge bearing. Because of lack of information and to keep the model sufficiently simple, the actual bridge bearings are not modeled in this work. This will have an effect on the simulated deformations of structural members located near these bearings, but should not influence the results for members remote from the bearings.

In the numerical model, each support is assumed to have unrestricted rotational degrees of freedom, representing a simply supported beam configuration. Therefore, each boundary condition is modelled by constraining displacements at only a single mesh node. Constraining the displacement of all nodes on the bearing contact surface would correspond to a clamped support. Initially the definition of the model boundary conditions was based on the design hypotheses. An overview of the boundary conditions initially used model is shown in the top illustration of Fig. 8. Experimental results however showed these to be a bad representation of the actual structural behavior. The boundary conditions were then calibrated such that a better agreement between the numerical and experimental results was achieved. An overview of the boundary conditions used in the final model is shown in the bottom illustration of Fig. 8. The calibration of the boundary conditions is discussed in detail in Sect. 5.3.

3.4 Mesh

The whole model has been constructed using 8-node doubly curved thick shell elements, with reduced integration (S8R). The model is meshed using a global mesh size of 75 mm.



(a) Simply supported bridge bearing at transverse girder end



(b) Bridge bearing at transverse girder end allowing longitudinal movement

Fig. 7. Bridge bearing at transverse girder end allowing longitudinal movement.

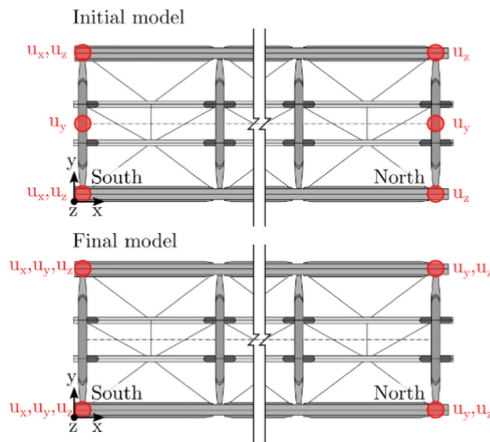


Fig. 8. The location of the bridge bearings and the associated constrained degrees of freedom, illustrated on a top view of the bottom structure.

This resulted in the best compromise between accuracy, modelling time and calculation time. Larger mesh sizes required additional partitioning in order to obtain a good quality mesh, whilst smaller mesh sizes did not result in significant differences in calculated vertical displacement. Decreasing the mesh size to 50 mm resulted in a maximum deviation of 0.3% in the vertical displacement of the longitudinal girder for a vertical unit load at the center of the bridge. The final global model contains 714 764 nodes and 245 218 elements.

3.5 Load Model

Train loads are applied to the finite element model in a progressive step wise manner. At each simulation step, the response of the structure is calculated in static, linear elastic conditions. Since the sleepers and rails are not modeled, the train loads are directly applied to the top flanges of the longitudinal girders. To avoid the introduction of large stress singularities in the model, vertical pressure loads are distributed over the area corresponding to the real contact area between the sleepers and the top flange of the longitudinal girder. Only vertical quasi-static loads are considered in this study.

The implementation of the rolling load has been realized using the Abaqus DLOAD subroutine, which allows the variation of loads in space and time. The subroutine is defined in a Fortran script; the number of simulations steps is defined in Abaqus.

4 Measurement Campaign

Within the scope of the SafeLife project, a total of 92 fiber optic Bragg grating (FBG) sensors have been installed, covering three sectors of the movable section indicated with a red box in Fig. 6. The global finite element model of the movable section of the Temse railway bridge. Both longitudinal girders have been instrumented with ten sensors on the top side of the bottom flanges. Only the most northern sensors on the longitudinal girders (blue n°1 in Fig. 9) were installed on the bottom side of the flanges for practical reasons. Transverse girders n°10 and n°11 have each been instrumented with 4×9 FBG sensors, two sensors lines on the top flange and two on the bottom flange. With these sensor lines the global mechanical behavior of the girders can be measured. The choice to instrument transverse girder n°11 was made to study the effects of the bridge bearings (i.e., the boundary conditions) on the behavior of this girder. The use of two sensor lines on both the top and bottom flanges allows for measurement of any asymmetric behavior of the transverse girders.

In Fig. 10, the strain measured as a function of time originating from a single sensor located on a longitudinal girder is shown. Distinct peaks can be observed corresponding to the bogies passing over the sensor location. As such, the number of train wagons can be determined, which can be used in future work to build a representative long term load spectrum for remaining life assessment.

4.1 Controlled Load Test

A controlled load test was been conducted with a locomotive of known weight and dimensions to for the purpose of validating the global model. A locomotive of type EM130 with known axle loads and geometry was used to conduct a series of controlled test runs. The wheel loads of the four axles at the moment of weighing are listed in Table 1. An additional 1200 L of fuel was added after the weighing, before the locomotive departed to the measurement site.

Figure 11 shows a schematic of the locomotive used for the controlled load test.

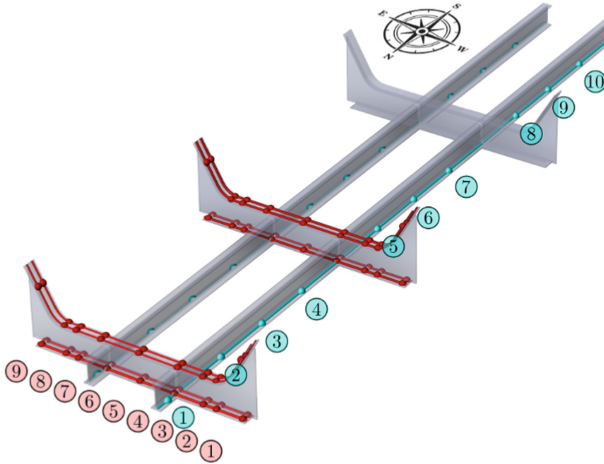


Fig. 9. Sensor lines and sensor locations on the transverse girders (red) and longitudinal girders (blue).

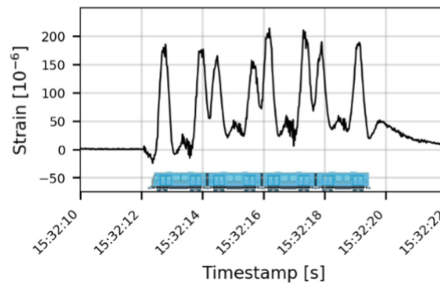


Fig. 10. Example strain measurement from sensor 5 on the western longitudinal girder.

Table 1. EM130 weight distribution before refueling as weighed by the weight bridge.

	Bogie 1		Bogie 2	
	Axle 1 (kg)	Axle 2 (kg)	Axle 3 (kg)	Axle 4 (kg)
Wheel 1	9370	9070	7420	7520
Wheel 2	8650	9170	7420	7550

5 Model Validation

5.1 Assumptions

A passage at 20 km/h with the EM130 locomotive is used for the validation of the model as the dynamic effects are negligible in this case. The EM130 locomotive has been included in the aforementioned DLOAD subroutine for the model validation.

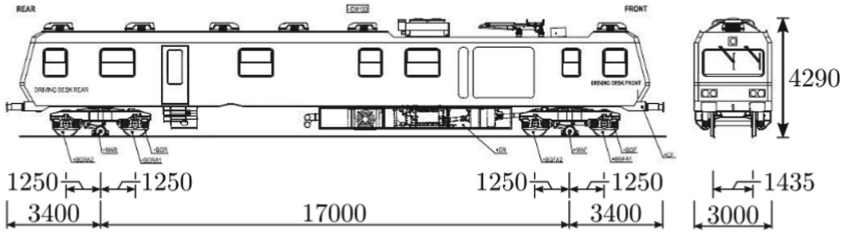


Fig. 11. Schematic of the EM130 locomotive used for the controlled load tests.

The loads are transferred from the rails to the longitudinal girders and then to the transverse girders. Thus, the first step in the validation is the comparison of the numerical and experimental strain results at the longitudinal girders.

5.2 Longitudinal Girders

Three statistical parameters are used to assess the model performance in this subsection and the two thereafter. The root mean square error (RMSE), maximum absolute error (MAE) and the relative absolute error (RAE). The first two are well-known, the RAE less so. The RAE can be defined as:

$$RAE = \frac{\sum_{i=1}^n |P_i - E_i|}{\sum_{i=1}^n |E_i - \bar{E}|} \quad (1)$$

where P_i is the value predicted by the finite element model for record i . E_i is the experimental value for record i and \bar{E} the mean of all considered experimental records. $RAE = 0$ represents a perfect fit. If $RAE > 1$, the model predictions are worse than a model that is simply equal to the mean of the experimental values. The advantage of the RAE metric is that it allows comparison of measurements with different magnitudes.

Figure 12 shows a comparison between numerical (in blue) and experimental (in black) strain histories, extracted at different sensor locations on the longitudinal girders during the passing of the EM130 train. The finite element simulation at sensor location 10, at the northern end of the girders near the bridge bearings, exhibits the worst agreement. Since $RAE > 1$ the model cannot describe the strain-time behavior close to the bridge bearings for either of the longitudinal girders. For the other sensor locations a much better agreement ($RAE = 0.4-0.5$) between the numerical and experimental behavior is found. Based on the statistical measures reported in it can be seen that the model agreement is better for the eastern longitudinal girder than for the western one. The most pronounced differences in strain occur when the peak values are reached. Since the sleepers and rails were not modeled, and the load is directly applied to the location where the strains are extracted from the model, the occurrence of larger peaks is not unexpected. But, given the observation that the differences remain small, the implemented load application can be justified as it significantly simplifies the model. A common observation for all sensors is that the numerical model overestimates the strains.

Figure 13 shows the comparison of the numerical and experimental strain distribution along the longitudinal girder, for a snapshot in time corresponding to 1.86 s in

Fig. 12. The comparison of these results will be influenced by a time shift between the experimental and numerical results because the locomotive did not travel at a constant speed whilst passing over the sensor locations. To illustrate the effect hereof, the area in between the results corresponding to the preceding and subsequent numerical time steps ($\Delta t = 0.09$ s) are plotted, representing an uncertainty band. The black dots correspond to the experimental strain results at different sensor locations. Each arrow corresponds to a single train axle. The disturbance that is observed in the numerical strain results, corresponds to the locations where the longitudinal girders pass through the transverse girder. The strain distribution shown in Fig. 13 clearly illustrates how loads are transferred from the longitudinal girder to the transverse girders as the deformation of the longitudinal girder sections in the neighboring non-loaded sectors remains limited. Overall an excellent agreement is found which is reflected by the statistical parameters reported in Fig. 13.

5.3 Transverse Girder 1

The first transverse girder, which is the most northern one of the movable section, is directly connected to the bridge bearings. Hence the agreement between the experimental and numerical results will strongly depend on how accurate the boundary conditions in the finite element model represent the real bridge bearings. Figure 14 shows the initial results for a comparison between the experimental and numerical strain distributions in the considered transverse girder when the front bogie of the train is at the center of the first instrumented sector. Although the finite element model appears to model the top flange behavior quite well, both the shape and magnitude of the numerical results for the bottom flange (shown in orange) are far off. The RAE of the residuals is greater than one, indicating the model cannot explain the experimental behavior. These results were obtained from a finite element model where the boundary conditions of the model were defined based on the hypothesized behaviors of the bridge bearings. For the bearing shown in Fig. 7a it was assumed that it only restricts vertical translations (z direction) and for the bearing shown in Fig. 7b vertical and transverse translations (z and y directions). The experimental results, however, show that considerable strain is building up at the edges, whilst the numerical values fade out to approximately zero strain. The assumption that the transverse movement of the girder is unrestricted seems false. Therefore, the boundary conditions representing the bearing shown in Fig. 7a were modified such that both vertical and transverse translations (z and y directions) were restricted. The full set of boundary conditions then corresponded to the overview shown in Fig. 8.

Figure 15 again shows the strain distributions for the transverse girder when the front bogie of the train is in the center of the first instrumented sector, but for the modified boundary conditions. As can be seen in the figure the overall shape and magnitude of the numerical and experimental strain distributions agree much better than before. This is also confirmed by the statistical parameters which are reported in the same figure. The model predictions have significantly improved for both the top flange ($RAE = 0.533 \rightarrow 0.618$) and the bottom flange ($RAE = 1.084 \rightarrow 0.444$). In both the numerical and experimental results a strain peak can be seen in the results of the top flange at 1.716 m. This corresponds to the location where the cover plate on the top flange of the transverse girder ends.

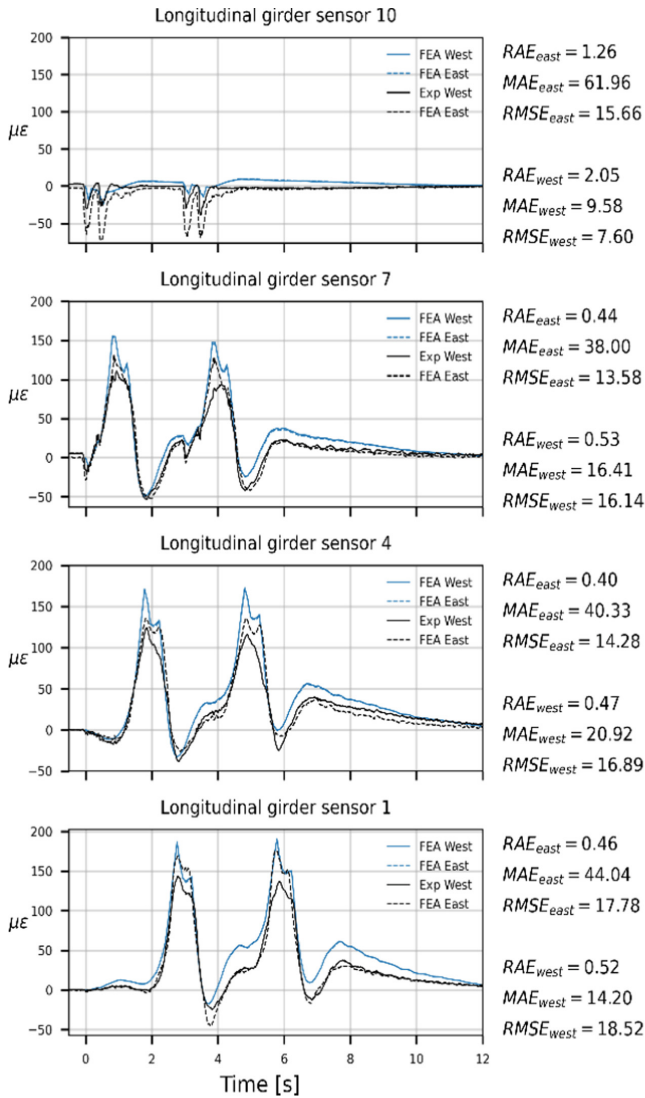


Fig. 12. Comparison between numerical (FEA) and experimental (Exp) strain histories for the longitudinal girders. Descriptive statistical measures of the residuals are shown next to the plots.

Despite this significant improvement, there remains an obvious discrepancy between the experimental and numerical strain distributions at the center of the transverse girder. The central sensors measure strain values significantly lower than those of the neighboring sensors; an effect that is not seen in the numerical measurements. Based on Fig. 16 it is assumed that this is most likely caused by the following. At the center of the bottom flange, a large steel coupling piece with pin has been attached to the bottom flange. When the movable section of the bridge closes, this pin (d) is guided in between two

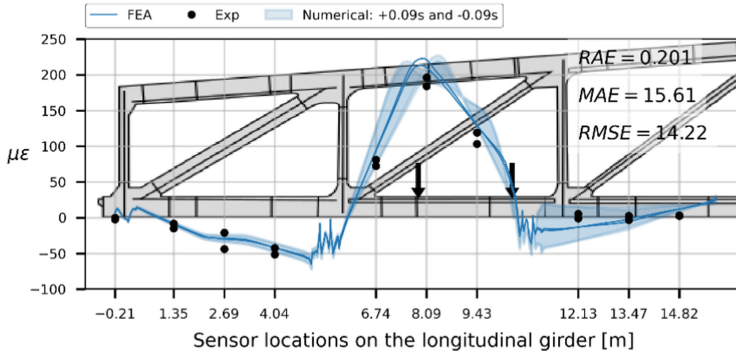


Fig. 13. Comparison between numerical and experimental longitudinal strains for the illustrated axle positions (arrows).

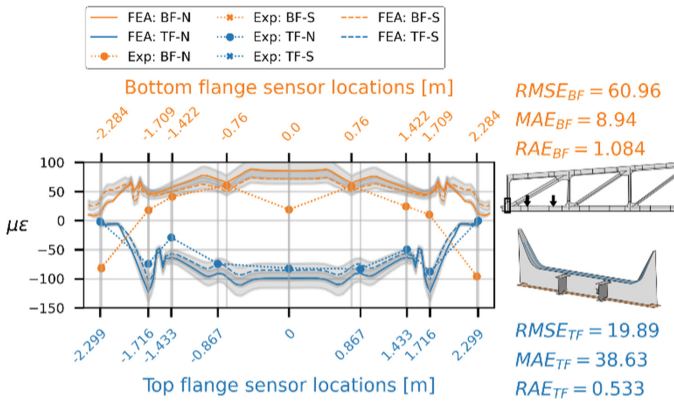


Fig. 14. Numerical and experimental results for the first transverse girder with the initial boundary conditions. Axle locations (black arrows) shown in the top right.

steel brackets (e) to ensure that the truss is positioned correctly on the bearings after closing. No structural drawings were available of this part, but it can be seen that the top of the coupling piece (c) is a rather thick steel plate. This plate seems to be both bolted and welded to the bottom flange. Furthermore, the steel plate at the top of the coupling piece is thicker than the flange (b) it is attached to. This will undoubtedly introduce considerable local stiffness, thereby explaining the lower strain at the center of the girder. As it does not affect the global structural behavior and no accurate information on the dimensions of this structural element was made available, it was not included in the model. It should be noted that no inspection was possible at the time due to practical constraints, therefore this conclusion is solely based on the Fig. 16.

Figure 17 shows the strain distribution of the transverse girder when the front bogie of the train is in the center of the second instrumented sector. For the top flange the agreement between the experimental and numerical results is a lot worse than when the train was positioned closer to the girder (i.e. Fig. 17) as reflected by the RAE. For the

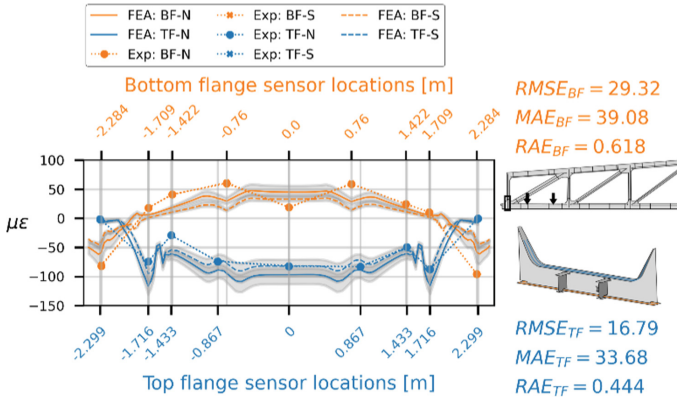


Fig. 15. Numerical and experimental results for the first transverse girder with the final boundary conditions. Axle locations (black arrows) shown in the top right.

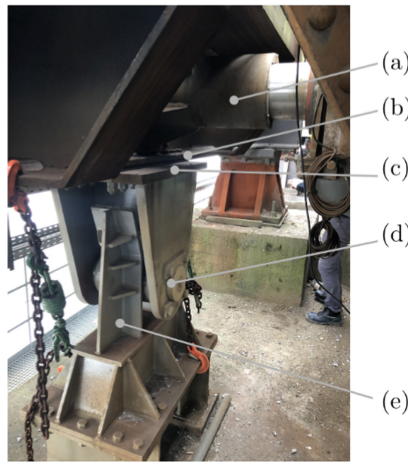


Fig. 16. Bridge guidance support structure

bottom flange, a large discrepancy between the experimental and part of the numerical results can be observed, with an RAE greater than one. According to the experimental measurements, both sides of the bottom flange are completely loaded in tension with almost identical strain values. I.e. the bottom flange of the girder is symmetrically loaded with respect to the girder web. The bottom flange in the model, however, exhibits a totally different behavior. The northern sensor line on the bottom flange is in tension at the girder sides and in compression in the center and vice versa for the southern sensor line. This opposing behavior in both sensor lines that lie on either side of the girder web indicates an out-of-plane bending deformation that is not present in reality.

At the northern side of the transverse girder, in the middle of the web, it can be seen that a large cylinder is attached to the web (Fig. 16a). Based on the photographs it seems to also exert pressure on the next bridge segment. No constructional drawings

nor detailed photographs are available of this part and it was therefore not included in the model. It is however without doubt that this component significantly changes the out-of-plane behavior of the transverse girder, explaining the discrepancy between the numerical and experimental results. It is important to note that this loading case results in overall small strain values. Besides, as this discrepancy does not affect the maximum principal strain range occurring in the flanges of the transverse girder, it will only have a marginal effect on the outcome of a fatigue assessment. Hence, a more detailed assessment of this discrepancy was considered outside the scope of this work. It can also be seen that experimental strain measurements for both sides of the flange are nearly identical, which is not the case for the numerical results. Since the difference in the numerical results between the flange sides is limited and these results are most likely heavily influenced by the absence of the cylinder and fork support in the model, this will not be discussed further.

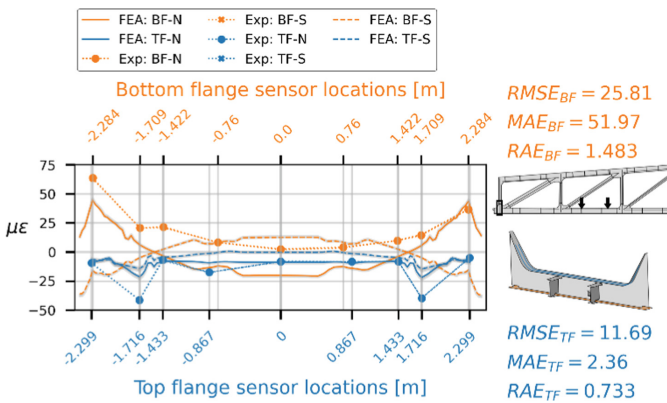


Fig. 17. Numerical and experimental results for the first transverse girder for axle locations (black arrows) shown in the top right.

When the train moves further on the bridge and the axle loads are located at a large distance away from the first transverse girder, the same observations can be made as for the two discussed cases, only with a further reduced magnitude of the strain.

5.4 Transverse Girder 2

Since the second transverse girder is located further away from the bridge bearings, it can be expected that the implementation of the boundary conditions and the causes for the discrepancies discussed in the previous subsection, will not affect the strain measurements on this second girder. Figures 18 and 19 show the strain distribution for the transverse girder when the front bogie of the train is at the center of the first and second instrumented sector, respectively. In both figures, an excellent agreement between the numerical and experimental results is found regarding both the shape and magnitude for every position of the train on the bridge in both the bottom and top flanges. This is reflected by the $RAE < 0.3$ and low $RMSE (< 10\%$ of the $|\mu\epsilon_{max}|$) for both top and bottom flanges.

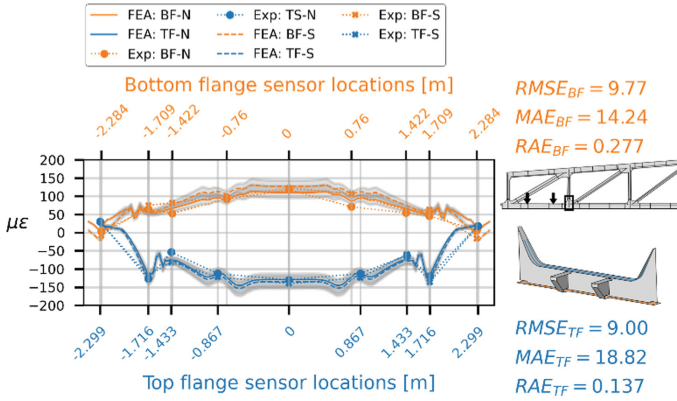


Fig. 18. Numerical and experimental results for the second transverse girder for the axle locations (black arrows) shown in the top right.

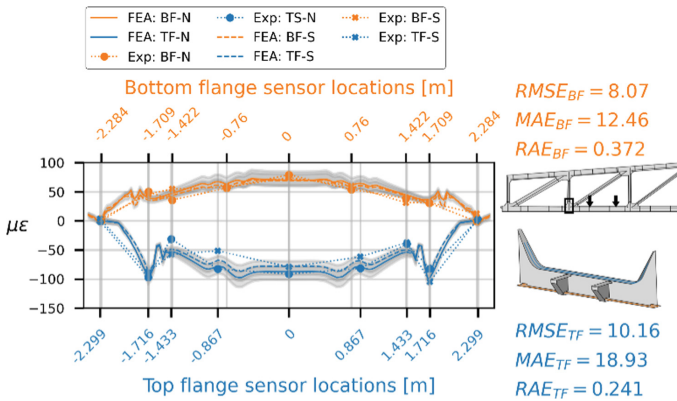


Fig. 19. Numerical and experimental results for the second transverse girder for the axle locations (black arrows) shown in the top right.

6 Discussion

Considering the comparisons between the numerical and experimental results for the longitudinal and transverse girders, an overall good agreement is found. One of the main observations is the discrepancy between the numerical and experimental results near the boundary condition at the north side of the bridge. Although the shape and magnitude of the numerical results in the northern transverse girder agrees well with the experimental results when the bogie is close to the considered girder, the agreement worsens when the bogie moves into a neighboring sector. This observation is however limited to the results of the bottom flange. The numerical results of the top flange are much better for all load scenarios. As discussed in the previous sections, the northern transverse girder is connected to two of the main bridge bearings, but also to two structural elements from which no technical drawings are available. A remarkable observation

that could be drawn from the combination of numerical and experimental results is the significant difference between the hypothesized (i.e., based on an interpretation of the design drawings) effect and the real effect of the bridge bearing on the structural behavior. Based on the experimental results it was concluded that the supports also restrict translations in the transverse directions. Even to the extent that, numerically, a completely restricted translation DOF in the transverse direction produced better results than a model that comprises the hypothesized boundary conditions (i.e., simply supported beam with free transverse movement). Evidently, this shows the importance of field measurements for a precise understanding of the mechanical behavior of all structural components. Furthermore, it also shows that the real interaction between the structural members and the bearings should be modeled if the local response of the structural elements near the bridge bearings is to be analyzed. Including the bearings explicitly would however significantly increase the complexity of the finite element model.

Irrespective of the conclusions in the previous paragraph, the results did also show that the difference between the real and hypothesized boundary conditions does not affect the overall structural response. Where restricting an extra translational DOF at the northern transverse girder had a major effect on the local strains for that structural element, no noticeable difference was observed in the neighboring girder. This could also, at least partly, be due to the bridge design. Both the numerical and experimental results showed that deformation of the structure is strongly constrained to the loaded sectors. Once a bogie passes to a neighboring sector, the strains in the transverse girders that do not border the loaded sector drop significantly. Thus, the truss design efficiently transfers the loads on the longitudinal girders through the closest transverse girders to the main trusses. Finally, it can be concluded that modeling the bridge bearings by constraining individual nodes is adequate if the goal is to model global behavior of the bridge. The finite element model can thus be considered validated for this purpose.

Both longitudinal girders were instrumented in a symmetrical manner with respect to the plane equidistant from and parallel to both girder webs. From weighing the train it is known that there is slight difference between left and right wheel loads. The maximum load difference between both sides of a single bogie is 3.5%. In the experimental results however, differences up to 20% were observed between the two longitudinal girders, which were not observed in the numerical results. Although the transverse wheel load difference will have an influence on the asymmetry in the measurement results to some extent, this is not the sole cause. The asymmetric measurement results observed in the transverse direction are hypothesized to be primarily caused by the influence of the pedestrian walkway, which was not included in the model. This is supported by the fact that the longitudinal girder with the lowest measured strains is located on the same side as the walkway. Secondary factors that could contribute to the observed asymmetry are enumerated hereafter. (1) The train wheels are never exactly at the center of the rails, thus locally introducing a small secondary moment that affects the stress in the longitudinal girder. (2) Construction tolerances were considerably larger at the time the bridge was constructed compared to contemporary bridges. (3) Over its lifetime the bridge has been locally repaired and repainted over five times because corrosion damage occurred. Therefore the design thickness of some structural members likely does not match the current state of the bridge. (4) The opening and closing of the bridge over the years

has likely introduced geometrical imperfections in the longitudinal girders which has resulted in a permanent asymmetry.

7 Conclusions

- The numerical strain values for the longitudinal girders show an excellent agreement with the experimental results, except for the sensors located closest to the bridge bearings.
- Deformation of the structure is strongly constrained to the loaded sectors. The truss design efficiently transfers the loads acting on the longitudinal girders through the closest transverse girders onto the main trusses.
- There is a significant difference between the hypothesized and the real implementation of the bridge bearings. Whereas an unconstrained movement in the transverse direction was expected, the real bridge bearing does not allow for this. The load transferred through the transverse girder does not overcome the frictional resistance between the top and bottom half of the support bearing, therefore acting as a rigid connection. This results in considerable strains in the bottom flange of the girder near the bearings.
- For the transverse girder that is directly connected to the bridge bearings, a good agreement was found between the numerical and experimental results for the top flange. For the bottom flange a good agreement was found when the bridge is loaded close to the considered transverse girder. When the load is further away this agreement worsens. Discrepancies were linked to photographic evidence of locally attached components (that were not identified on the design drawings) that result in local stiffening and restraint of displacements. Hence model validation of a model intended for fatigue life analysis should be based on full strain trace data and not only consider the time steps at maximum load.
- For the second transverse girder, the numerical and experimental results show an excellent agreement in both shape and magnitude for every position of the train on the bridge. Based on the observation that significant deformations are mainly constrained within the loaded sectors and the negligible influence of the difference between the real and theoretical boundary conditions on the global structural behavior, this conclusion can be extended to the other transverse girders that are not located close to the boundary conditions.
- Asymmetric measurement results were observed in the transverse direction that could not only be assigned to the transverse load difference, possibly indicating that the real structural behavior is asymmetric. This was not observed in the numerical model.

The results show that combination of a full-scale finite element model with a well thought out optical fiber sensor layout provides considerable insight in the real structural behavior of a complex bridge. The installed monitoring system will remain in operation to aggregate data concerning the crossings of different trains during daily operation (speed, direction, number of axles, axle loads and frequency). This data will be used in the future to construct representative long term load histograms suitable for fatigue analysis.

References



1. R. Haghani, PANTURA D5.3 Needs for maintenance and refurbishment of bridges in urban environments, 2013.
2. I. Olofsson, L. Elfgrén, B. Bell, B. Paulsson, E. Niederleithinger, J. Sandager Jensen, G. Feltrin, B. Täljsten, C. Cremona, R. Kiviluoma, J. Bien, Assessment of European railway bridges for future traffic demands and longer lives – EC project “Sustainable Bridges,” Structure and Infrastructure Engineering. 1 (2005) 93–100. <https://doi.org/10.1080/15732470412331289396>.
3. R.J. Woodward, D.W. Cullington, A.F. Daly, P.R. Vassie, P. Haardt, R. Kashner, R. Astudillo, C. Velando, B. Godart, C. Cremona, B. Mahut, A. Raharinaivo, I. Markey, L. Bevc, I. Perus, BRIME Final Report, Europe, 2001.
4. B. Kühn, A. Nussbaumer, R. Helmerich, S. Herion, M.H. Kolstein, S. Walbridge, B. Androic, O. Dijkstra, Ö. Bucak, Assessment of existing steel structures. A guideline for estimation of the remaining fatigue life, 2008. <https://doi.org/10.1080/15732470500365562>.
5. A. Dinas, Th.N. Nikolaidis, C.C. Baniotopoulos, Sustainable Restoration Criteria for a Historical Steel Railway Bridge, Procedia Environ Sci. 38 (2017) 578–585. <https://doi.org/10.1016/j.proenv.2017.03.131>.
6. Y.E. Zhou, Assessment of bridge remaining fatigue life through field strain measurement, Journal of Bridge Engineering. 11 (2006) 129–139. [https://doi.org/10.1061/\(asce\)1084-0702\(2006\)11:6\(737\)](https://doi.org/10.1061/(asce)1084-0702(2006)11:6(737)).
7. C.C. Spyrakos, I.G. Raftoyiannis, J.C. Ermopoulos, Condition assessment and retrofit of a historic steel-truss railway bridge, J Constr Steel Res. 60 (2004) 1213–1225. <https://doi.org/10.1016/j.jcsr.2003.11.004>.
8. J. Ermopoulos, C.C. Spyrakos, Validated analysis and strengthening of a 19th century railway bridge, Eng Struct. 28 (2006) 783–792. <https://doi.org/10.1016/j.engstruct.2005.10.006>.
9. R. Malm, A. Andersson, Field testing and simulation of dynamic properties of a tied arch railway bridge, Eng Struct. 28 (2006) 143–152. <https://doi.org/10.1016/j.engstruct.2005.07.011>.
10. B.O. Caglayan, K. Ozakgul, O. Tezer, Fatigue life evaluation of a through-girder steel railway bridge, Eng Fail Anal. 16 (2009) 765–774. <https://doi.org/10.1016/j.engfailanal.2008.06.018>.
11. J. Leander, A. Andersson, R. Karoumi, Monitoring and enhanced fatigue evaluation of a steel railway bridge, Eng Struct. 32 (2010) 854–863. <https://doi.org/10.1016/j.engstruct.2009.12.011>.
12. A. Fatemi, L. Yang, Cumulative fatigue damage and life prediction theories: a survey of the state of the art for homogeneous materials, Int J Fatigue. 20 (1998) 9–34. [https://doi.org/10.1016/S0142-1123\(97\)00081-9](https://doi.org/10.1016/S0142-1123(97)00081-9).
13. K. Hectors, W. de Waele, Cumulative damage and life prediction models for high-cycle fatigue of metals: A review, Metals (Basel). 11 (2021) 1–32. <https://doi.org/10.3390/met11020204>.
14. Q.H. Zhang, Y.Z. Bu, Q. Li, Review on fatigue problems of orthotropic steel bridge deck, Zhongguo Gonglu Xuebao/China Journal of Highway and Transport. 30 (2017).
15. M. Aygül, M. Bokesjö, M. Heshmati, M. Al-Emrani, A comparative study of different fatigue failure assessments of welded bridge details, Int J Fatigue. 49 (2013) 62–72. <https://doi.org/10.1016/j.ijfatigue.2012.12.010>.
16. G. Alencar, A. de Jesus, J.G.S. da Silva, R. Calçada, Fatigue cracking of welded railway bridges: A review, Eng Fail Anal. 104 (2019) 154–176. <https://doi.org/10.1016/j.engfailanal.2019.05.037>.
17. X.W. Ye, Y.H. Su, J.P. Han, A state-of-the-art review on fatigue life assessment of steel bridges, Math Probl Eng. 2014 (2014). <https://doi.org/10.1155/2014/956473>.

18. A. Hobbacher, Recommendations for fatigue design of welded joints and components, 2nd edition, Springer, 2016.
19. Z.X. Li, T.Q. Zhou, T.H.T. Chan, Y. Yu, Multi-scale numerical analysis on dynamic response and local damage in long-span bridges, *Eng Struct.* 29 (2007) 1507–1524. <https://doi.org/10.1016/j.engstruct.2006.08.004>.
20. H. Wang, A. Li, T. Guo, S. Ma, Accurate stress analysis on rigid central buckle of long-span suspension bridges based on submodel method, *Science in China, Series E: Technological Sciences.* 52 (2009) 1019–1026. <https://doi.org/10.1007/s11431-009-0070-z>.
21. H. Wang, A. Li, R. Hu, J. Li, Accurate stress analysis on steel box girder of long span suspension bridges based on multi-scale submodeling method, *Advances in Structural Engineering.* 13 (2010) 727–740. <https://doi.org/10.1260/1369-4332.13.4.727>.
22. M. Mashayekhi, E. Santini-Bell, Fatigue assessment of a complex welded steel bridge connection utilizing a three-dimensional multi-scale finite element model and hotspot stress method, *Eng Struct.* 214 (2020) 110624. <https://doi.org/10.1016/j.engstruct.2020.110624>.
23. D. Ribeiro, R. Calçada, R. Delgado, M. Brehm, V. Zabel, Finite element model updating of a bowstring-arch railway bridge based on experimental modal parameters, *Eng Struct.* 40 (2012) 413–435. <https://doi.org/10.1016/j.engstruct.2012.03.013>.
24. A. Meixedo, D. Ribeiro, J. Santos, R. Calçada, M. Todd, Progressive numerical model validation of a bowstring-arch railway bridge based on a structural health monitoring system, *J Civ Struct Health Monit.* 11 (2021) 421–449. <https://doi.org/10.1007/s13349-020-00461-w>.
25. M.I. Friswell, J.E. Mottershead, Finite Element Model Updating in Structural Dynamics, Springer Netherlands, Dordrecht, 1995. <https://doi.org/10.1007/978-94-015-8508-8>.
26. R. Maaskant, T. Alavie, R.M. Measures, G. Tadros, S.H. Rizkalla, A. Guha-Thakurta, Fiber-optic Bragg grating sensors for bridge monitoring, *Cem Concr Compos.* 19 (1997) 21–33. [https://doi.org/10.1016/s0958-9465\(96\)00040-6](https://doi.org/10.1016/s0958-9465(96)00040-6).
27. D. Inaudi, Application of fiber optic sensors to structural monitoring, *Trends in Optical Non-Destructive Testing and Inspection.* 4763 (2000) 459–472. <https://doi.org/10.1016/b978-008043020-1/50031-4>.
28. W. de Waele, J. Degrieck, P. de Baets, W. Moerman, L. Taerwe, Feasibility of integrated optical fibre sensors for monitoring of composite structures, *Insight.* 45 (2003).
29. T.H.T. Chan, L. Yu, H.Y. Tam, Y.Q. Ni, S.Y. Liu, W.H. Chung, L.K. Cheng, Fiber Bragg grating sensors for structural health monitoring of Tsing Ma bridge: Background and experimental observation, *Eng Struct.* 28 (2006) 648–659. <https://doi.org/10.1016/j.engstruct.2005.09.018>.
30. C. Barbosa, N. Costa, L.A. Ferreira, F.M. Arajo, H. Varum, A. Costa, C. Fernandes, H. Rodrigues, Weldable fibre Bragg grating sensors for steel bridge monitoring, *Meas Sci Technol.* 19 (2008). <https://doi.org/10.1088/0957-0233/19/12/125305>.
31. F. Matta, F. Bastianini, N. Galati, P. Casadei, A. Nanni, Distributed Strain Measurement in Steel Bridge with Fiber Optic Sensors: Validation through Diagnostic Load Test, *Journal of Performance of Constructed Facilities.* 22 (2008) 264–273. [https://doi.org/10.1061/\(asce\)0887-3828\(2008\)22:4\(264\)](https://doi.org/10.1061/(asce)0887-3828(2008)22:4(264)).
32. M. Enckell, B. Glisic, F. Myrvoll, B. Bergstrand, Evaluation of a large-scale bridge strain, temperature and crack monitoring with distributed fibre optic sensors, *J Civ Struct Health Monit.* 1 (2011) 37–46. <https://doi.org/10.1007/s13349-011-0004-x>.
33. B.J.A. Costa, J.A. Figueiras, Evaluation of a strain monitoring system for existing steel railway bridges, *J Constr Steel Res.* 72 (2012) 179–191. <https://doi.org/10.1016/j.jcsr.2011.12.006>.
34. J. Xu, Y. Dong, Z. Zhang, S. Li, S. He, H. Li, Full scale strain monitoring of a suspension bridge using high performance distributed fiber optic sensors, *Meas Sci Technol.* 27 (2016). <https://doi.org/10.1088/0957-0233/27/12/124017>.
35. K. van der Kooi, N.A. Hoult, H. Le, Monitoring an In-Service Railway Bridge with a Distributed Fiber Optic Strain Sensing System, *Journal of Bridge Engineering.* 23 (2018) 05018007. [https://doi.org/10.1061/\(asce\)be.1943-5592.0001281](https://doi.org/10.1061/(asce)be.1943-5592.0001281).

36. G. Kouroussis, C. Caucheteur, D. Kinet, G. Alexandrou, O. Verlinden, V. Moeyaert, Review of trackside monitoring solutions: From strain gages to optical fibre sensors, *Sensors (Switzerland)*. 15 (2015) 20115–20139. <https://doi.org/10.3390/s150820115>.
37. C. Du, S. Dutta, P. Kurup, T. Yu, X. Wang, A review of railway infrastructure monitoring using fiber optic sensors, *Sens Actuators A Phys.* 303 (2020) 111728. <https://doi.org/10.1016/j.sna.2019.111728>.



Comparative Study on Frequency-Based Cable Parameter Identification Algorithms

Ceshi Sun¹ , Gang Zheng² , Jigang Yu¹, Peng Chen¹, and Chuan Zeng¹

¹ College of Civil Engineering, Chongqing Jiaotong University, Chongqing, People's Republic of China

² State Key Laboratory of Mountain Bridge and Tunnel Engineering, Chongqing Jiaotong University, Chongqing, People's Republic of China
zhenggang@cqjtu.edu.cn

Abstract. Frequency-based algorithms are prevalently used to estimate bridge cable tensions. However, for the non-negligible bending stiffness, the eigenproblem involves a cumbersome procedure of solving transcendental equations, which aroused various treatment skills for simplification or optimization, yielding diverse identification algorithms with different effectiveness, efficiency, and applicability that bring difficulties and confusion in making an appropriate choice for users. Therefore, it is necessary to systematically illustrate these algorithms' intrinsic relations, differences, and application characteristics. A comprehensive comparative study on five representative algorithms, including analytical and empirical formulas and optimal algorithms, is carried out in this paper by parallelly identifying cable parameters on 24 real-life bridge cables worldwide available in the literature. Results show that: (i) Algorithms of empirical formulas are always based on the modified chord theory or tensioned-simply-supported beam model, resulting in the solution of two simultaneous equations by providing two measured frequencies in case of unknown bending stiffness. (ii) Proper identification procedure is significant in applying an algorithm. (iii) The FROCPI algorithm is the simplest method with good consistency and broad applicability for long and short cables.

Keywords: Bridge cable · Frequency-based algorithm · Cable tension identification · Bending stiffness identification

1 Introduction

Measuring and monitoring cable tension is a crucial aspect of engineering structures during construction and operation [1, 2]. The frequency-based method is the most widely used technique for cable tension testing [3, 4], with the chord theory model being the most fundamental [5]. However, the measurement accuracy of the chord theory model is not high for short cables with non-ideal hinges, leading researchers to develop cable tension testing algorithms that consider bending stiffness for several decades. Unfortunately, this approach presents another challenging problem as the frequency equation is a transcendental equation that is difficult to solve and not convenient for engineering applications [6].

To address this issue, world-wide scholars have proposed various cable tension identification algorithms, including empirical formulas based on approximation and optimization methods. For example, in 1996, Zui [7] proposed the classic segmented function empirical formula. In 1998, Tabatabai [8] presented an empirical formula based on a frequency surface obtained by solving transcendental equations in advance. Ren [9] and Fang [10] provided approximate calculation formulas based on a revised tensioned simply supported beam model. Ceballos and Prato [11] proposed an approximate algorithm based on an elastic support tensile simply supported beam model. Huang [12] identified bending stiffness using a fitting formula for the frequency curve and simultaneously solving two equations corresponding to measured frequencies. Yu [13] proposed an approximate algorithm that considers end elastic constraints to identify both bending stiffness and cable tension. In a recent research, Sun et al. proposed a new decoupling algorithm based on frequency-ratio-offset: FROCPI [14]. The measurement accuracy, efficiency, and applicability of these algorithms vary, resulting in the need to examine their suitability for practical applications.

This paper analyzes the testing accuracy, efficiency, and applicability of five representative algorithms used to simultaneously identify cable tension and bending stiffness for 24 inclined cables of real-life bridges discussed in the literature. The inherent connections behind these algorithms are examined and the reasons for the superiority or inferiority of their testing accuracy are analyzed.

2 Selected Representative Identification Algorithms

This paper conducted a comparative study on five algorithms, namely the method proposed by Tabatabai [8], Fang [10], Huang [12], Yu [13], and the FROCPI [14], for cable force identification. The equations and application description of these algorithms are presented in Table 1, with some of them simplified to neglecting sag effect and elastic boundary conditions. It should be emphasized that this study aims to simultaneously identify cable tension and bending stiffness utilizing only two measured frequencies, rather than solely focusing on cable tension identification. Consequently, some algorithms require solving simultaneous equations. Moreover, in consideration of algorithmic practicality and simplicity, only those that can be easily solved by hand were chosen, as opposed to those necessitating iteration or optimization with a computer. For the convenience of later description, the five algorithms are briefly denoted by A_1 , A_2 , A_3 , A_4 and A_5 in sequence.

3 Comparative Study

3.1 Selected Cable Data from Literature

Given that this comparative study evaluates the identification efficiency of cable tension and bending stiffness, particular emphasis was placed on selecting real-life bridge cable data from references that simultaneously identify these two parameters. Furthermore, since the algorithms presented in Table 1 required two measured frequencies to carry out the identification, the cable data listed in the references must ensure the measurement of

Table 1. Five cable tension identification algorithms.

Algorithm	Equation for identification	Description
Tabatabai (A ₁)	$\frac{\omega_n}{\omega_{ns}} = 1 + \frac{2}{\xi} + \frac{\left(4 + \frac{n^2\pi^2}{2}\right)}{\xi^2}$ <p>where $\omega_{ns} = \frac{n\pi}{L} \sqrt{\frac{T}{m}}$, $\xi = L\sqrt{\frac{T}{EI}}$</p>	<p>If EI is unknown, two measured frequencies must be used to solve simultaneous equations for T and EI</p>
Fang (A ₂)	$T = 4\pi^2 ml^2 \frac{f_n^2}{\gamma_n^2} - \frac{EI}{l^2} \gamma_n^2$ <p>where $\gamma_n = n\pi + A\psi_n + B\psi_n^2$,</p> $\psi_n = \frac{1}{\chi_n \gamma_n} = \sqrt{\frac{EI}{m\omega_n^2 l^4}}$ $A = -18.9 + 26.2n + 15.1n^2,$ $B = \begin{cases} 290.0 & (n = 1) \\ 0 & (n \geq 1) \end{cases}$	<p>If EI is unknown, two measured frequencies must be used to solve simultaneous equations for T and EI</p>
Huang (A ₃)	$T = K_n ml^2 \omega_n^2 / (n\pi)^2$ $\sqrt{EI} =$ $\frac{-(B_i \phi_i - r B_j \phi_j) + \sqrt{(B_i \phi_i - r B_j \phi_j)^2 - 4(A_i \phi_i^2 - r A_j \phi_j^2)(r - 1)}}{2(A_i \phi_i^2 - r A_j \phi_j^2)}$ <p>where, $K_n = -A_n \lambda_n^2 - B_n \lambda + 1$,</p> $\begin{cases} A_n = 98.2n^4 + 87.64n^3 + 65.37n^2 \\ B_n = 9.31n + 1.72 \end{cases} \quad n =$ <p>1, 2, ..., ∞,</p> $\lambda = \frac{1}{\alpha_n \beta_n} = \sqrt{\frac{EI}{m\omega_n^2 l^4}}, \frac{\omega_j^2}{j^2} / \frac{\omega_i^2}{i^2} = r,$ $\phi_n = \frac{1}{\sqrt{m\omega_n l^2}}$	<p>The calculation formulas for T and EI have been obtained through the simultaneous equations of two measured frequencies, but the formulas are rather complicated</p>

(continued)

Table 1. (continued)

Algorithm	Equation for identification	Description
Yu (A ₄)	$T = 4(\rho A)l^2 \left[\frac{p^4 f_n^2 - n^4 f_p^2}{n^2 p^2 (p^2 - n^2)} \right] \times (\lambda_0^2)$ $\sqrt{EI / Tl^2} \equiv \delta$ <p>where, $\lambda_0 \equiv (1 - 2\delta)$,</p> $\delta = \frac{1}{\pi} \sqrt{\frac{p^2 f_n^2 - n^2 f_p^2}{n^4 f_n^2 - p^4 f_p^2}}$	Equation of T has been obtained using two measured frequencies, and the relation between T and EI , denoted as δ , has also been derived. Then, by substituting the obtained T into δ , EI can be determined
FROCPI (A ₅)	$\tilde{T} = (\tilde{m}\alpha_x^2 / \alpha_t^2)$ $\tilde{EI} = (\tilde{m}\alpha_x^4 / \alpha_t^2)$ <p>where, $\alpha_x = \tilde{l} / l, \alpha_t = \omega_i / \tilde{\omega}_i$</p>	The decoupled formulas for T and EI have been provided

at least two frequencies. Of course, some references provide more than two measured frequencies, which involves the problem of frequency order selection [14]. This paper does not consider the optimal or worst selection of frequency order for the time being, but selects frequencies uniformly according to the following two principles: (i) Since in using the FROCPI algorithm, the application precondition that $\gamma_{ij} > 0$ must satisfy, the frequency pairs that violating this condition are not selected in this paper; (ii) Among the measured frequencies provided in the literature, one relatively high-order frequency and one relatively low-order one are selected as much as possible. The cable parameters and measured frequencies collected from various references are shown in Table 2.

3.2 Identified Results and Comparison

Cable Tension

The identified cable tensions and errors corresponding to reference values are listed in Table 3. Overall, the identification results of the five algorithms are quite consistent, indicating good consistency. However, the identified cable tension errors by Yu and Tabatabai’s algorithms for Cable 24 are respectively 78.2% and 19.6%, which is much higher than the other three algorithms. This is mainly because the length of this cable is too short, only 2.56 m. In order to compare the overall identification performance of these five algorithms, the root mean square (RMS) of the cable tension identification error for these 24 cables was calculated, which are 6.21, 5.25, 4.81, 16.32 and 5.01 respectively. It should be noted that the errors of the cable tensions of the Hwamyung Bridge (Cables 1–8) and Yongjiang Bridge (Cables 9–13) were either all positive or all negative for the five algorithms, indicating systematic rather than random errors, meaning that there may be errors of parameters in these two references.

Table 2. Cable parameters from various literature.

Number		f_i (Hz)	f_j (Hz)	l (m)	T (kN)	EI (kNm ²)	m (kg/m)
Hwamyung Cable-Stayed Bridge [15]	1	6.024(3)	10.095(5)	55.917	3245.000	/	66.189
	2	7.263(3)	12.354(5)	55.917	4854.000	/	66.189
	3	7.239(3)	9.717(4)	55.917	4791.000	/	66.189
	4	7.330(3)	9.857(4)	55.917	4859.000	/	66.189
	5	7.605(3)	13.001(5)	45.568	3515.000	/	66.189
	6	2.557(1)	15.381(6)	45.568	3186.000	/	66.189
	7	8.881(3)	15.040(5)	45.568	4693.000	/	66.189
	8	9.021(3)	15.064(5)	45.568	4620.000	/	66.189
Yongjiang Bridge [16]	9	2.509(2)	8.82(7)	102.436	2706.263	1227.600	40.846
	10	6.058(4)	9.116(6)	84.576	2139.452	747.240	32.522
	11	5.85(3)	9.799(5)	66.731	2201.069	747.240	32.522
	12	8.55(3)	20.493(7)	48.877	2497.620	747.240	32.522
	13	19.317(4)	24.433(5)	31.011	2803.083	747.240	32.522
Qijiang Bridge [17]	14	11.332(5)	15.913(7)	61.928	2330.000	235.008	30.400
	15	7.817(7)	11.225(10)	115.505	4700.000	1314.213	72.600
	16	4.284(5)	8.609(10)	156.520	6055.000	1931.215	85.900
Yamen Bridge [18]	17	3.1471(1)	19.1922(6)	40.850	2223.000	340.200	34.900
	18	2.3968(2)	14.579(12)	104.230	2948.000	598.200	48.000
	19	2.3391(3)	11.796(15)	165.280	3873.000	837.300	58.900
	20	4.9418(5)	13.3029(8)	74.500	2265.000	414.300	38.6000
Hedong & Sanshan west bridges [12]	21	15.2224(6)	20.4313(8)	47.663	2000.000	292.500	35.400
	22	0.6615(1)	6.6377(10)	184.144	4050.000	1047.150	68.400
	23	15.6797(5)	33.4922(10)	30.323	825.000	147.570	24.556
	24	60.0586(1)	140.332(2)	2.560	1000.000	147.570	24.556

Note The number in parentheses is the frequency order. The bold font of T and EI are identified values from references, while others are reference values given in the corresponding reference

Bending Stiffness

The identified bending stiffnesses and errors corresponding to reference values are listed in Table 4. Cables 1–8 have no error data due to the lack of a reference for bending stiffness. Compared with the cable tension identification results shown in Table 3, the discreteness of the bending stiffness identified by the five algorithms is much greater. In particular, the identified bending stiffnesses of A_2 are significantly larger, and that of A_3 are significantly smaller than the other three algorithms, while the results of A_1 , A_4 and A_5 are in good agreement. Similarly, for Hwamyung Bridge and Yongjiang Bridge, the identification values are either all positive or all negative. The root mean square (RMS)

Table 3. Identified cable tensions and errors.

Number		Results									
		Identified cable tension (kN)					Error (%)				
		A ₁	A ₂	A ₃	A ₄	A ₅	A ₁	A ₂	A ₃	A ₄	A ₅
Hwamyung Cable-Stayed Bridge [15]	1	3207	3175	3232	3289	3207	-1.2	-2.2	-0.4	1.4	-1.2
	2	4433	4384	4479	4660	4433	-8.7	-9.7	-7.7	-4.0	-8.7
	3	4471	4411	4524	4669	4473	-6.7	-7.9	-5.6	-2.5	-6.6
	4	4528	4467	4583	4755	4530	-6.8	-8.1	-5.7	-2.1	-6.8
	5	3182	3146	3214	3365	3181	-9.5	-10.5	-8.6	-4.3	-9.5
	6	3539	3511	3555	3580	3539	11.1	10.2	11.6	12.4	11.1
	7	4460	4412	4506	4661	4461	-5.0	-6.0	-4.0	-0.7	-5.0
	8	4862	4817	4888	4935	4862	5.2	4.3	5.8	6.8	5.2
Yongjiang Bridge [16]	9	2648	2624	2660	2684	2648	-2.2	-3.0	-1.7	-0.8	-2.1
	10	2075	2065	2087	2111	2075	-3.0	-3.5	-2.5	-1.3	-3.0
	11	2120	2100	2137	2173	2121	-3.7	-4.6	-2.9	-1.3	-3.6
	12	2376	2358	2396	2463	2376	-4.9	-5.6	-4.1	-1.4	-4.9
	13	2611	2592	2633	2747	2610	-6.8	-7.5	-6.1	-2.0	-6.9
Qijiang Bridge [17]	14	2332	2326	2343	2368	2332	0.1	-0.2	0.6	1.6	0.1
	15	4697	4696	4715	4761	4696	-0.1	-0.1	0.3	1.3	-0.1
	16	6071	6063	6092	6138	6071	0.3	0.1	0.6	1.4	0.3
Yamen Bridge [18]	17	2215	2191	2233	2282	2216	-0.3	-1.4	0.4	2.7	-0.3
	18	2940	2924	2951	2980	2940	-0.3	-0.8	0.1	1.1	-0.3
	19	3865	3857	3875	3899	3866	-0.2	-0.4	0.1	0.7	-0.2
	20	2267	2262	2278	2307	2267	0.1	-0.1	0.6	1.8	0.1
Hedong & Sanshan west bridges [12]	21	1979	1977	1989	2021	1979	-1.1	-1.2	-0.5	1.0	-1.1
	22	4016	4000	4026	4049	4017	-0.8	-1.2	-0.6	0.0	-0.8
	23	802	799	804	835	799	-2.8	-3.1	-2.6	1.2	-3.1
	24	1196	1044	903	1782	1078	19.6	4.4	-9.7	78.2	7.8

of the bending identification error for cables 9–24 was also calculated, which are 18.84, 19.81, 23.07, 18.45 and 18.83 respectively.

3.3 Discussion

The mechanical models of the above algorithms are all a tensioned simply-supported Euler beam with clamped ends. However, as mentioned before, it is inconvenient for

Table 4. Identified bending stiffnesses and errors.

Number		Results									
		Identified bending stiffness (kNm ²)					Error (%)				
		A ₁	A ₂	A ₃	A ₄	A ₅	A ₁	A ₂	A ₃	A ₄	A ₅
Hwamyung Cable-Stayed Bridge [15]	1	710	937	622	720	699	/	/	/	/	/
	2	3778	4235	3551	3927	3716	/	/	/	/	/
	3	2830	3426	2532	2906	2768	/	/	/	/	/
	4	3666	4302	3328	3784	3583	/	/	/	/	/
	5	2268	2499	2164	2375	2235	/	/	/	/	/
	6	109	189	71	110	108	/	/	/	/	/
	7	1963	2249	1817	2027	1929	/	/	/	/	/
	8	249	413	208	251	247	/	/	/	/	/
Yongjiang Bridge [16]	9	554	886	463	558	549	-54.9	-27.8	-62.3	-54.6	-55.3
	10	487	621	437	492	482	-34.9	-16.9	-41.5	-34.2	-35.5
	11	613	820	535	621	603	-18.0	9.8	-28.4	-16.9	-19.2
	12	807	900	764	833	800	8.0	20.5	2.3	11.4	7.0
	13	710	770	685	742	702	-5.0	3.0	-8.3	-0.7	-6.1
Qijiang Bridge [17]	14	233	275	214	235	231	-1.1	17.2	-8.9	-0.1	-1.9
	15	1308	1407	1250	1324	1303	-0.5	7.0	-4.9	0.8	-0.9
	16	1940	2240	1810	1955	1930	0.5	16.0	-6.3	1.2	-0.1
Yamen Bridge [18]	17	358	444	290	365	353	5.3	30.4	-14.9	7.2	3.6
	18	643	779	592	651	640	7.5	30.1	-1.0	8.8	7.1
	19	856	1003	802	863	854	2.3	19.7	-4.2	3.1	2.0
	20	426	480	399	432	423	2.9	15.8	-3.6	4.2	2.1
Hedong & Sanshan west bridges [12]	21	221	238	211	225	220	-24.5	-18.5	-27.8	-23.0	-24.9
	22	962	1410	744	966	957	-8.2	34.6	-29.0	-7.7	-8.6
	23	143	151	147	153	146	-3.4	2.2	-0.1	3.7	-0.8
	24	124	134	154	164	131	-15.9	-9.1	4.4	11.0	-11.3

Note Symbol “/” means no error value was obtained because no reference value is listed in the reference

directly practical applications due to difficult in solving a transcendental frequency equation. Therefore, scholars have adopted different processing or simplification methods to propose different algorithms.

Tabatabai determines the key variables ξ and λ^2 that governing the system, then uses numerical methods to obtain the frequency surface, and finally fits a non-dimensional simple relationship for ratio of the n -th order in-plane frequency and the corresponding

out-of-plane frequency. The algorithm is suitable for cases where $\lambda^2 < 3.1$ and $\xi > 50$. However, to simultaneously identify cable tension and bending stiffness, two measured frequencies must be used to solve two equations simultaneously.

Fang's algorithm is actually based on the revisionist ideology of the tensioned simply-supported beam model. That is using the coefficients γ_n obtained by fitting a quadratic parabola with the parameter Ψ_n to replace $n\pi$ in the frequency equation of the tensioned simply-supported beam system. Likewise, simultaneous equation solving is required to additionally identify the bending stiffness using this method.

Huang's algorithm is based on the revisionist ideology of the string theory model, where a correction coefficient K_n is multiplied onto the frequency equation of the string as the cable tension, which needs to be determined based on the frequency equation under different boundary conditions. In the case of clamped-clamped boundary conditions, a quadratic parabolic fit is also used to determine K_n . Similarly, Huang has also provided the formula for calculating bending stiffness through simultaneous equation solving.

Yu's algorithm involves using two simultaneous frequency equations from the tensioned simply-supported beam model to eliminate the term containing the bending stiffness, and to obtain the relationship between cable tension and the two frequencies and their orders. The concept of effective vibration length is then introduced to correct the expression of cable tension, taking into account the boundary conditions at both ends.

It can be seen that the above four algorithms all use curve fitting methods based on the modified string or tensioned simply-supported beam models to obtain empirical formulas for cable tension. Due to the coupling between bending stiffness and cable tension, in situations where the actual bending stiffness is unknown (often difficult to determine in practical engineering), it is necessary to establish two equations simultaneously to obtain more accurate cable tension identification values. On the contrary, the FROCIPI algorithm does not adopt any curve fitting or modification ideology, but instead carries out interpolation calculation based on pre-solution data of the frequency transcendental equation, and then uses the decoupling algorithm to calculate cable tension and bending stiffness separately, making it not an empirical or simplified formula. As all the five algorithms share the same mechanical model, i.e., the tensioned simply-supported Euler beam with clamped ends, their difference lay on the specific identification procedure. Thus, proper identification procedure is significant for creating an algorithm.

4 Conclusions

This paper compares the process and results of identifying the cable tensions and bending stiffness of real-life bridge cables using five different algorithms, and the following main conclusions are drawn:

- (1) The first four algorithms are all based on the modified chord theory or tensioned-simply-supported beam model, and use curve fitting and other methods to obtain approximate empirical formulas for identifying cable forces and bending stiffness within a certain parameter range. The FROCIPI algorithm, however, is based on transcendental equations pre-solved data, rather than any model modification or curve fitting, and therefore cannot be considered as an empirical formula.

- (2) In actual bridge engineering, the cable bending stiffness is often unknown, so at least two measured frequencies are required to simultaneously identify the cable tension and bending stiffness. This leads to the fact that the first four algorithms must solve two simultaneous equations, while the FROCP algorithm is a decoupling algorithm that does not require this process, making it slightly simpler.
- (3) Except for Yu and Tabatabai's algorithms, which have a large identification error for some short cables, the accuracy of the five algorithms in cable tension testing is overall consistent, but Fang and Huang's algorithms have a significantly lower accuracy in identifying bending stiffness compared to the other three methods.


References

1. Amezcua-Sanchez, J.P., Valtierra-Rodriguez, M., Adeli, H.: Wireless smart sensors for monitoring the health condition of civil infrastructure. *Scientia Iranica*. (2018).
2. Amezcua-Sanchez, J.P., Adeli, H.: Nonlinear measurements for feature extraction in structural health monitoring. *Scientia Iranica*. 26, 3051–3059 (2019).
3. Zhijun Li, Hyo Seon Park, Hojjat Adeli: New method for modal identification of super high-rise building structures using discretized synchrosqueezed wavelet and Hilbert transforms. *The Structural Design of Tall and Special Buildings*. 26, e1312 (2017).
4. Yongding Tian, Cheng Zhang, Shang Jiang, Jian Zhang, Wenhui Duan: Noncontact cable force estimation with unmanned aerial vehicle and computer vision. *Computer-Aided Civil and Infrastructure Engineering*. 2020, 1–16 (2020).
5. Wenju Zhao, Guangwei Zhang, Jian Zhang: Cable force estimation of a long-span cable-stayed bridge with microwave interferometric radar. *Computer-Aided Civil and Infrastructure Engineering*. 35, 1419–1433 (2020).
6. Sébastien Lalonde, Reynald Guilbault, Frédéric Légeron: Modeling multilayered wire strands, a strategy based on 3D finite element beam-to-beam contacts - Part I: Model formulation and validation. *International Journal of Mechanical Sciences*. 126, 281–296 (2017).
7. Hiroshi Zui, Tohru Shinke, Yoshio Namita: Practical formulas for estimation of cable tension by vibration method. *Journal of Structural Engineering*. 122, 651–656 (1996).
8. Armin B. Mehrabi, Habib Tabatabai: Unified finite difference formulation for free vibration of cables. *Journal of Structural Engineering*. 124, 1313–1322 (1998).
9. Wei-Xin Ren, Gang Chen, Wei-Hua Hu: Empirical formulas to estimate cable tension by cable fundamental frequency. *Structural Engineering and Mechanics*. 20, 363–380 (2005).
10. Zhi Fang, Jian-qun Wang: Practical Formula for Cable Tension Estimation by Vibration Method. *Journal of Bridge Engineering*. 17, 161–164 (2012).
11. Marcelo A. Ceballos, Carlos A. Prato: Determination of the axial force on stay cables accounting for their bending stiffness and rotational end restraints by free vibration tests. *Journal of Sound and Vibration*. 317, 127–141 (2008).
12. Yong-Hui Huang, Ji-Yang Fu, Rong-Hui Wang, Quan Gan, Ai-Rong Liu: Unified Practical Formulas for Vibration-Based Method of Cable Tension Estimation. *Advances in Structural Engineering*. 18, 405–422 (2015).
13. Chih-Peng Yu: Tension prediction for straight cables based on effective vibration length with a two-frequency approach. *Engineering Structures*. 222, 111121 (2020).
14. Sun, C., Zheng, G., Zhang, X., Cai, W., Gao, W., Chen, P., Zhou, X.: A decoupled algorithm for cable parameter identification by frequency-ratio-offset. *Computer-Aided Civil and Infrastructure Engineering*. 12980 (2023).

15. Soojin Cho, Jinsuk Yim, Sung Woo Shin, Hyung-Jo Jung, Chung-Bang Yun, Ming L. Wang: Comparative Field Study of Cable Tension Measurement for a Cable-Stayed Bridge. *Journal of Bridge Engineering*. 18, 748–757 (2013).
16. Xu, B., Dan, D., Yu, X.: Real-time online intelligent perception of time-varying cable force based on vibration monitoring. *Engineering Structures*. 270, 114925 (2022).
17. Li, X.: Practical identification algorithm for bending stiffness in cable tension measurements of stay-cables. *China Railway Survey and Design*. 1, 61–64 (2011) (In Chinese).
18. Su, C., Xu, Y., Han, D.: Parameter analysis and identification of bending stiffness of cables during tension measurements by frequency method. *Journal of Highway and Transportation Research and Development*. 22(3), 75–78 (2005) (In Chinese).



Simulation of Photoacoustic Wave Generation and Propagation in Fluid-Solid Coupled Media Using Finite Integration Technique

Kazuyuki Nakahata^(✉) , Akihiro Miki, and Taizo Maruyama

Ehime University, Bunkyo 3, Matsuyama, Ehime 790-8577, Japan
nakahata@cee.ehime-u.ac.jp

Abstract. Photoacoustic (PA) waves using the transmission of a laser beam are being introduced in recent nondestructive ultrasonic testing. Not parameters related to acoustic, optic, and thermal phenomena but also spatial and time conditions of the laser irradiation affect the generation of the PA wave. Therefore, an appropriate design and preparation are required for reliable laser ultrasonic testing. Numerical simulation can be an effective tool for predicting PA signals. This study aims to develop a numerical model to simulate PA wave generation and propagation. Here, the heat conduction and the elastic wave equations were solved in a coupled manner based on discretization by the finite integration technique (FIT). We simulated two cases where the laser were irradiated on the target material in the air and underwater. The amplitude of the longitudinal PA wave propagating in the depth direction in the case of laser irradiation in water became more extensive than in the case of laser irradiation in the air. This reason was that the presence of water in the upper part caused a reaction force against the stress generated by thermal expansion. The FIT simulation was validated by experimental measurement of the PA wave.

Keywords: Laser ultrasonics · Photoacoustic wave · Coupled finite integration technique · Simulation · Experimental measurement · Solid-water interface

1 Introduction

Laser ultrasonics in nondestructive testing (NDT) is a technology for transmitting and receiving ultrasonic waves in a non-contact manner. Since laser ultrasonics enables us to generate ultrasonic waves at a distance from the target object, it is effective for inspection of structural components in infrastructure [1, 2]. Laser ultrasonics uses different mechanisms for transmitting and receiving ultrasonic waves [3]. For the transmission, a pulse laser beam is irradiated onto the surface of the target to generate an ultrasonic wave into the target. On the other hand, a continuous laser pulse is irradiated onto the target surface to receive ultrasound waves. At that time, surface deformation is detected using interference between reflected and reference light. A combination of ultrasonic transmission and reception makes a complete non-contact inspection [4, 5] possible.

There are two ways to generate the ultrasonic wave using the laser as shown in Fig. 1. One is a laser ablation that needs intense pulse energy (Fig. 1(a)). Although the laser ablation mode might deface the target material, several NDT case studies using the ablation mode [6] and the modeling [7] have been reported. The other is a photoacoustic (PA) mode, sometimes called a thermal expansion mode, that uses small laser energy compared with laser ablation (Fig. 1(b)). In the PA mode, the sudden change and confinement of heat generate the stress wave with an ultrasonic frequency range. Diagnostic imaging methods [8] using the PA wave are being introduced actively in the medical field. A photoacoustic microscopy (PAM) using the PA wave was developed as a tool for visualizing vascular networks [9], and recent applications propose a measurement of the oxygen saturation in the vascular [10] based on the difference of light absorption by wavelength. In the human body, the laser penetrates the interior of the body so that PA waves are generated from internal light absorbers such as blood, lipids, lymph, and among others [11]. However, the laser cannot penetrate a deeper portion of industrial materials such as metal and concrete, and the PA wave generates near the surface. Therefore, the generating position of the PA wave and its spreading characteristics in the NDT application are different from those in the medical one.

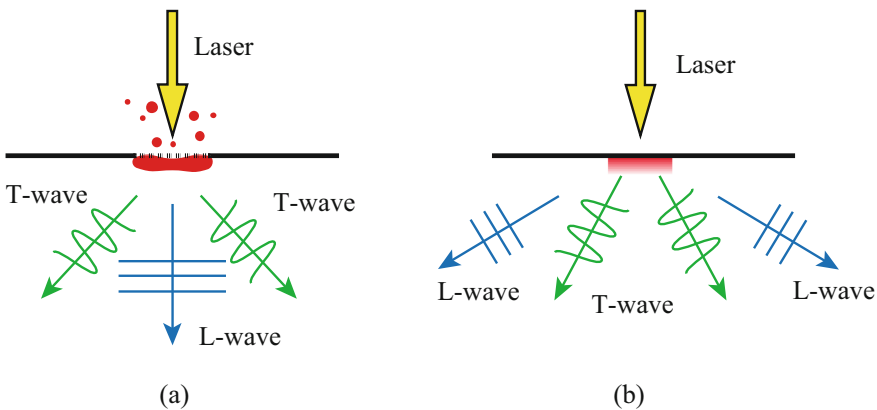


Fig. 1. Ultrasonic wave generation by (a) laser ablation and (b) photoacoustic phenomenon

Theoretical and experimental investigations for the laser ultrasonics for the NDT were carried out widely and diligently in the 1980s [12]. Since entering the 21st century, numerical models for laser ultrasonics have been proposed due to advances in hardware performance and computational techniques. To model the generation and propagation of PA waves induced by laser, it is necessary to solve the coupled equations of heat conduction and wave equation. Yoshikawa and Nishimura [13] analyzed the coupled problem of thermal expansion of a material due to laser irradiation and elastic wave generation in a semi-infinite three-dimensional domain using the boundary element method. Effects of a laser pulse duration and a beam radius in PA wave generation were investigated using the finite element method [14]. Although the laser pulse duration is on the order of nanoseconds, the propagation time of the PA wave in a solid is several tens of microseconds in general nondestructive testing. This time-scale difference sometimes

needs technical and computational operations to simulate the entire process from generating and propagating PA waves in the target area [15]. Nevertheless, the actual PA signal might be received with a low frequency due to the material attenuation and the bandwidth limitation of the receiving sensor. Therefore, a simplified model based on the directivity pattern [16] shown in Fig. 2 is often incorporated into the simulation without solving the coupled equation rigorously. Figure 2 shows the analytical representation of the radiation patterns of longitudinal (L) and transverse (T) waves generated in laser irradiation. This was derived assuming the surface is traction free, and light penetration is very small compared to the acoustic wavelength. In reality, however, the PA waves generated by pulsed laser irradiation highly depend on the laser intensity, pulse duration, spot size, physical properties of the material, interface conditions, etc. Therefore, it should be carefully confirmed whether the model in Fig. 2 is applicable in advance.

We here present a simulation method of the PA wave. In this study, we couple both the heat conduction equation and the wave equation in the framework of the finite integration technique [17] (FIT) to simulate the generation of thermal stresses and the propagation of the PA wave rigorously. Here we model the PA wave generation and propagation in solid using the coupled FIT approach. Sometimes, the laser is used in fluid as water, so it should be helpful to investigate the PA wave by considering the effect with or without a fluid layer. Further, a validation of the simulation is discussed using the PA wave measurement in the later section. Finally, we summarize our research findings and mentions future lines of work.

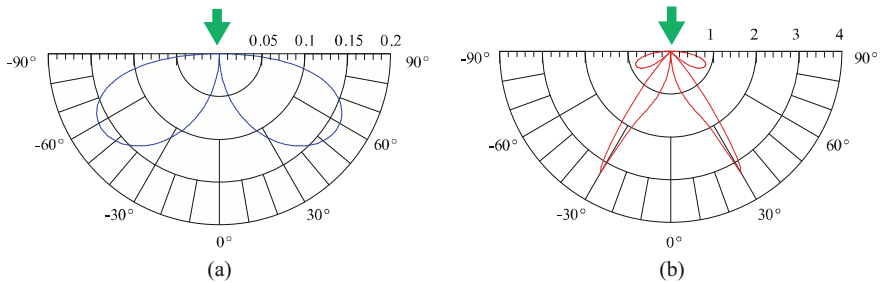


Fig. 2. (a) Directivities of (a) L-wave and (b) T-wave generated by photoacoustic phenomena [16].

2 Modeling of Generation and Propagation of PA Wave

2.1 The Governing Equations for Elastic Wave Propagation

The Cartesian coordinates (x_1, x_2, x_3) are used in this formulation. Since index notation is convenient for the expression in discretization, the summation convention rule is applied below. Let \mathbf{x} and t be the position vector and time, respectively. We denote the stress and the strain in solid/fluid by $\sigma_{ij}(\mathbf{x}, t)$ and $\varepsilon_{ij}(\mathbf{x}, t)$, respectively. When the target

material is illuminated with a laser pulse with an energy less than the melting threshold of the material, a transient stress field is excited due to the thermoelastic expansion.

$$\sigma_{ij} = C_{ijkl}\varepsilon_{kl} - \beta_{ij}T \quad (1)$$

Above equation is known as the Duhamel-Neumann relation [12], which is used for the constitutive relation for a linear elastic body with a change in temperature $T(\mathbf{x}, t)$. Assuming the material is isotropic, the elastic stiffness tensor $C_{ijkl}(\mathbf{x})$ in Eq. (1) is expressed using the Lamé constants $\lambda(\mathbf{x})$, $\mu(\mathbf{x})$ and Kronecker's delta tensor δ_{ij} as follows

$$C_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \quad (2)$$

In addition, $\beta_{ij}(\mathbf{x})$ in Eq. (1) can be written using the linear expansion coefficient $\alpha(\mathbf{x})$ for an isotropic material as in the following equation:

$$\beta_{ij} = (3\lambda + 2\mu)\alpha\delta_{ij} \quad (3)$$

Substituting Eqs. (2) and (3) into Eq. (1) and replacing the strain $\varepsilon_{ij}(\mathbf{x}, t)$ by the displacement $u_i(\mathbf{x}, t)$, we obtain

$$\sigma_{ij} = \lambda \frac{\partial u_k}{\partial x_k} \delta_{ij} + \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) - (3\lambda + 2\mu)\alpha\delta_{ij}T \quad (4)$$

For the discretization of the FIT, we provide an expression for the time derivative of the above equation as

$$\dot{\sigma}_{ij} = \lambda \frac{\partial v_k}{\partial x_k} \delta_{ij} + \mu \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) - (3\lambda + 2\mu)\alpha\delta_{ij}\dot{T} \quad (5)$$

The dot ($\dot{\sigma}_{ij}$) on the character means the derivative in terms of time. The stress $\sigma_{ij}(\mathbf{x}, t)$ and velocity $v_i(\mathbf{x}, t)$ satisfy the following equation of motion

$$\rho\dot{v}_i = \frac{\partial \sigma_{ij}}{\partial x_j} \quad (6)$$

where $\rho(\mathbf{x})$ is the density.

In the FIT for wave propagation, we solve for the stress and velocity in Eqs. (5) and (6). In the case of PA waves, the time evolution of temperature must also be taken into account. The L-wave and T-wave velocities c_L and c_T in a solid, respectively, are expressed as follows

$$c_L = \sqrt{\frac{\lambda + 2\mu}{\rho}}, c_T = \sqrt{\frac{\mu}{\rho}} \quad (7)$$

To represent the acoustic wave in a fluid, $\mu(\mathbf{x}) = 0$ in Eq. (5). Also, since shear stress does not act in a fluid, $\sigma_{ij}(\mathbf{x}, t) = 0$ ($i \neq j$) in Eq. (6).

2.2 The Governing Equations for Heat Conduction

When the laser is irradiating the material, the temperature $T(\mathbf{x}, t)$ is governed by the following heat conduction equation:

$$\rho c \dot{T} = -\frac{\partial q_i}{\partial x_i} + Q \quad (8)$$

where $c(\mathbf{x})$ is the specific heat at constant volume and $Q(\mathbf{x}, t)$ is the heat source. In Eq. (8), $q_i(\mathbf{x}, t)$ is the heat flux and satisfies the following Fourier's law

$$\frac{1}{k} q_i = -\frac{\partial T}{\partial x_i} \quad (9)$$

where $k(\mathbf{x})$ is the thermal conductivity.

To calculate the generation and propagation of PA waves due to laser irradiation, the velocity $v_i(\mathbf{x}, t)$, stress $\sigma_{ij}(\mathbf{x}, t)$, temperature $T(\mathbf{x}, t)$, and heat flux $q_i(\mathbf{x}, t)$ are solved using Eqs. (5), (6), (8), and (9). When the laser irradiates in the target area for a short time, the stress changes due to the rapid temperature rise, as shown in Eq. (5). Here, the deformation by the stress change has little effect on the temperature, so we solve a one-way coupled analysis in which the temperature change contributes to the deformation.

2.3 The Finite Integration Technique

Although the discretization for elastic wave equations was described in detail in past research [17], a coupled approach of the heat conduction and wave equations is focused on below. For simplicity, we consider a two-dimensional (2D) plane strain field (x_1, x_2) . Integrating Eq. (5) over volume V and applying Gauss' divergence theorem, we obtain

$$\int_V \dot{\sigma}_{11} dV = \int_S (\lambda + 2\mu) v_1 n_1 dS + \int_S \lambda v_2 n_2 dS - \int_V (3\lambda + 2\mu) \alpha \dot{T} dV \quad (10)$$

where, \mathbf{n} is the outward normal on V . As the integration volume V in Eq. (10), we consider a square cell with a side length of Δx . Assuming that stress and temperature are constant values in the integration cell, we obtain

$$\dot{\sigma}_{11} \Delta x^2 = (\lambda + 2\mu) [v_1^{(R)} - v_1^{(L)}] \Delta x + \lambda [v_2^{(U)} - v_2^{(D)}] \Delta x - (3\lambda + 2\mu) \alpha \dot{T} \Delta x^2 \quad (11)$$

where superscripts denote the positions of the physical quantities of the integral volume, as shown in Fig. 3(a). The material constants λ , μ , and ρ are assumed to be uniform in the integral cell of stress. Rewriting Eq. (11), we obtain the following equation.

$$\dot{\sigma}_{11} = \frac{(\lambda + 2\mu)}{\Delta x} [v_1^{(R)} - v_1^{(L)}] + \frac{\lambda}{\Delta x} [v_2^{(U)} - v_2^{(D)}] - (3\lambda + 2\mu) \alpha \dot{T} \quad (12)$$

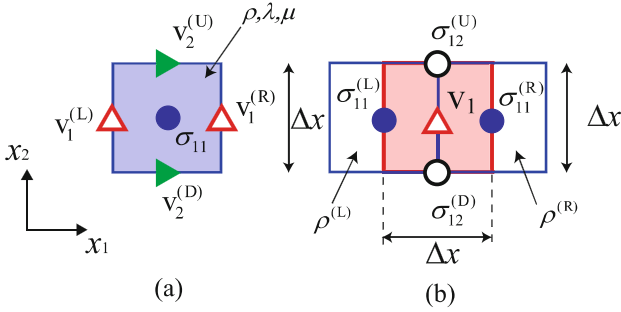


Fig. 3. Integration cells for (a) normal stress σ_{11} and (b) particle velocity v_1

Similarly, integrating and discretizing Eq. (6) yields

$$\dot{v}_1 = \frac{1}{\bar{\rho}_1 \Delta x} \left[\sigma_{11}^{(R)} - \sigma_{11}^{(L)} + \sigma_{12}^{(U)} - \sigma_{12}^{(D)} \right] \tag{13}$$

Since the density is given in the σ_{11} integration cell, as shown in Fig. 3(a), the average value $\bar{\rho} = \frac{1}{2}(\rho^{(L)} + \rho^{(R)})$ is used in the integration of v_1 , as shown in Fig. 3(b).

Next, we show the discretization of Eq. (8). Integrating Eq. (8) over volume V and applying Gauss' divergence theorem, we obtain

$$\int_V \rho c \dot{T} dV = - \int_S (q_1 n_1 + q_2 n_2) dS + \int_V Q dV \tag{14}$$

As shown in Fig. 4(a), the integration cell for T is denoted by blue color. From Eq. (5), it is convenient for the discretization to have a common grid for stress $\sigma_{ij}(\mathbf{x}, t)$ and temperature $T(\mathbf{x}, t)$. The parameters c , ρ , and k related to heat conduction are defined in the integral cell of T . Since T is assumed to be constant in the cell, Eq. (8) becomes.

$$\dot{T} = \frac{1}{\rho c \Delta x} \left[q_1^{(L)} - q_1^{(R)} + q_2^{(D)} - q_2^{(U)} \right] + \frac{Q}{\rho c} \tag{15}$$

Using Eq. (9), the discretization of q_1 is carried out in the same way. As shown in Fig. 4(b), k is defined by the integration cell for T , then we have

$$q_1 = \frac{\bar{k}_1}{\Delta x} \left[T^{(L)} - T^{(R)} \right] \tag{16}$$

where the averaging calculation is performed as $\bar{k}_1 = 2 / \left(\frac{1}{k^{(L)}} + \frac{1}{k^{(R)}} \right)$.

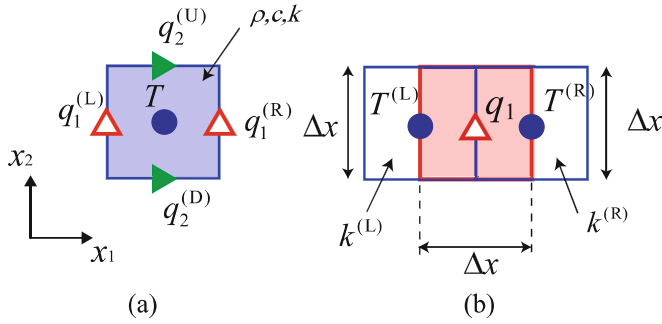


Fig. 4. Integration cells for (a) temperature T and (b) heat flux q_1

2.4 Coupled Analysis

The physical quantities are arranged on a grid, as shown in Fig. 5. Velocity and heat flux, normal stress, and temperature are placed on the same grid. The stresses and velocities are updated using the central difference approximation. In the time domain, the stress components σ_{ij} are allocated at half-time steps, while the velocities v_i are allocated at full-time steps. The following time discretization yields an explicit leap-frog scheme:

$$\{\sigma_{ij}\}^{z+1/2} = \{\sigma_{ij}\}^{z-1/2} + \Delta t \{\dot{\sigma}_{ij}\}^z \quad (17)$$

$$\{v_i\}^{z+1} = \{v_i\}^z + \Delta t \{\dot{v}_i\}^{z+1/2} \quad (18)$$

where Δt is the time step and the superscript z is the integer of the time step.

The temperature effect must also be considered in the calculation of $\dot{\sigma}_{ii}$ in Eq. (5). As shown in Eq. (17), the updates of the stresses are performed at the integer step z . Therefore, the temperature updates are computed in the integer step by using the forward-difference approximation as:

$$\{T\}^{z+1} = \{T\}^z + \Delta t \{\dot{T}\}^z \quad (19)$$

The entire calculation process is summarized below.

- Step 1: Determine the initial and boundary conditions [17].
- Step 2: Calculate heat transfer. Set the heat source Q and then calculate using Eq. (15). Also, update the temperature and heat flux using Eqs. (19) and (16), respectively.
- Step 3: Calculate the stresses. Using the data T^z obtained in Step 2, calculate the stress $\dot{\sigma}_{ij}^z$ in Eq. (12) and then obtain $\sigma_{ij}^{z+1/2}$ using Eq. (17).
- Step 4: Substitute $\sigma_{ij}^{z+1/2}$ into Eq. (13) to obtain \dot{v}_i^{z+1} . Then, substitute \dot{v}_i^{z+1} into Eq. (18) to obtain the velocity field $v_i^{z+1/2}$.
- Step 5: Repeat Steps 2–4 for the estimated time step, and output the calculation results at the required interval.

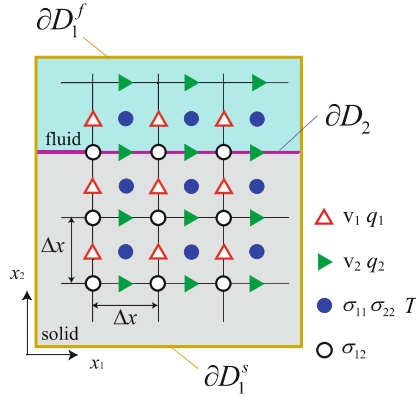


Fig. 5. Allocation of physical quantities in the 2D FIT

2.5 Stability Conditions for Coupled FIT

Since Eqs. (17)–(19) are explicitly updated, the numerical stability conditions must be fulfilled based on Von Neumann’s stability analysis [18]. In the case of the 2D wave equation, the time step Δt must satisfy the following conditions:

$$\Delta t \leq \frac{1}{c_{L\max}} \frac{\Delta x}{\sqrt{2}} \tag{20}$$

where $c_{L\max}$ is the velocity of the fastest wave in the medium. To solve the 2D heat conduction equation with stability,

$$\Delta t \leq \frac{\rho c}{k} \frac{(\Delta x)^2}{4} \tag{21}$$

is required.

The laser pulse duration is short in the PA wave generation, so the time step Δt should be small. This means that the increment number becomes necessarily large, so the update with different time steps between Eqs. (17) and (19) is not beneficial from the standpoint of computational time. In this method, the solutions of the wave equation and the heat conduction equation are updated with the common Δt , which satisfies both Eqs. (20) and (21).

3 Numerical Simulation

3.1 Numerical Models

Figure 6(a) shows numerical model A when the laser is irradiated from air to aluminum. The simulation area (x_1, x_2) is a 5×4 mm rectangle with 0.1 mm absorption layers (perfectly matched layers) on the left, right and bottom sides of the region. Figure 6(b) shows numerical model B when the laser irradiation is from water. In model B, a water layer of 1 mm is placed on top of the aluminum. The material constants are listed in

Table 1. The cell size is set to $\Delta x = 0.5 \mu\text{m}$, and the time step $\Delta t = 0.05 \text{ ns}$. The center of the laser irradiation is $(x_1, x_2) = (2.5 \text{ mm}, 4.0 \text{ mm})$, and the spot radius of the laser and the pulse laser duration are defined as r_0 and t_0 , respectively. The heat source Q is fed into the cell in the spot radius on the aluminum surface. In reality, the PA wave is generated in water by laser irradiation. However, the pressure of the PA wave generated on the aluminum surface is much greater than that in water because of the Grüneisen coefficient [11], representing the efficiency of light-to-heat conversion. In this study, therefore, we ignore the generation of PA waves in water.

Table 1. Material and thermal parameters for aluminum and water.

Material	Aluminum	Water
Density ρ (kg m^{-3})	2688	996.6
Longitudinal wave velocity c_L (m s^{-1})	6400	1470
Transverse wave velocity c_T (m s^{-1})	3150	–
Linear expansion coefficient α (K^{-1})	2.31×10^{-5}	1.00×10^{-4}
Specific heat c ($\text{J kg}^{-1} \text{K}^{-1}$)	905	4179
Thermal conductivity k ($\text{W m}^{-1} \text{K}^{-1}$)	237	0.6104

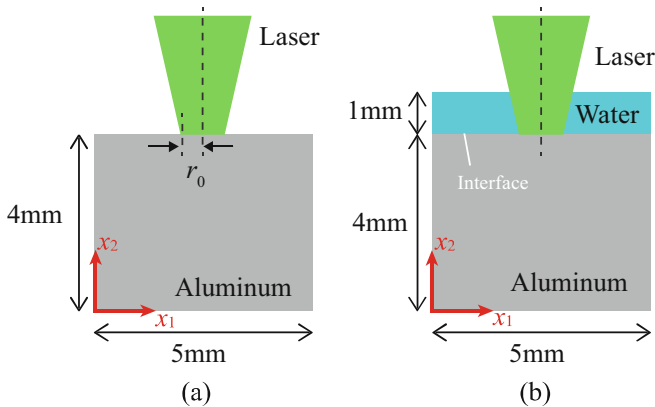


Fig. 6. (a) Numerical model of laser irradiation from air to aluminum body (Model A), (b) numerical model of laser irradiation from water to aluminum body (Model B)

3.2 Simulation Result for Model A

Figure 7 shows the simulation results in Model A when the laser pulse duration t_0 is 7 ns, and the spot radius r_0 is 0.05 mm. The plots of the Mises stress at $t = 320$ and 480 ns after laser irradiation are shown in Fig. 7. The stress values are normalized by dividing the maximum value throughout the entire time step. First, it is shown that the L-wave is

generated mainly in the direction of 60° , not in the x_2 axis direction. On the other hand, the T-wave has more significant stress in the direction of 30° than the L-wave. It can be seen that the wave propagation directions in Fig. 7 are in good agreement with the directivity pattern, as shown in Fig. 2.

Figure 8 shows the simulation results when the spot radius r_0 is 0.428 mm. Here we consider two types of laser pulse duration t_0 . One is 7 ns, and the other is 30 ns. In Fig. 8(a), an L-wave wavefront with a small amplitude can be seen in the x_2 axis direction. The generation of this L-wave wave is caused by the finite spot radius. On the other hand, the Rayleigh waves in the x_1 direction are smaller than those in Fig. 7. Figure 8(b) shows the simulation results in the case of the laser pulse duration $t_0 = 30$ ns while keeping the spot radius $r_0 = 0.428$ mm. The L-wave is hardly seen in the x_2 axis direction. The amplitude of T-waves becomes smaller than in Fig. 8(a).

Let us summarize these simulation results. The key point to increase the gain of the PA wave is that light-thermal conversion takes place in a brief period. Also, the directivity pattern of the PA wave shown in Fig. 2 is applicable when the pulse duration is short, and the spot area is small. Nevertheless, it was demonstrated that the L-wave could be generated in the bottom direction if the laser spot area gets to a specific size.

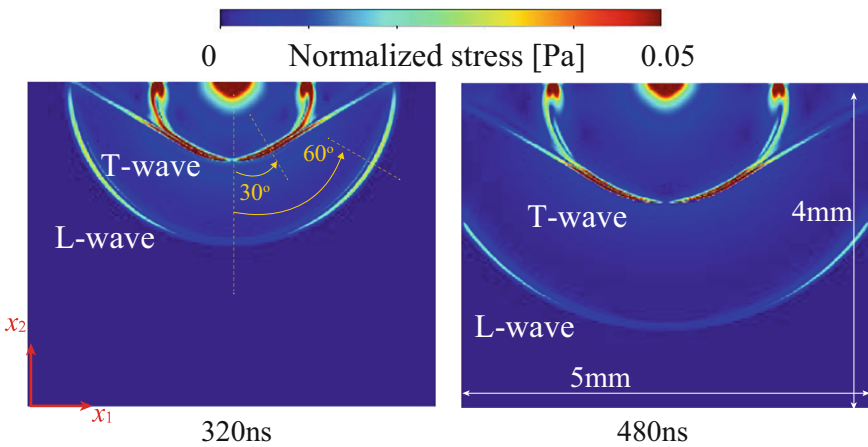


Fig. 7. Snapshots of PA wave generated by laser irradiation in the case of the laser pulse duration $t_0 = 7$ ns and spot radius $r_0 = 0.05$ mm in Model A.

3.3 Simulation Result for Model B

Figure 10 shows the simulation results for Model B when the laser pulse duration t_0 is 7 ns, and the spot radius r_0 is 0.428 mm. It can be seen that pressure (P) waves in water as well as the L- and T-waves in aluminum are generated. The amplitude of the L-wave generated in the x_2 axis direction in aluminum is more significant than that of the laser irradiation in the air. This is considered because the presence of water at the top layer exerts a reaction force on the vertical stress in the solid against the thermal expansion, and then the amplitude of the PA wave in the x_2 axis direction becomes large.

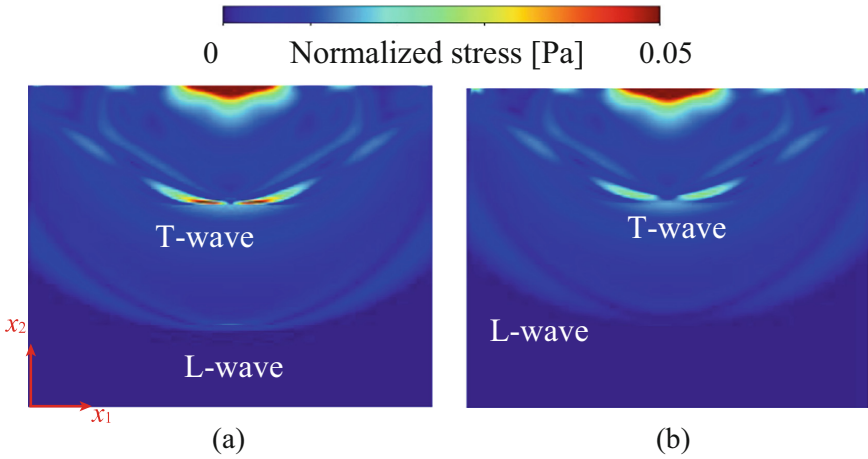


Fig. 8. Snapshots of the PA wave at $t = 480$ ns in the case of the spot radius $r_0 = 0.428$ mm in Model A. (a) Laser pulse duration $t_0 = 7$ ns and (b) $t_0 = 30$ ns.

In the next section, this phenomenon is validated by experimental measurement of the PA wave (Fig. 9).

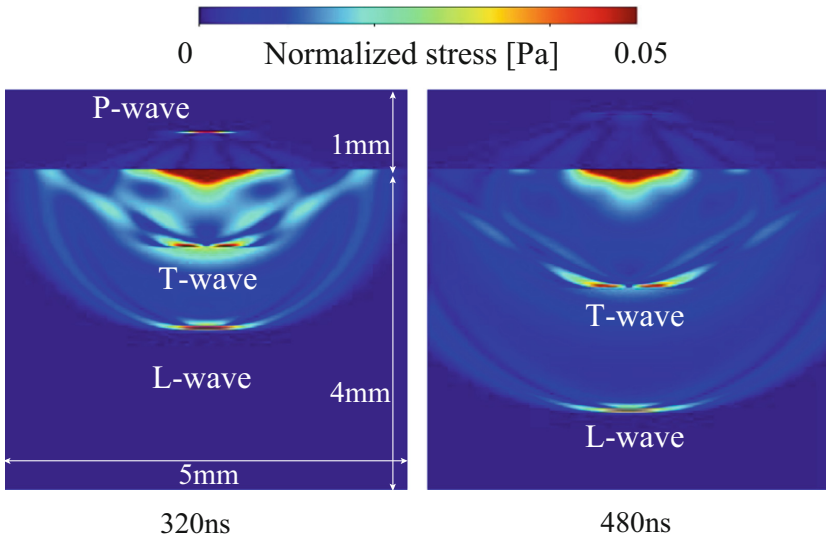


Fig. 9. Snapshots of PA wave generated by laser irradiation in water (Model B). The laser pulse duration $t_0 = 7$ ns and spot radius $r_0 = 0.428$ mm.

4 Validation by PA Wave Measurement

In this experiment, the laser irradiation is scanned, and the generated PA waves are received by a fixed ultrasonic probe. Although the frequency range of the PA wave is wide in reality, the PA waves are modulated by the bandwidth of the receiving probe. Therefore, this experiment cannot be directly compared with the previous section’s simulation. Still, the generation of L- and T-waves and their spreading are helpful to validate the simulation.

4.1 Experimental Setup and Specimen

Figure 10 shows the system configuration for acquiring photoacoustic waves. The generation of laser light is driven by a compact Q-switched Nd:YAG laser operating at approximately 1.0 mJ pulse energy (Nano L90-100, Litron). The laser light source emits a laser with a wavelength of 532 nm and a pulse duration of 7.0 ns. The generated PA waves are detected using an ultrasonic probe (V112-RM, Olympus) with a radius of approximately 6.4 mm and a center frequency of 10 MHz. The measured signals are digitized at a sampling rate of 100 MS/s. Figure 11 shows the laser scan area on an aluminum specimen ($c_L = 6400$ m/s and $c_T = 3150$ m/s). The size of both specimens is $100 \times 100 \times 10$ mm. Figures 11(a) and (b) show the cases of laser irradiation from the air and water, respectively. The laser spot radius is approximately 0.43 mm. We scan for 60 mm in the x direction at a pitch of 0.05 mm while recording the waveform at the probe for each irradiation.

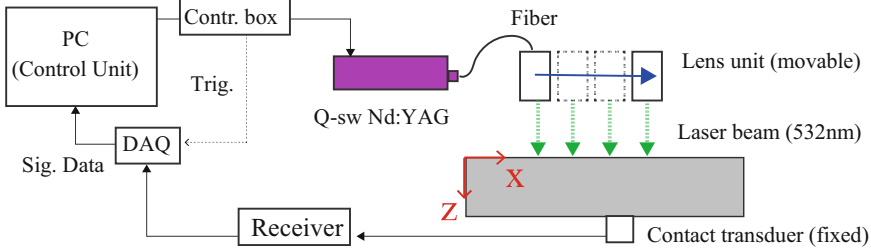


Fig. 10. Flaw diagram of PA measurement system.

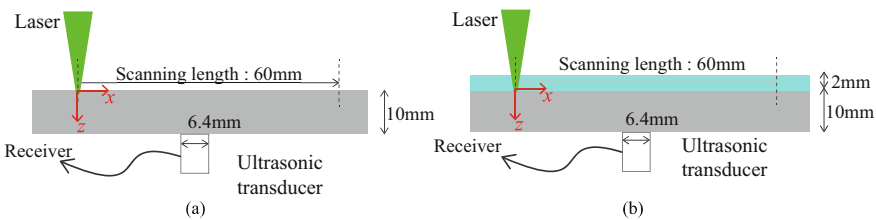


Fig. 11. Experiment in the case of laser irradiation from (a) air and (b) water

4.2 Experimental Results

In Fig. 12, time history waveforms at each irradiation in the x -direction are plotted throughout the scan. The vertical and horizontal axes are the arrival time of the signal and scanning point, respectively. The L-wave arrives at approx. $1.6 \mu\text{s}$, and the T-wave arrives later in Fig. 12. From Fig. 12(a), we can observe the L-wave propagating in the bottom direction of the specimen. The L-wave propagates in the vertical direction at the simulation result in Fig. 8(a), and the measurement shows similar results to the simulation. Figure 12(b) shows the measurement results when a laser is irradiated onto aluminum from water. The amplitude of the L wave is larger than the measurement result in the case of laser irradiation from the air. As the simulation results are shown in Fig. 9, we can observe the amplitude of the L-wave propagating in the vertical direction becomes more significant in water in Fig. 12(b). The above measurement results show the same tendency concerning the L- and T-waves generation as the simulation models A and B.

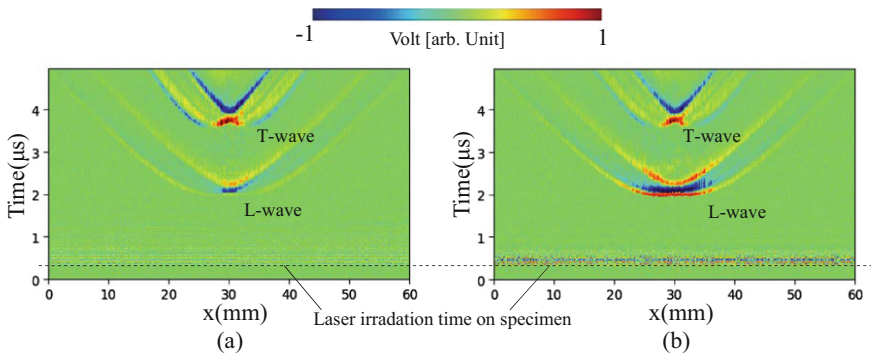


Fig. 12. Time history waveforms at each irradiation in the x -directional scan. The vertical and horizontal axes are the arrival time of the signal and scanning point, respectively. The color shows the amplitude of waveforms in the case of laser irradiation from (a) air and (b) water.

5 Concluding Remarks

In this study, simulations were presented to simulate the PA wave generation and propagation in solid and water. To model the phenomena, we numerically solved the coupled heat conduction equation and wave equation using the FIT. It was shown that the PA waves generated by laser irradiation varied depending on the spot size, laser pulse duration, the material surface condition, and others. We simulated two cases where the laser was irradiated on the target material in the air and underwater. The amplitude of the L-wave propagating in the bottom direction in the case of laser irradiation in water becomes more extensive than in the case of laser irradiation in the air. This reason was assumed that the presence of water in the upper part generated a reaction force against the stress caused by thermal expansion. A similar trend was observed in the measurement; therefore, the proposed FIT modeling was well-validated by the experiment.

We did not reach a direct comparison between the simulation and the experiment in this study. The comparison requires a measurement device such as a laser Doppler vibrometer, which can receive the PA signal at a narrow area. In the future, we would like to extend the 2D simulation to large-scale 3D calculation and model the generation of ultrasonic waves by ablation.

Acknowledgement. This work was supported by JSPS KAKENHI Grant Number 21H01420.

References




1. Wakata, S., Hosoya, N., Hasegawa, N., and Nishikino, M.: Defect detection of concrete in infrastructure based on Rayleigh wave propagation generated by laser-induced plasma shock waves, *Int. J. Mech. Sci.* 218, 107039 (2022).
2. Shimada, Y., Kotyaev, O., Kurahashi, S., Yasuda, N., Misaki, N., Takayama, Y., and Soga, T.: Development of laser-based remote sensing technique for detecting defects of concrete lining, *Electron. Comm. Jpn.* 102, 12–18 (2019).
3. Monchalín, J.P.: Laser-ultrasonics: principles and industrial application, *Proceedings in Ultrasonic and Advanced Methods for Nondestructive Testing and Material Characterization*, 79–115 (2007).
4. Nomura, K., Deno, S., Matsuida, T., Otaki, S., and Asai, S.: In situ measurement of ultrasonic behavior during lap spot welding with laser ultrasonic method, *NDT & E Int.* 130, 102662 (2022).
5. An, Y.K., Park, B., and Sohn, H.: Complete noncontact laser ultrasonic imaging for automated crack visualization in a plate, *Smart Mater. Struct.* 22(2), 025022 (2013).
6. Choi, S. and Jhang, K.Y.: Internal defect detection using laser-generated longitudinal waves in ablation regime, *J. Mech. Sci. Technol.* 32, 4191–4200 (2018).
7. Scherr, J.F., Kollofrath, J., Popovics, J.S., Bühling, B., and Grosse, C.U.: Detection of delaminations in concrete plates using a laser ablation impact echo technique, *J. Nondestruct. Eval.* 42, 11 (2023).
8. Wang, L., and Hu, S.: Photoacoustic tomography: in vivo imaging from organelles to organs, *Science* 335, 1458–1462 (2012).
9. Zhang, H.F., Maslov, K., Stoica, G., and Wang, L.V.: Functional photoacoustic microscopy for high-resolution and noninvasive in vivo imaging, *Nat. Biotechnol.* 24(7), 848–851 (2006).
10. Saijo, Y., Ida, T., Iwazaki, H., Miyajima, J., Tang, H., Shintate, R., Sato, K., Hiratsuka, T., Yoshizawa, S., and Umemura, S.: *Proceedings in Photons Plus Ultrasound: Imaging and Sensing 2019*, 108783E (2019).
11. Okawa, S., Hirasawa, T., Kushibiki, T., and Ishihara, M.: Numerical evaluation of linearized image reconstruction based on finite element method for biomedical photoacoustic imaging, *Opt. Rev.* 20, 442–451 (2013).
12. Scruby, C. B., Dewhurst, R. J., Hutchins, D. A., and Palmer, S. B.: Quantitative studies of thermally-generated elastic waves in laser-irradiated metals, *J. Appl. Phys.* 51, 6210–6216 (1980).
13. Yoshikawa, H., Ohta, Y., and Nishimura, N.: Crack determination using laser-measured horizontal and vertical velocity waveforms of ultrasound, *Struct. Eng. / Earthquake Eng.* 23(2), 279–285 (2006).
14. Soltani, P. and Akbareian, N.: Finite element simulation of laser generated ultrasound waves in aluminum plates, *Lat. Am. J. Solids. Struct.* 11, 1761–1776 (2014).

15. Zhu, Z., Sui, H., Yu, L., Zhu, H., Zhang, J., and Peng, J.: Effective defect features extraction for laser ultrasonic signal processing by using time-frequency analysis, in *IEEE Access* 7, 128706–128713 (2019).
16. Scruby, C. B.: Some applications of laser ultrasound, *Ultrasonics* 27(4), 195–209 (1989).
17. Fellingner, P., Marklein, R., Langenberg, K.J., and Klaholz, S.: Numerical modeling of elastic wave propagation and scattering with EFIT - elastodynamic finite integration technique, *Wave Motion* 21(1), 47–66 (1995).
18. Isaacson, E. and Keller, B. H.: *Analysis of Numerical Methods*, Courier Corporation, 1994.

Numerical and Analytical Methods in Engineering Mathematics



The Singular Function Boundary Integral Method for Solving Three-Dimensional Laplacian Problems with Conical Vertex Singularities

Hind Lamsikine¹ , Otmane Souhar¹ , and Georgios C. Georgiou² 

¹ Université Chouaib Doukkali, El Jadida, Morocco

hindmsikin@gmail.com, souhar.o@ucd.ac.ma

² University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus

georgios@ucy.ac.cy

Abstract. The Singular Function Boundary Integral Method (SFBIM) is extended to solve three-dimensional Laplacian problems with conical vertex singularities. The solution is approximated by the leading terms of the local asymptotic series in spherical coordinates. In order to calculate the unknown singular coefficients, i.e., the vertex stress intensity factors, the Laplacian problem is discretized by applying Galerkin's principle. The governing equation is weighted by the local functions over the domain and the volume integrals are then reduced to surface ones by means of Green's second identity. Given that the local solution satisfies identically the boundary conditions over the conical surface causing the vertex singularity, the dimension of the problem is reduced by one and the boundary integrals need to be calculated only far from the vertex singularity, which yields a considerable reduction of the computational cost. Neumann boundary conditions are weakly imposed and Dirichlet conditions are applied by means of Lagrange multiplier functions. The latter are approximated by means of finite elements over the corresponding boundary parts and the corresponding nodal values are thus additional unknowns. Preliminary results for a model are presented. The advantages and the limitations of the method are also discussed.

Keywords: Laplace equation · Vertex stress intensity factors · Singular function boundary integral method

1 Introduction

Boundary singularities are encountered in many engineering applications, e.g., in fracture and fluid mechanics. In two dimensions, for example, the solutions of the Laplace or the biharmonic equation are singular in the neighborhood of a corner or of a boundary point where there is an abrupt change in the boundary condition (e.g., a crack). The local solution is commonly expressed as an asymptotic expansion, the coefficients of which are generally referred to as singular coefficients [1]. In fracture mechanics, these

coefficients are the stress intensity factors (SIFs), the values of which determine whether a crack will propagate or not. Different methods have been proposed for the calculation of the SIFs [2]. In three dimensions, there exist different types of boundary singularities, such as corner, edge and conical vertex singularities [3–5].

Boundary singularities are also known to limit the regularity of the solution causing serious convergence difficulties to standard numerical methods, such as the finite element and finite-difference methods [6]. These difficulties may not be restricted in the neighborhood of the singularity, propagating in the rest of the problem domain. Special methods have been developed in order to resolve the issues [2, 7].

The Singular Function Boundary Integral Method (SFBIM) is a special method developed for resolving elliptic boundary value problems with boundary singularities [8]. The SFBIM is a Trefftz method meaning that the solution over the entire problem domain is approximated by the leading terms of the local asymptotic expansion [1, 8]. The boundary value problem is discretized weighting the governing elliptic equation in the Galerkin sense by means of the singular functions. The main advantages of the method are: (a) reduction of the problem dimension by one by means of Green's second identity, (b) direct calculation of the leading singular coefficients, as primary unknowns (no postprocessing of the numerical solution), (c) numerical integrations are only carried out in the rest of the boundary, i.e., far from the singularity, given that the singular functions satisfy the boundary conditions at the boundaries causing the singularities; (d) exponential convergence [1, 9]. Another feature of the method is that Dirichlet boundary are imposed by means of Lagrange multipliers [8, 10].

The method was originally developed for two-dimensional Laplacian and biharmonic problems with boundary singularities. Georgiou et al. [8] applied it to the (Laplacian) Motz problem, and reported extremely accurate estimates for the leading singular coefficients. Subsequently, the method has been applied to Laplacian problems with re-entrant corners [11] and to biharmonic elasticity [12] and fluid mechanics problems [13].

The SFBIM has also been applied to three-dimensional Laplace problems with a boundary straight-edge singularity caused by two intersecting flat planes [10]. The local solution has been derived by Yosibash et al. [4], who demonstrated that this involves the eigenpairs of the corresponding two-dimensional problem. The coefficients of the local asymptotic expansion are known as edge flux intensity factors (EFIFs). The application of the SFBIM on a test problem showed that convergence is quite fast and that accurate estimates of the EFIFs are obtained [10].

The objective of the present work is to extend the SFBIM to three-dimensional Laplacian problems with vertex singularities. Despite their importance in elasticity, electromagnetism, and heat transfer, studies of elliptic problems involving vertex singularities are quite limited; see Zaltzman and Yosibash [5] and references therein. Zaltzman and Yosibash [5] derived analytical expressions for the eigenpairs of the solution to the Laplace equation in the neighborhood of a vertex in a three-dimensional domain. They also extended the modified Steklov eigen-formulation for the computation of vertex eigenpairs employing p /spectral finite element methods and presented results for various test problems.

The local solution around of a conical vertex singularity derived by Zaltzman and Yosibash [5] is briefly discussed in Sect. 2. In Sect. 3, the SFBIM is formulated for a

model Laplacian problem. Some preliminary results are presented in Sect. 4. Finally, the conclusions of this work are presented in Sect. 5.

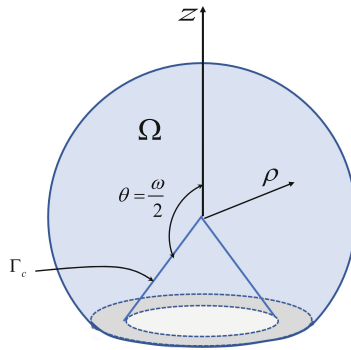


Fig. 1. A three-dimensional domain Ω with a rotationally symmetric conical vertex boundary.

2 A 3D Laplacian Problem with Vertex Singularities

Consider a three-dimensional domain Ω with a rotationally symmetric conical vertex as illustrated in Fig. 1. Employing spherical coordinates (ρ, θ, φ) centered at the vertex so that the z -axis coincides with the symmetry axis of the cone and assume that the angle θ corresponding to the conical surface Γ_c is $\omega/2$. Zaltzman and Yosibash [5] demonstrated that a function $u(\rho, \theta, \varphi)$, satisfying the Laplace equation,

$$\nabla^2 u = \frac{1}{\rho^2} \frac{\partial}{\partial \rho} \left(\rho^2 \frac{\partial u}{\partial \rho} \right) + \frac{1}{\rho^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial u}{\partial \theta} \right) + \frac{1}{\rho^2 \sin^2 \theta} \frac{\partial^2 u}{\partial \varphi^2} = 0 \text{ in } \Omega \quad (1)$$

with either homogeneous Dirichlet,

$$u(\rho, \omega/2, \varphi) = 0 \text{ on } \Gamma_c \quad (2)$$

or homogeneous Neumann boundary condition,

$$\partial u(\rho, \omega/2, \varphi) / \partial n = 0 \text{ on } \Gamma_c \quad (3)$$

is of the form

$$u(\rho, \theta, \varphi) = \sum_{m=0}^{\infty} \sum_{\ell=1}^{\infty} \rho^{\nu_{\ell}^{(m)}} [\alpha_{m,n} \sin(m\varphi) + \beta_{m,n} \cos(m\varphi)] P_{\nu_{\ell}}^m(\cos \theta) \quad (4)$$

where $P_{\nu}^m(x)$ is the first-kind associated Legendre function of degree ν and order m and $\alpha_{m,n}$ and $\beta_{m,n}$ are the singular coefficients or vertex stress intensity factors. The degrees ν_{ℓ} are the positive roots of

$$P_{\nu}^m(\cos(\omega/2)) = 0 \quad (5)$$

in the case of homogeneous Dirichlet condition (Eq. (2)) on Γ_c , or the roots of

$$\left. \frac{1}{\rho} \frac{dP_v^m(\cos \theta)}{d\theta} \right|_{\theta=\omega/2} = 0 \tag{6}$$

in the case of homogeneous Neumann condition. These roots are easily calculated using Mathematica or MATLAB. To this end, the following formula for the associated Legendre function P_v^m [14] is quite useful:

$$P_v^m(\cos \theta) = \frac{\sqrt{2} (-1)^m \Gamma(v + m + 1)}{\sqrt{\pi} \Gamma(v + m + 1) \Gamma(m + 1/2) (\sin \theta)^m} \times \int_0^\theta \frac{\cos(v + 1/2)x}{(\cos x - \cos \theta)^{1/2-m}} dx \tag{7}$$

where $0 < \theta < \pi$, m is integer, and v is real. Equation (4) holds in the neighborhood of the vertex singularity and the constants $\alpha_{m,n}$ and $\beta_{m,n}$ are determined from the boundary conditions on the rest of the domain, i.e., on $\Omega \setminus \Gamma_c$.

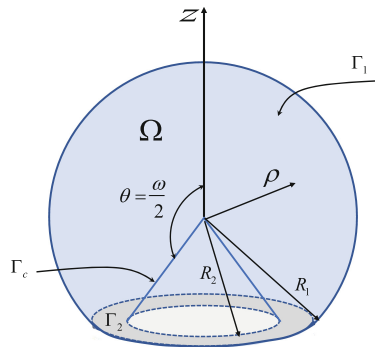


Fig. 2. Domain and bounding surfaces of the model Laplacian problem considered in Sect. 3.

3 Implementation of the SFBIM

In this section, we consider the following model Laplacian problem with a vertex singularity in a three-dimensional domain Ω

$$\left. \begin{aligned} \nabla^2 u &= 0 \text{ in } \Omega \\ u &= 0 \text{ on } \Gamma_c \\ u &= 1 \text{ on } \Gamma_1 \\ \frac{\partial u}{\partial z} &= 0 \text{ on } \Gamma_2 \end{aligned} \right\} \tag{8}$$

where $\partial\Omega = \Gamma_c \cup \Gamma_1 \cup \Gamma_2$. As illustrated in Fig. 2, the surface Γ_1 is a part of the surface of a sphere of radius R_1 , while Γ_2 is the surface of a flat hollow circle.

In the singular function boundary integral method, the solution of the problem in Eq. (8) is approximated by the leading terms of the local solution in the neighborhood of the singularity, given by Eq. (4). Let

$$\bar{u} = \sum_{m=0}^{N_m} \sum_{\ell=1}^{N_\ell} \sum_{s=1}^2 A_{m\ell}^{(s)} W_{m\ell}^{(s)} \tag{9}$$

where the singular functions $W_{m\ell}^{(s)}$ are defined by

$$W_{m\ell}^{(1)} = \sin(m\varphi)\rho^{v_\ell^m} P_{v_\ell}^m \text{ and } W_{m\ell}^{(2)} = \cos(m\varphi)\rho^{v_\ell^m} P_{v_\ell}^m \tag{10}$$

and $A_{m\ell}^{(s)}$ are the approximations of the vertex stress intensity factors to be determined. The total number of these coefficients is $N_\alpha = 2(N_m + 1)N_\ell$.

The singular functions $W_{m\ell}^{(s)}$ satisfy Laplace’s equation and the boundary conditions along the conical surface Γ_c . These functions are used to weight the governing equation over the domain Ω in the Galerkin sense, yielding the following N_α discretized equations:

$$\int_{\Omega} \nabla^2 \bar{u} W_{kw}^{(\sigma)} dV = 0, \quad k = 0, 1, \dots, N_m, \quad w = 1, \dots, N_\ell, \quad \sigma = 1, 2 \tag{11}$$

Given that the singular functions $W_{kw}^{(\sigma)}$ satisfy Laplace’s equation, double application of Green’s theorem leads to the following boundary integral equations:

$$\int_{\partial\Omega} \left(\frac{\partial \bar{u}}{\partial n} W_{kw}^{(\sigma)} - \bar{u} \frac{\partial W_{kw}^{(\sigma)}}{\partial n} \right) dS = 0 \tag{12}$$

where the indices are the same as in Eq. (11) and n denotes the normal direction to the boundary. Since the singular functions satisfy the boundary condition along the conical surface causing the vertex singularity, the integrand in Eq. (12) vanishes on Γ_c . Hence, one can write

$$\int_{\Gamma_1} \left(\frac{\partial \bar{u}}{\partial \rho} W_{kw}^{(\sigma)} - \bar{u} \frac{\partial W_{kw}^{(\sigma)}}{\partial \rho} \right) dS + \int_{\Gamma_2} \left(-\frac{\partial \bar{u}}{\partial z} W_{kw}^{(\sigma)} + \bar{u} \frac{\partial W_{kw}^{(\sigma)}}{\partial z} \right) dS = 0 \tag{13}$$

Note that the first term in the second integrand vanishes due to the homogeneous Neumann boundary condition on Γ_2 . In order to impose the Dirichlet boundary condition on Γ_1 , a Lagrange multiplier function

$$\lambda(\theta, \varphi) = \left. \frac{\partial \bar{u}}{\partial \rho} \right|_{\rho=R_1} \tag{14}$$

is introduced, which is approximated by means of polynomial basis functions Ψ_i :

$$\lambda(\theta, \varphi) = \sum_{i=1}^{N_\lambda} \lambda_i \Psi_i(\theta, \varphi) \tag{15}$$

where $\lambda_i, i = 1, \dots, N_\lambda$ are the unknown discrete Lagrange multipliers. The additional algebraic equations needed for these additional unknowns are obtained by weighting the Dirichlet condition on Γ_1 with the basis functions Ψ_i . Thus, one obtains a system of $N_a + N_\lambda$ algebraic equations:

$$\int_{\Gamma_1} \left(\lambda W_{kw}^{(\sigma)} - \bar{u} \frac{\partial W_{kw}^{(\sigma)}}{\partial \rho} \right) dS + \int_{\Gamma_2} \bar{u} \frac{\partial W_{kw}^{(\sigma)}}{\partial z} dS = 0, \tag{16}$$

$$k = 0, 1, \dots, N_m, \quad w = 1, \dots, N_\ell, \quad \sigma = 1, 2$$

and

$$\int_{\Gamma_1} \bar{u} \Psi_j dS = \int_{\Gamma_1} \Psi_j dS, \quad j = 1, \dots, N_\lambda \tag{17}$$

Let $X_1 = (A_{01}^{(1)}, A_{01}^{(2)}, A_{02}^{(1)}, \dots, A_{N_a N_\lambda}^{(1)}, A_{N_a N_\lambda}^{(2)})^T$ and $X_2 = (\lambda_1, \dots, \lambda_{N_\lambda})^T$ be the vectors containing the unknown singular coefficients and Lagrange multipliers, respectively. Then, the system of Eqs. (16) and (17) can be written in the following block form:

$$\begin{bmatrix} K & L \\ L^T & O \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} B \\ C \end{bmatrix} \tag{18}$$

It is easily verified that the stiffness matrix is symmetric. It should be noted that the system is singular when $N_\lambda > N_a$.

4 Results and Discussion

Following Zaltzman and Yosibash [5], we consider their first test problem where homogeneous Dirichlet condition holds over the conical surface Γ_c . Moreover, it is assumed that the rest of the boundary $\partial\Omega \setminus \Gamma_c$ is spherical. Therefore, the problem to be solved becomes:

$$\left. \begin{aligned} \nabla^2 u(\rho, \theta, \varphi) &= 0 \text{ in } \Omega \\ u(\rho, \omega/2, \varphi) &= 0 \text{ on } \Gamma_c \\ u(R_1, \theta, \varphi) &= g(\theta, \varphi) \text{ on } \partial\Omega \setminus \Gamma_c \end{aligned} \right\} \tag{19}$$

Below, we show preliminary results obtained for $\omega = 6\pi/4$. All eigenvalues have been computed numerically. We consider the relative errors of the leading two vertex stress intensity factors, which are tabulated in Table 1. In the present preliminary computations, the Lagrange multipliers functions are approximated by constants basis functions. One observes that the values of the vertex stress intensity factors converge rapidly with N_λ . Figure 3 shows the relative errors which occur in the computed values of the two vertex stress intensity factors β_1^N and β_2^N versus N_λ when $N_a = 30$. Again, fast convergence is observed.

Table 1. Convergence of the first two vertex stress intensity factors of the problem in Eq. (19) with N_λ when $N_a = 30$.

N_λ	β_1^N	β_2^N
2	0.16541800385138092	0.2638862956512482
3	-0.159152901686569	0.9966369688466924
5	0.2662356310588355	0.30900580588055543
8	0.28804394848386233	0.15372385115262274
Exact	0.28	0.1

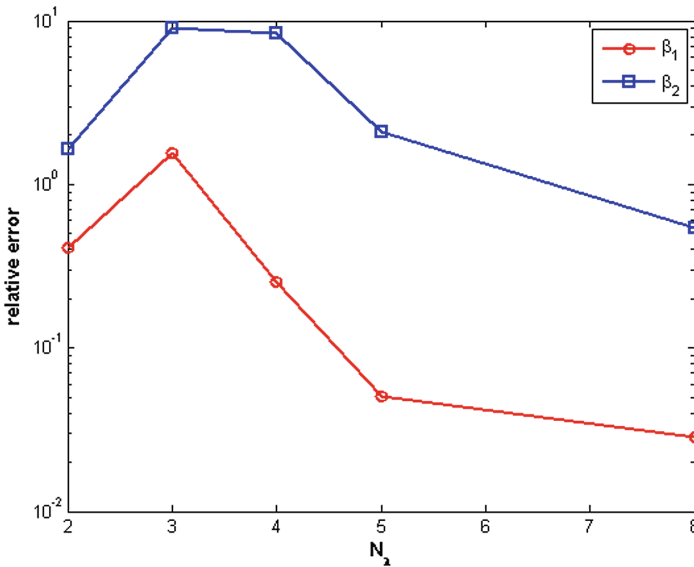


Fig. 3. Preliminary computations of the relative errors of the first two vertex stress intensity factors of the problem in Eq. (19) when $N_a = 30$.

5 Conclusions




The Singular Function Boundary Integral Method (SFBIM) has been used to solve three-dimensional Laplacian test problems with conical vertex singularities. The convergence of the method has been investigated by varying the number of Lagrange multipliers. With the SFBIM, the leading vertex stress intensity factors of the local asymptotic solution expansion are the main unknowns of the formulation of the method. The latter are calculated explicitly, which is important advantage over the other numerical methods, which necessitate the post-processing of the numerical solution to get the vertex stress intensity factors. Preliminary computations show that the method converges fast with the number of Lagrange multipliers.

References

1. Christodoulou, E., Elliotis, M., Xenophontos, C., Georgiou, G., Analysis of the singular function boundary integral method for a biharmonic problem with one boundary singularity, *Numer. Methods PDEs.* 28, 749–767 (2012).
2. Brenner, S.C., Multigrid methods for the computation of singular solutions and stress intensity factors I: corner singularities, *Math. Comput.* 68, 559–583 (1999).
3. Stephan, E., Whiteman, J.R., Singularities of the Laplacian at corners and edges of three-dimensional domains and their treatment with finite element methods, *Math. Methods. App. Sci.* 10, 339–350 (1988).
4. Yosibash, Z., Actis, R., Szabó, B., Extracting edge flux intensity functions for the Laplacian, *Int. J. Numer. Methods Eng.* 53, 225–242 (2002).
5. Zaltzman, T., Yosibash, Z., Vertex singularities associated with conical point for the 3D Laplace equation, *Numer. Methods PDEs.* 27, 662–679 (2011).
6. Brenner, S.C., Scott, L.R., *The Mathematical Theory of Finite Element Methods*, Springer, New York, 1994.
7. Babuška, I., von Petersdoof, T., Anderson, B., Numerical treatment of vertex singularities and intensity factors for mixed boundary value problems for the Laplace equation in R^3 , *SIAM J. Numer. Anal.* 31, 1265–1288 (1994).
8. Georgiou, G.C., Olson, L., Smyrlis, G., A singular function boundary integral method for the Laplace equation, *Commun. Numer. Meth. Eng.* 12, 127–134 (1996).
9. Xenophontos, C., Elliotis, M., Georgiou, G., A singular function boundary integral method for Laplacian problems with singularities, *SIAM J. Sci. Comp.* 28, 517–532 (2006).
10. Christodoulou, E.C., Elliotis, M., Xenophontos, C., Georgiou, G.C., The singular Function Boundary Integral Method For 3-D Laplacian problems with a boundary straight-edge singularity, *Appl. Math. Comp.* 219, 1073–1081 (2012).
11. Elliotis, M., Georgiou, G., Xenophontos, C., The solution of a Laplacian problem over an L-shaped domain with a singular function boundary integral method, *Comm. Numer. Methods Eng.* 18, 213–222 (2002).
12. Elliotis, M., Georgiou, G., Xenophontos, C., The singular function boundary integral method for biharmonic problems with crack singularities, *Eng. Anal. Bound. Elem.* 31, 209–215 (2007).
13. Elliotis, M., Georgiou, G., Xenophontos, C., Solution of the planer Newtonian stick-slip problem with a singular function boundary integral method, *Int. J. Numer. Methods Fluids.* 48, 1000–1021 (2005).
14. Lebedev, N.N., *Special Functions and Their Applications*, Dover, New York, 1972.



Analytical Solutions of Axial Annular Newtonian Flows with Dynamic Wall Slip

Meryieme EL Farragui¹, Otmane Souhar¹ , and Georgios C. Georgiou²  

¹ Université Chouaib Doukkali, El Jadida, Morocco
souhar.o@ucd.ac.ma

² University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus
georgios@ucy.ac.cy

Abstract. We consider the cessation of annular Poiseuille and annular Couette flows of a Newtonian fluid, under the assumption that wall slip occurs following a dynamic slip law, i.e., a law involving a slip-relaxation parameter. The relaxation time-dependent term forces the eigenvalue parameter to appear in the boundary conditions and, thus, the resulting spatial problems correspond to Sturm-Liouville problems different from their static-Navier slip counterparts. The orthogonality conditions of the associated eigenfunctions for both flows of interest are derived and analytical closed-form solutions are then obtained. For comparison purposes, the analytical solutions corresponding to the static Navier slip condition, are derived. Comparisons are also made with the plane Poiseuille and Couette flows when the annular gap becomes small. As expected, flow dynamics becomes slower in the presence of wall slip and this effect is accentuated by increasing the slip-relaxation parameter.

Keywords: Annular Poiseuille flow · Annular Couette flow · Newtonian fluid · Navier slip · Dynamic slip

1 Introduction

The classical no-slip boundary condition, dictating that fluid particles adjacent to a (fixed or moving) wall stick to it, is violated in many important macroscopic flows of complex and also simple fluids [1, 2]. Let us denote by u_w^* the slip velocity, defined as the relative velocity of the fluid particles with respect to that of the wall. Wall slip reduces the required pressure drop and finds application in microfluidics [3]. However, it also has undesired effects, causing, for example, instabilities in certain flows of industrial importance [4, 5]. It also affects measurements in viscometric experiments and needs to be taken into account, in order to determine the true rheology of the fluid under study [6]. The mismatch of viscosity data obtained from different rheometers or from the same rheometer using different radii or gaps is often attributed to wall slip. Different techniques are used in order to get correct estimates of the rheological parameters, such as the Mooney method for capillary rheometers [7] and those suggested by Yoshimura and Prud'homme [8] for circular Couette and parallel disk rheometers.

In general, the slip velocity depends not only on the fluid properties and the flow conditions, such as the shear and normal stresses (which include pressure) and the temperature, but also on the physical properties of the wall/fluid interface [1]. Navier's law [9] assumes that the slip velocity varies linearly the wall shear stress, τ_w^* , and introduces a single material parameter β^* that incorporates all other effects:

$$u_w^* = \frac{\tau_w^*}{\beta^*} \quad (1)$$

The no-slip boundary condition corresponds to the limiting case where $\beta^* \rightarrow \infty$. The above equation is static, i.e., the slip velocity is independent of the history of the fluid particles. More involved static slip equations have been proposed in the literature, such as the power-law generalization of the Navier law, non-monotonic slip equations, and equations predicting slip above a stress threshold [1].

Of special interest to the present work are dynamic slip equations that apply to time-dependent flows and include a relaxation time λ^* . The dynamic extension of Eq. (1) takes the form

$$u_w^* + \lambda^* \frac{\partial u_w^*}{\partial t^*} = \frac{\tau_w^*}{\beta^*} \quad (2)$$

where t^* is the time. It should be pointed out that in steady-state flows, Eq. (2) is equivalent to the Navier-slip Eq. (1).

Kaoullas and Georgiou [10] used separation of variables to derive analytical solutions for the start-up and cessation of Newtonian plane and axisymmetric Poiseuille and plane a circular Couette flows with dynamic wall slip following Eq. (2). They pointed out that since the eigenvalue parameter also appears in the boundary conditions, the resulting spatial problem corresponds to a Sturm–Liouville problem different from that found using the static Navier law (1). The main observation is that dynamic wall slip damps the flow development even more than static wall slip. The same conclusion was also made by Abou-Dina et al. [11], who obtained equivalent analytical solutions of the start-up Newtonian Couette flow with dynamic wall slip along the fixed wall using separation of variables and also the one-sided Fourier transform methods.

The objective of the present work is to derive analytical solutions of two Newtonian flows in an annular tube with dynamic wall slip: (a) the start-up of annular Poiseuille flow, i.e., the flow caused by imposing a constant pressure-gradient; and (b) the start-up annular Couette flow, i.e., the flow resulting by suddenly moving the inner cylinder in the axial direction. Both flows find industrial applications, e.g., in the oil industry. To the best of our knowledge, the effects of dynamic wall slip on these two flows have not been investigated in the literature.

The analytical solutions of the two flows of interest are derived in Sects. 2 and 3. In Sect. 2, the cessation of annular Poiseuille flow is analysed. The cessation of annular Couette flow is studied in Sect. 3. Results for both flows are presented and discussed in Sect. 4. Concluding remarks are provided in Sect. 5.

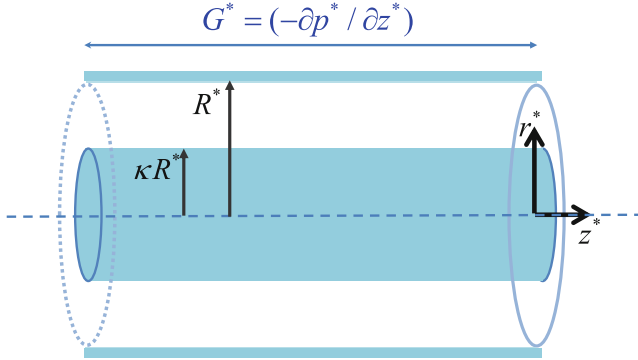


Fig. 1. Geometry of annular Poiseuille flow; the flow is driven by a pressure gradient $G^* = (-\partial p^*/\partial z^*)$.

2 Annular Poiseuille Flow

We consider the annular Poiseuille flow, i.e., the pressure-driven flow of a Newtonian fluid in an infinitely long annular tube. Let the radii of the inner and outer cylinders be κR^* and R^* , where $0 < \kappa < 1$, as illustrated in Fig. 1. Using cylindrical coordinates (r^*, θ, z^*) and assuming that the flow is incompressible and axisymmetric and gravity is negligible, the velocity $u_z^* = u_z^*(r^*, t^*)$ and the z-component of the momentum equation becomes [12]:

$$\rho^* \frac{\partial u_z^*}{\partial t^*} = G^* + \eta^* \frac{1}{r^*} \frac{\partial}{\partial r^*} \left(r^* \frac{\partial u_z^*}{\partial r^*} \right) \tag{3}$$

where $G^* = (-\partial p^*/\partial z^*)$ is the pressure gradient, and ρ^* and η^* are respectively the fluid density and viscosity, both of which are constant.

In the present work, we assume that dynamic wall slip occurs along the two cylinders, i.e., at $r^* = \kappa R^*$ and $r^* = R^*$, and denote the corresponding slip velocities by u_{w1}^* and u_{w2}^* :

$$\left. \begin{aligned} u_z^*(\kappa R^*, t^*) &= u_{w1}^*(t^*) \\ u_z^*(R^*, t^*) &= u_{w2}^*(t^*) \end{aligned} \right\} \tag{4}$$

Slip obeys the dynamic-slip law in Eq. (2), where

$$\tau_w^* = \left| \tau_{rz}^* \right|_w = \left| \eta \frac{\partial u_z^*}{\partial r^*} \right|_w \tag{5}$$

Given that the velocity near the two cylinders is smaller than in the core of the annulus, the velocity derivative is positive near the inner cylinder and negative near the outer cylinder. Therefore, applying the dynamic-slip law we obtain the following boundary conditions:

$$\left. \begin{aligned} u_z^*(\kappa R^*, t^*) + \lambda^* \frac{\partial u_z^*(\kappa R^*, t^*)}{\partial t^*} &= \frac{\eta^*}{\beta^*} \frac{\partial u_z^*(\kappa R^*, t^*)}{\partial r^*} \\ u_z^*(R^*, t^*) + \lambda^* \frac{\partial u_z^*(R^*, t^*)}{\partial t^*} &= -\frac{\eta^*}{\beta^*} \frac{\partial u_z^*(R^*, t^*)}{\partial r^*} \end{aligned} \right\} \tag{6}$$

2.1 Steady-State Solution

Let us denote by $\bar{u}_z^* = \bar{u}_z^*(r^*)$ the steady-state velocity. Combining the simplified Eqs. (3) and (6), one obtains the following boundary value problem:

$$\left. \begin{aligned} \frac{d}{dr^*} \left(r^* \frac{d\bar{u}_z^*}{dr^*} \right) &= -\frac{G^* r^*}{\eta^*} \\ \bar{u}_z^*(\kappa R^*, t^*) &= \frac{\eta^*}{\beta^*} \frac{d\bar{u}_z^*(\kappa R^*, t^*)}{dr^*} \\ \bar{u}_z^*(R^*, t^*) &= -\frac{\eta^*}{\beta^*} \frac{d\bar{u}_z^*(R^*, t^*)}{dr^*} \end{aligned} \right\} \tag{7}$$

It is straightforward to integrate and apply the boundary conditions [13]. One then obtains the following solution:

$$\bar{u}_z^*(r^*) = \frac{G^* R^{*2}}{4\eta^*} \left[1 + 2(1 - \sigma^2)B - \frac{r^{*2}}{R^{*2}} - 2\sigma^2 \ln \frac{R^*}{r^*} \right] \tag{8}$$

where

$$B \equiv \frac{\eta^*}{\beta^* R^*} \tag{9}$$

is the dimensionless slip number and

$$\sigma^2 = \frac{1 - \kappa^2 + 2(1 + \kappa)B}{2[\ln(1/\kappa) + (1 + 1/\kappa)B]} \tag{10}$$

It should be noted that $r_M^* = \sigma R^*$ is the radius at which the steady-state velocity is maximum. The two slip velocities are also easily calculated:

$$\bar{u}_{w1}^* = \frac{G^* R^{*2} B}{2\kappa \eta^*} (\sigma^2 - \kappa^2) \tag{11}$$

and

$$\bar{u}_{w2}^* = \frac{G^* R^{*2} B}{2\eta^*} (1 - \sigma^2) \tag{12}$$

For the steady-state volumetric flow rate $\bar{Q}^* = 2\pi \int_0^{R^*} \bar{u}_z^*(r^*) r^* dr^*$, one gets

$$\bar{Q}^* = \frac{\pi G^* R^{*4}}{8\eta^*} \left\{ 1 - \kappa^4 + 4(1 + \kappa^3)B - \frac{[1 - \kappa^2 + 2(1 + \kappa)B]^2}{\ln(1/\kappa) + (1 + 1/\kappa)B} \right\} \tag{13}$$

Scaling the velocity $\bar{u}_z^*(r^*)$ by the mean velocity in the annular tube,

$$\bar{V}^* = \frac{G^* R^{*2}}{8\eta^*(1 - \kappa^2)} \left\{ 1 - \kappa^4 + 4(1 + \kappa^3)B - \frac{[1 - \kappa^2 + 2(1 + \kappa)B]^2}{\ln(1/\kappa) + (1 + 1/\kappa)B} \right\} \tag{14}$$

we obtain the following expression for the dimensionless velocity:

$$\bar{u}_z(r) \equiv \frac{\bar{u}_z^*(r^*)}{\bar{V}^*} = \frac{2(1 - \kappa^2)[1 + 2(1 - \sigma^2)B - r^2 - 2\sigma^2 \ln \frac{1}{r^*}]}{1 - \kappa^4 + 4(1 + \kappa^3)B - \frac{[1 - \kappa^2 + 2(1 + \kappa)B]^2}{\ln(1/\kappa) + (1 + 1/\kappa)B}} \tag{15}$$

where $r = r^*/R^*$. The no-slip solution is recovered by setting $B = 0$. The dimensionless slip velocities are given by

$$\bar{u}_{w1} = \frac{4(1 - \kappa^2)(\sigma^2 - \kappa^2)B}{\kappa \left\{ 1 - \kappa^4 + 4(1 + \kappa^3)B - \frac{[1 - \kappa^2 + 2(1 + \kappa)B]^2}{\ln(1/\kappa) + (1 + 1/\kappa)B} \right\}}, \quad \bar{u}_{w2} = \kappa \bar{u}_{w1} \tag{16}$$

2.2 Cessation Flow

We consider now the cessation flow, assuming that initially ($t^* = 0$) the velocity is the steady-state solution in Eq. (15) and that the pressure gradient suddenly ($t^* > 0$) vanishes. We work with dimensionless equations scaling r^* by R^* , u_z^* by the mean steady-state velocity \bar{V}^* , and the time t^* by ρ^*R^{*2}/η^* . The resulting initial boundary value problem can be written as follows:

$$\left. \begin{aligned} \frac{\partial u_z}{\partial t} &= \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u_z}{\partial r} \right), \quad \kappa \leq r \leq 1, \quad t \geq 0 \\ u_z(\kappa, t) + \Lambda \frac{\partial u_z(\kappa, t)}{\partial t} &= B \frac{\partial u_z(\kappa, t)}{\partial r}, \quad t \geq 0 \\ u_z(1, t) + \Lambda \frac{\partial u_z(1, t)}{\partial t} &= -B \frac{\partial u_z(1, t)}{\partial r}, \quad t \geq 0 \\ u_z(r, 0) &= \bar{u}_z(r), \quad \kappa \leq r \leq 1 \end{aligned} \right\} \tag{17}$$

where

$$\Lambda \equiv \frac{\lambda^* \eta^*}{\rho^* R^{*2}} \tag{18}$$

is the dimensionless slip relaxation number. To solve problem (17) we use the method of separation of variables. The following expression is obtained for the velocity:

$$u_z(r, t) = \sum_{n=1}^{\infty} A_n Z_0^{(n)}(\alpha_n r) e^{-\alpha_n^2 t} \tag{19}$$

where $Z_i^{(n)}(x) \equiv J_i(x) + \beta_n Y_i(x)$, $i = 0, 1$, and $J_i(x)$ and $Y_i(x)$ are the i th-order Bessel functions of the first and second kind, respectively, and (α_n, β_n) are the n th solution of the system:

$$\left. \begin{aligned} Z_0^{(n)}(\kappa \alpha_n)(1 - \Lambda \alpha_n^2) &= -B \alpha_n Z_1^{(n)}(\kappa \alpha_n) \\ Z_0^{(n)}(\alpha_n)(1 - \Lambda \alpha_n^2) &= B \alpha_n Z_1^{(n)}(\alpha_n) \end{aligned} \right\} \tag{20}$$

The constants A_n are determined upon application of the initial condition in Eq. (17) and the proper orthogonality condition:

$$A_n = \frac{4(1 - \kappa^2)(1 - \Lambda\alpha_n^2)}{B\alpha_n^2[(B + 2\Lambda)Z_1^{(n)2}(\alpha_n) + \kappa(2\Lambda - \kappa B)Z_1^{(n)2}(\kappa\alpha_n)] + (1 - \Lambda\alpha_n^2)^2[Z_1^{(n)2}(\alpha_n) - \kappa^2 Z_1^{(n)2}(\kappa\alpha_n)]} \times \left\{ \begin{aligned} &\Lambda\alpha_n^4[2B(1 - \sigma^2)Z_1^{(n)}(\alpha_n) - \kappa(2B(1 - \sigma^2) + 1 - \kappa^2 - 2\sigma^2 \ln(1/\kappa))Z_1^{(n)}(\kappa\alpha_n)] \\ &+ 2(1 - \Lambda\alpha_n^2)[2 + B(1 - \sigma^2)\alpha_n^2][Z_1^{(n)}(\alpha_n) - \kappa Z_1^{(n)}(\kappa\alpha_n)] \\ &- 2B\alpha_n^2[(1 - \sigma^2)Z_1^{(n)}(\alpha_n) - (\sigma^2 - \kappa^2)Z_1^{(n)}(\kappa\alpha_n)] \end{aligned} \right\} \alpha_n^3 \left\{ 1 - \kappa^4 + 4(1 + \kappa^3)B - \frac{[1 - \kappa^2 + (1 + \kappa)B]^2}{\ln(1/\kappa) + (1 + 1/\kappa)B} \right\} \quad (21)$$

The two slip velocities are simply given by:

$$u_{w1}(t) = u_z(\kappa, t) = \sum_{n=1}^{\infty} A_n Z_0^{(n)}(\kappa\alpha_n) e^{-\alpha_n^2 t} \quad (22)$$

and

$$u_{w2}(t) = u_z(1, t) = \sum_{n=1}^{\infty} A_n Z_0^{(n)}(\alpha_n) e^{-\alpha_n^2 t} \quad (23)$$

Also, the volumetric flow rate is

$$Q(t) = \sum_{n=1}^{\infty} \frac{A_n}{\alpha_n} [Z_1^{(n)}(\alpha_n) - \kappa Z_1^{(n)}(\kappa\alpha_n)] e^{-\alpha_n^2 t} \quad (24)$$

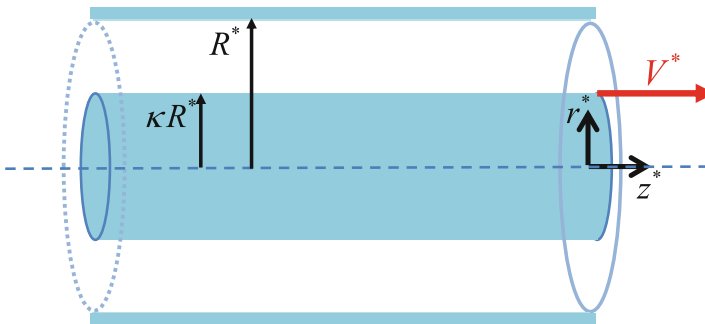


Fig. 2. Geometry of annular Couette flow; the flow is driven by the motion of the inner cylinder.

3 Annular Couette Flow

In this section, the annular Couette flow is considered, i.e., the flow between two coaxial cylinders of infinite length and radii κR^* and R^* , which is driven by the axial motion of the inner cylinder.

3.1 Steady-State Solution

We first consider the steady-state flow in which the inner cylinder moves at a constant speed V^* , as illustrated in Fig. 2. Given that the pressure gradient is zero, integrating the steady-state version of Eq. (3) gives $\bar{u}_z^*(r^*) = c_1^* \ln r^* + c_2^*$, where the constants c_1^* and c_2^* are determined from the boundary conditions:

$$\left. \begin{aligned} u_z^*(\kappa R^*) &= V^* - u_{w1}^* \\ u_z^*(R^*) &= u_{w2}^* \end{aligned} \right\} \tag{25}$$

Since the velocity is a decreasing function of the radial distance, the shear stress $\tau_{rz}^* = \eta^* du_z^*/dr^*$ is negative everywhere in the flow domain, and thus Eq. (25) takes the form:

$$\left. \begin{aligned} u_z^*(\kappa R^*) &= V^* + \frac{\eta^*}{\beta^*} \frac{du_z^*(\kappa R^*)}{dr^*} \\ u_z^*(R^*) &= -\frac{\eta^*}{\beta^*} \frac{du_z^*(R^*)}{dr^*} \end{aligned} \right\} \tag{26}$$

where the static version of Eq. (2), i.e., Navier slip, has been employed. It is straightforward to show that the steady-state velocity is given by

$$\bar{u}_z^*(r^*) = \frac{B + \ln(R^*/r^*)}{(1 + 1/k)B + \ln(1/\kappa)} V^* \tag{27}$$

The volumetric flow rate is found to be

$$Q^* = \frac{\pi(1 - \kappa^2)R^{*2}V^*[1 + 2B - 2\ln(1/\kappa)/(1/\kappa^2 - 1)]}{2[(1 + 1/k)B + \ln(1/\kappa)]} \tag{28}$$

3.2 Cessation Flow

We consider the cessation flow, i.e., we assume that at $t^* = 0$ the velocity is given by Eq. (27), i.e., the steady-state solution, and suddenly the inner cylinder stops moving. We work with dimensionless equations scaling lengths and time as in Sect. 2 and the velocity u_z^* by the initial speed V^* of the inner cylinder. The resulting initial boundary value problem is identical to that in Eq. (17). Therefore, the velocity is given by Eq. (25), where (α_n, β_n) are the n th positive solution of Eq. (20) and the constants A_n are found by applying the initial condition

$$u_z(r, 0) = \bar{u}_z(r) = \frac{B + \ln(1/r)}{(1 + 1/k)B + \ln(1/\kappa)} \tag{29}$$

Using the proper orthogonality condition, one finds that

$$A_n = - \frac{2\kappa(1 - \Lambda\alpha_n^2)Z_1^{(n)}(\kappa\alpha_n)/\alpha_n}{\left\{ \begin{aligned} &B\alpha_n^2 \left[(B + 2\Lambda)Z_1^{(n)2}(\alpha_n) + \kappa(2\Lambda - \kappa B)Z_1^{(n)2}(\kappa\alpha_n) \right] \\ &+ (1 - \Lambda\alpha_n^2)^2 \left[Z_1^{(n)2}(\alpha_n) - \kappa^2 Z_1^{(n)2}(\kappa\alpha_n) \right] \end{aligned} \right\}} \tag{30}$$

It should be noted that the two slip velocities and the volumetric flow rate are given by Eqs. (22)–(24).

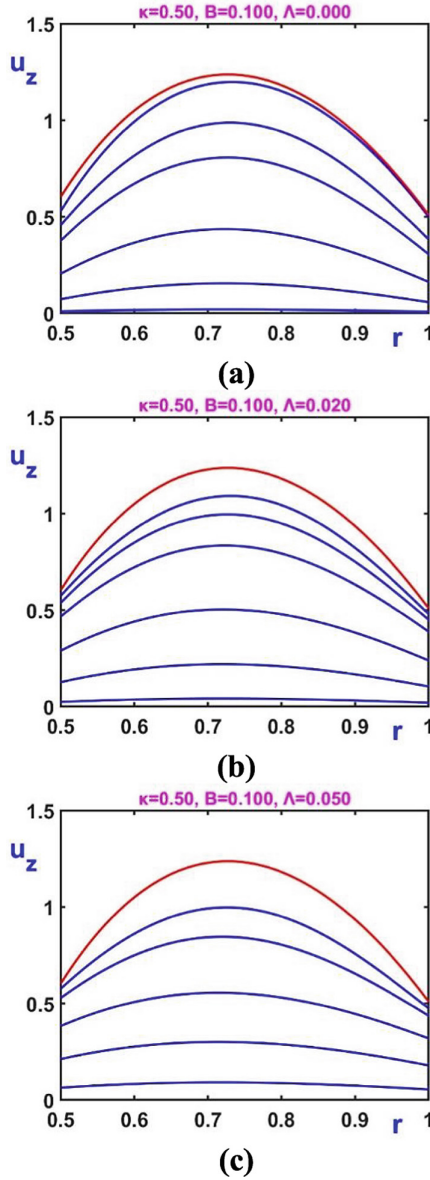
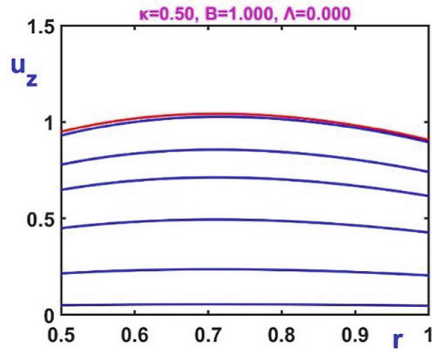
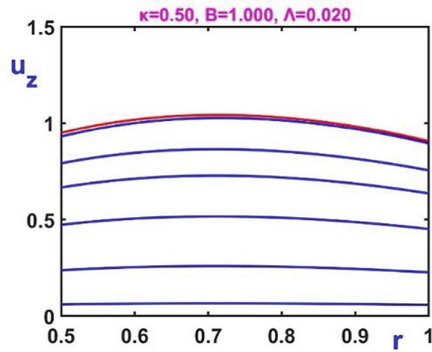


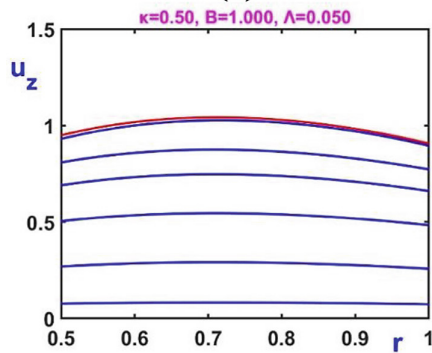
Fig. 3. Evolution of the dimensionless velocity in cessation of annular Poiseuille flow with dynamic wall slip when $B = 0.1$ (moderate slip) and $\kappa = 0.5$: (a) $\Lambda = 0$ (Navier slip; $t = 0, 0.01, 0.02, 0.05, 0.1, 0.2$); (b) $\Lambda = 0.02$ ($t = 0, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2$); (c) $\Lambda = 0.05$ ($t = 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.4$). The red profile is the initial steady-state solution.



(a)



(b)



(c)

Fig. 4. Evolution of the dimensionless velocity in cessation of annular Poiseuille flow with dynamic wall slip when $B = 1$ (strong slip) and $\kappa = 0.5$: (a) $\Lambda = 0$ (Navier slip; $t = 0, 0.05, 0.1, 0.2, 0.4, 0.8$); (b) $\Lambda = 0.02$ ($t = 0, 0.05, 0.1, 0.2, 0.4, 0.8$); (c) $\Lambda = 0.05$ ($t = 0, 0.05, 0.1, 0.2, 0.4, 0.8$). The red profile is the initial steady-state solution.

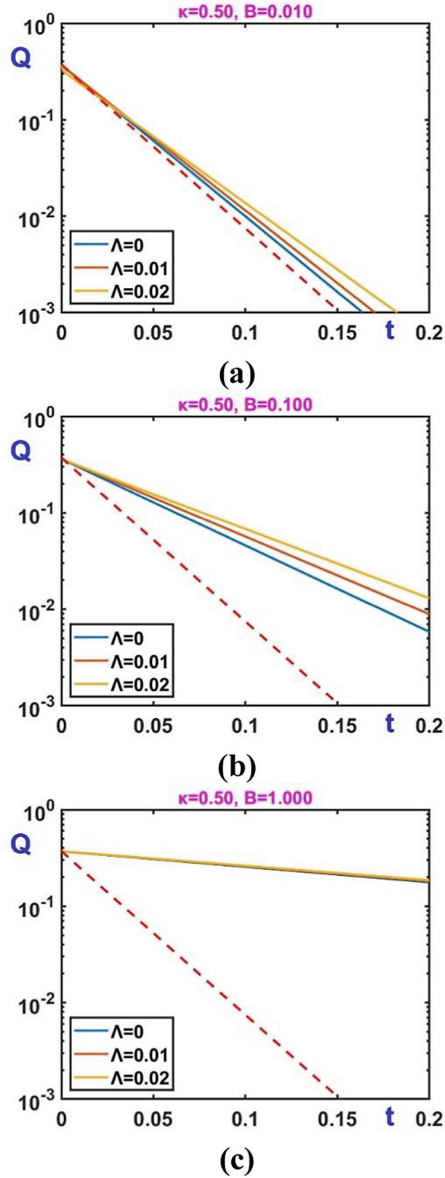


Fig. 5. Evolution of the dimensionless volumetric flow rate in cessation of annular Poiseuille flow with dynamic wall slip when $\kappa = 0.5$ and $\Lambda = 0$ (Navier slip), 0.01 and 0.02: (a) $B = 0.01$ (weak slip); (b) $B = 0.1$ (moderate slip); (c) $B = 1$ (strong slip). The red dashed line corresponds to the no-slip case.

4 Results and Discussion

Results have been obtained for various values of the radii ratio κ and wide ranges of the slip parameters B and Λ . The Fourier modes, i.e., the solutions of the system in Eq. (20) are easily determined using standard methods, marching along the positive x -axis calculating all the roots till the desired number N is reached. Our numerical experiments revealed that considering $N = 1000$ terms of the solution in Eq. (19) was sufficient to ensure series convergence. It should be noted that the series was not convergent at very small times (of the order of 10^{-6}) resulting in oscillations, that persist even when much more terms (up to $N = 10^6$) are employed. However, at such small times the time-dependent solution does not differ much from the initial steady-state solutions in Eqs. (15) and (27) for the Poiseuille and Couette flows, respectively. Next, results for $\kappa = 0.5$ are presented and discussed.

4.1 Annular Poiseuille Flow

The effect of the relaxation parameter Λ on the evolution of the velocity is illustrated in Figs. 3 and 4, where results for $B = 0.1$ (moderate slip) and 1 (strong slip) are shown. Note that the top plots correspond to $\Lambda = 0$, i.e., to (static) Navier slip. In agreement with the literature [10, 11], the evolution of the velocity becomes slower as the value of Λ is increased. When slip is strong, which is the case in Fig. 4, the velocity profiles become rather flat and cessation is much slower and the effect of Λ is not so pronounced. In other words, dynamic wall slip is not important when slip is very strong.

The combined effects of the slip and relaxation numbers on the evolution of the flow are also illustrated in Fig. 5, where the calculated volumetric flow rates for different values of the two parameters are plotted. One observes that the evolution of $Q(t)$ becomes slower when B or Λ are increased and that the effect of Λ is more pronounced when slip is weak (Fig. 5a) or moderate (Fig. 5b).

4.2 Annular Couette Flow

The evolution of the velocity in the case of no or static wall slip is illustrated in Fig. 6, where results for $\Lambda = 0$ and $B = 0, 0.1$ and 1 are shown. As expected, the initial steady-state velocity profile tends to become flat and cessation is damped as wall slip becomes stronger. It is also clear that initially only the flow adjacent the inner cylinder of the annulus is affected. The inner slip velocity u_{w1} is initially much bigger than u_{w2} . As the phenomenon is developed the difference between the two slip velocities diminishes, as they both tend to zero, and the velocity distribution tends to become flatter and symmetric.

The effect of the relaxation parameter is illustrated in Fig. 7, where results for $B = 0.1$ (moderate slip) and different values of Λ are shown. Again, cessation is slowed down and the velocity distribution tends to become flat near the inner cylinder. The evolution of the two slip velocities in the three cases of Fig. 7 is shown in Fig. 8. The combined effects of B and Λ on the volumetric flow rate are shown in Fig. 9. The role of the relaxation parameter is important only when slip is weak or moderate.

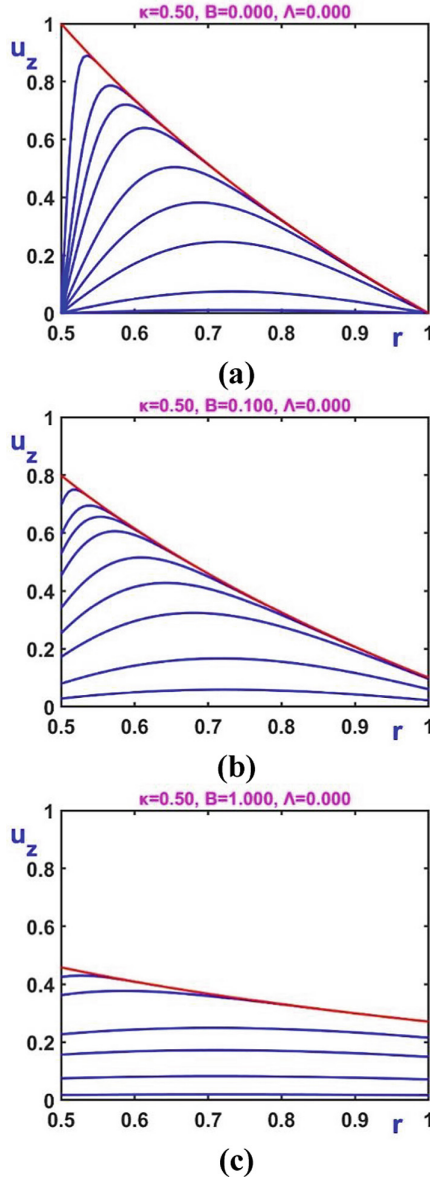
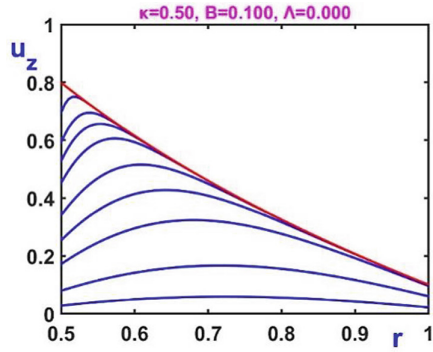
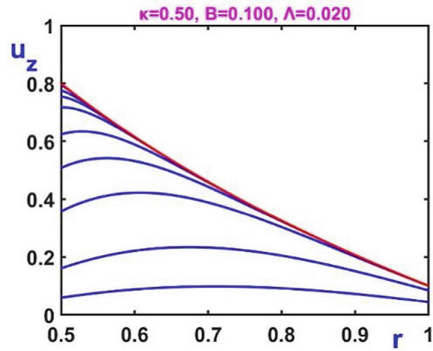


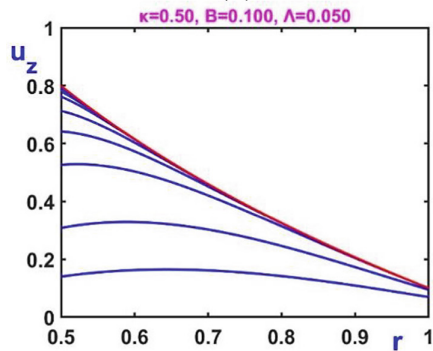
Fig. 6. Evolution of the dimensionless velocity in cessation of annular Couette flow with Navier slip when $\kappa = 0.5$: (a) $B = 0$ (no slip; $t = 0, 0.0005, 0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1$); (b) $B = 0.1$ (moderate slip; same times as in (a)); (c) $B = 1$ (strong slip; $t = 0, 0.001, 0.01, 0.1, 0.2, 0.4, 0.8$). The red profile is the initial steady-state solution.



(a)



(b)



(c)

Fig. 7. Evolution of the dimensionless velocity in cessation of annular Couette flow with dynamic wall slip when $\kappa = 0.5$ and $B = 0.1$: (a) $\Lambda = 0$ (Navier slip; $t = 0, 0.0005, 0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1$); (b) $\Lambda = 0.02$ ($t = 0, 0.0001, 0.0005, 0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1$); (c) $B = 1$ (same times as in (b)). The red profile is the initial steady-state solution.

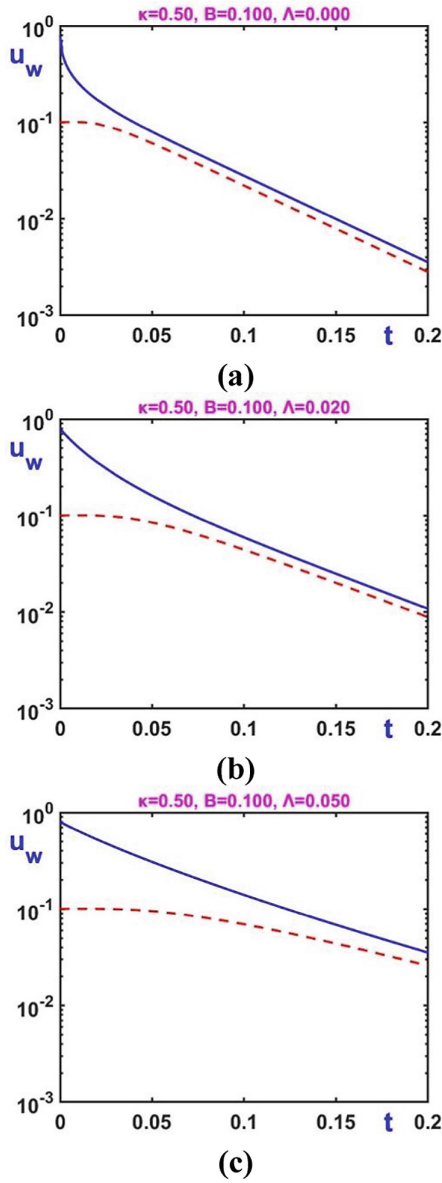


Fig. 8. Evolution of the two slip velocities u_{w1} (solid blue) u_{w2} (dashed red) in cessation of annular Couette flow when $\kappa = 0.5$ and $B = 0.1$: (a) $\Lambda = 0$ (Navier slip); (b) $\Lambda = 0.02$; (c) $\Lambda = 0.05$.

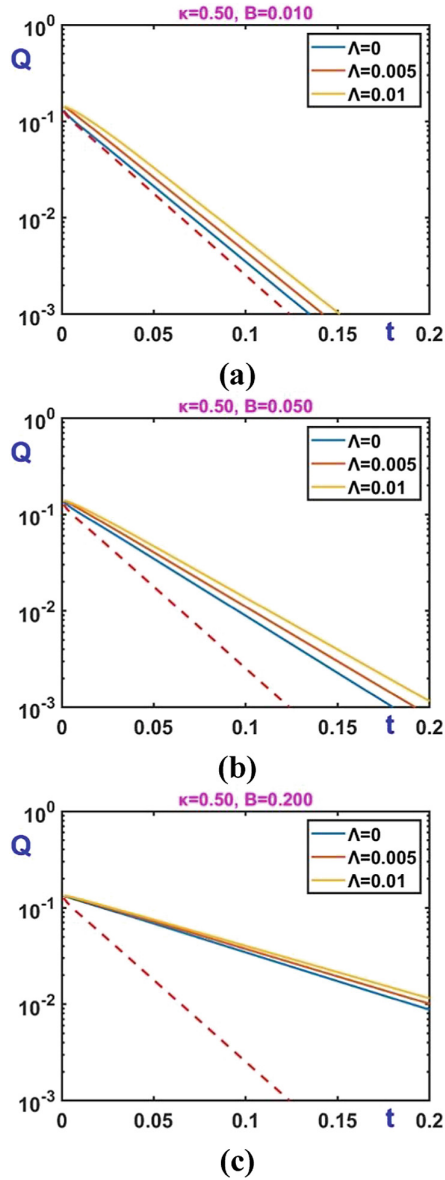


Fig. 9. Evolution of the dimensionless volumetric flow rate in cessation of annular Poiseuille flow with dynamic wall slip when $\kappa = 0.5$ and $\Lambda = 0$ (Navier slip), 0.005 and 0.01: (a) $B = 0.01$ (weak slip); (b) $B = 0.1$ (moderate slip); (c) $B = 1$ (strong slip). The red dashed line corresponds to the no-slip case.

5 Conclusions

We used separation of variables to derive analytical solutions for the cessation of annular Poiseuille and Couette flows of a Newtonian fluid exhibiting dynamic wall slip. It has been demonstrated that, under the assumption that the same slip law applies along both cylinders of the annulus, the two solutions share the same Fourier modes, the only difference being in the initial condition. Representative results for a radii ratio $\kappa = 0.5$ have been presented showing that cessation is damped by the slip and relaxation parameters, in agreement with previous studies in the literature for other flows [10, 11]. It has also been demonstrated that the effect of the slip relaxation parameter is not significant when slip is strong.

References

1. Hatzikiriakos, S.G., Wall slip of molten polymers, *Prog. Polym. Sci.* 37, 624–643 (2012).
2. Neto, C., Evans, D.R., Bonaccorso, E., Butt, H.J., Craig, V.S.J., Boundary slip in Newtonian liquids: a review of experimental studies, *Rep. Prog. Phys.* 68, 2859–2897 (2005).
3. Stone, H.A., Stroock, A.D., Ajdari, A., Engineering flows in small devices: microfluidics toward a lab-on-a-chip, *Annu Rev Fluid Mech* 36, 381–411 (2004).
4. Denn, M.M., Extrusion instabilities and wall slip, *Annu. Rev. Fluid Mech.* 33, 265–287 (2001).
5. Georgiou, G., Stick-slip instability. Chapter 6 in: *Polymer Processing Instabilities: Control and Understanding*, S.G. Hatzikiriakos and K. Migler (Eds.), pp. 161–206, Mercel Dekker, Inc (2004).
6. Moud, A.A., Piette, J., Danesh, M., Georgiou, G.C., Hatzikiriakos, S.G., Apparent slip in colloidal suspensions, *J. Rheology* 66(1), 79–90 (2022).
7. Mooney, M., Explicit formulas for slip and fluidity, *J. Rheology* 2, 210–222 (1931).
8. Yoshimura, A., Prud’homme, A.K., Wall slip corrections for Couette and parallel disk viscometers, *J. Rheology* 32, 53–67 (1988).
9. Navier, C.L.M.H., Sur les lois du mouvement des fluids, *Mem. Acad. R. Sci. Inst. Fr.* 6, 389–440 (1827).
10. Kaoullas, G., Georgiou, G.C., Start-up Newtonian Poiseuille and Couette flows with dynamic wall slip, *Meccanica* 50, 1747–1760 (2015).
11. Abou-Dina, M.S., Helal, M.A., Ghaleb, A., Kaoullas, G., Georgiou, G.C. Newtonian plane Couette flow with dynamic wall slip, *Meccanica* 55, 1499–1507 (2020).
12. Papanastasiou, T., Georgiou, G., Alexandrou, A., *Viscous Fluid Flow*, CRC Press, Boca Raton (2000).
13. Gryparis E., Georgiou, G.C., Annular Poiseuille flow of Bingham fluids with wall slip, *Phys. Fluids* 34, 033103 (2022).



A Mathematical Approach for Recreation Non-symmetric 2D Railway Alignments

Gerardo Casal¹, Alberte Castro², Duarte Santamarina¹(✉),
and Miguel E. Vázquez-Méndez^{1,3}

¹ Departamento de Matemática Aplicada, Universidade de Santiago de Compostela,
Escola Politécnica Superior de Enxeñaría, 27002 Lugo, Spain

`xerardo.casal@usc.es`, `duarte.santamarina@usc.es`

² Departamento de Enxeñaría Agroforestal, Universidade de Santiago de
Compostela, Escola Politécnica Superior de Enxeñaría, 27002 Lugo, Spain

`alberte.castro@usc.es`

³ CITMaga, 15782 Santiago de Compostela, Spain

`miguelernesto.vazquez@usc.es`

Abstract. The constant passage of trains on the railways tracks causes, in the course of time, geometric deviations that must be corrected periodically by means of a track calibration process. It consists of designing a new horizontal alignment, called recreated alignment, as close as possible to the deformed center track fulfilling also the technical constraints according to the operational requirements of the railway. In recent years different models have been proposed to address this task. This paper proposes, firstly, a new geometrical model for the definition of horizontal alignments to deal with non symmetric transition curves at both sides of a circular curve and secondly, a two-stage optimization algorithm to compute the recreated alignment. The usefulness of this model is tested with two academic examples showing its good behavior. The paper ends with a conclusions section in which a brief description of future work is also presented.

1 Introduction

The passage of trains on the railways over time causes deviations of the tracks with respect to its original position. The geometric deviations can be measured by means of a *track surveying* process, which provides, among other relevant information, a set of coordinate points along the track center line. This track deviations must be amended periodically to guarantee the safety of train convoys as well as the comfort of passengers. It is therefore necessary to create a horizontal alignment rectifying the layout by a track calibration process. The movements made by the recalibrating machine (called tamping and lining machine) are limited in size therefore the new alignment would have to move the track as little as possible from its deformed state. During this calibration the track center line is adjusted by adopting a recreated horizontal alignment (HA) defined as close as possible to the surveyed points considering that current technical constraints

must be fulfilled according to the operational requirements of the railway (Fig. 1).



Fig. 1. Railway deviation, track surveying machine and tamping and lining machine.

This recreation process is a complex job, usually performed in the literature by a two steps process. The first step consists in the analysis of the existing alignment by means of the identification of all its geometrical elements (tangents, circular curves and transition curves). This is a difficult task due to the fact that, even knowing the number of curves in a track section is not always easy to identify properly to which of the three elements belong the measured points and locate the boundaries between them (the so-called tangency points). Several methods have been proposed for this identification by analyzing the changes on the curvature graphs using circular curves (see [8, 9, 13]), using spline techniques [1, 5] or using the heading direction graph, which is less noisy (see [2, 10, 15]).

The second step consists in, once the elements present in the path are known, the aim is to obtain an alignment with those elements that fits, as much as possible, to the existing (deformed) path, creating a new alignment complying with all the technical specifications. In order to accomplish this task some authors have used Image processing [6, 7, 16]. A more recent approach use some artificial vision techniques, such as edge detectors and Hough transform, to identify the geometrical elements of existing alignments from satellite or photographic images. Recently, [11, 12] propose a new methodology for HA recreation called swing iterations where an iterative process of reclassification of the measured points previously assigned to a certain type of geometrical element is carried out for the identification of the alignment components. Later, the geometrical parameters of the alignment obtained from the identification stage is optimized by means of a genetic algorithm (GA) developed by the authors [11].

In this work two significant advances in this recreation technique are presented. The first one is linked to the parametrization of the track. An alignment is formed by an adequate combination of three types of geometrical elements: tangents, circular curves and transition curves. Generally, the design of newly constructed railway lines try to use always symmetrical transition curves (usually clothoids) at each side of the circular curves. However, many old railways line

routes that have had several remodeling (e.g. to eliminate level crossings or due to small route redesigns) ended up with transition curves that do not fit into this standard. Hence, it should be taken into account the possibility of the existence of non-symmetric transition curves in order to define an appropriate recreated alignment. Taking this fact into consideration, we have modified the symmetric parametrization presented in [3, 18] in terms of the arc length parameter to this non-symmetric scenario.

The second advance is linked to the definition of the recreated design. We propose a two-stage optimization method to recover a HA close to the surveyed points in which there is no need to do the identification of the existing alignment by means of its geometrical elements. The only requirement is to know the number of curves involved in the recreation. The algorithm is capable of doing this identification automatically. In the first stage we propose to create an initial layout for the second step, close to the alignment (see [3, 17, 20]). This is done by using a *coarse* (easy and fast computing) objective function which minimizes the y -distance between a feasible alignment and a spline obtained by the set of points measured along the railway line acquired by track surveying, a sort of least square horizontal alignment fitting. This first step has two intentions: the first one is to obtain a valid HA close to the points to easily identify to which section each point might belong (circular curve, tangent or clothoid) and additionally to obtain an initial seed for the second optimization stage (more complex and computationally expensive). In this second stage the problem is to minimize the orthogonal distance of a HA to the points measured by track surveying.

Finally, to prove the good behavior of the method, numerical results are presented in two academic test. The paper ends (Sect. 4) with some interesting conclusions and a brief outline of future work.

2 Mathematical Model

2.1 Geometrical Model

The horizontal alignment (HA) of a road or a railway is a 2D curve formed by the appropriate combination of tangents, circular curves and transition curves (clothoids) joining two terminals $\mathbf{P}_a, \mathbf{P}_b \in \mathbb{R}^2$. In previous works by the authors (see [3, 18]), a model with N symmetrical curves of type clothoid-circular curve-clothoid with the vertex on the convex side of the curve was considered. There, a HA was determined by its end points $\mathbf{P}_a, \mathbf{P}_b \in \mathbb{R}^2$, their main tangent intersection points (IPs) $\mathbf{V}_j = (X_j, Y_j) \in \mathbb{R}^2$, and the radii $R_j > 0$ and angles $\omega_j \geq 0$ of the circular curves. However, it has been noticed that many old railway lines which have been remodeled over the years to eliminate level crossings or due to small route redesigns (e.g. bypass avoiding villages where the convoys no longer stops, etc.), turned out with transition curves that do not fit into this standard. Hence, it should be taken into account the possibility of the existence of non-symmetric transition curves in order to define an appropriate recreated alignment. In this new scenario it is necessary to introduce a new degree of

freedom on each curve to take into consideration the non-symmetry and being able to properly and uniquely define the alignment (see Fig. 2). In this case the decision variable would be represented as follows

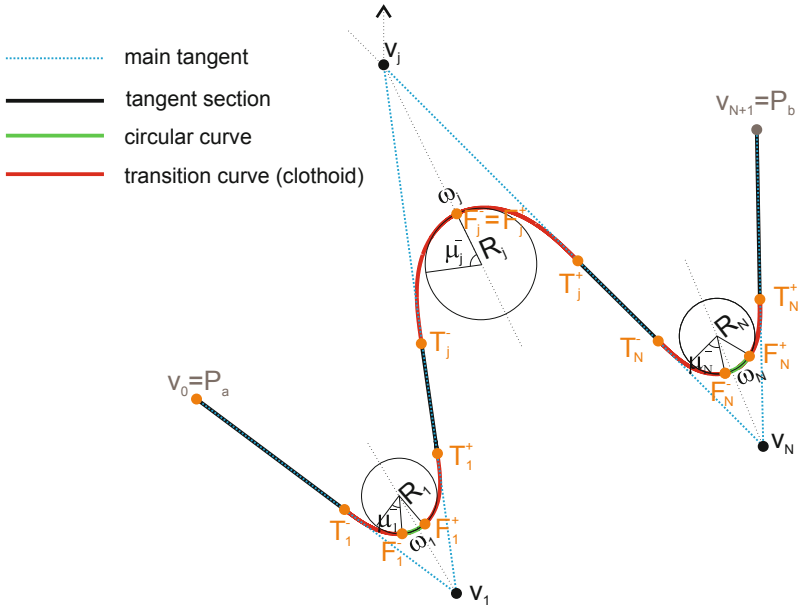


Fig. 2. Geometrical model of a non-symmetric HA between terminals \mathbf{P}_a and \mathbf{P}_b . In black the decision variables, in orange the tangency points between clothoids and tangents ($\mathbf{T}_j^-, \mathbf{T}_j^+$) and between the clothoids and the circular curve ($\mathbf{F}_j^-, \mathbf{F}_j^+$), respectively.

$$\mathbf{x}^N = (X_1, Y_1, R_1, \omega_1, \mu_1^-, X_2, Y_2, R_2, \omega_2, \mu_2^-, \dots, X_N, Y_N, R_N, \omega_N, \mu_N^-) \in \mathbb{R}^{5N}. \tag{1}$$

where μ^- is the angle bore by the ingoing clothoid. This decision variable \mathbf{x}^N provides the whole alignment, its key points (tangency points) and other relevant values throughout all the layout. In the following we show how to compute the tangency points and these other important magnitudes in the j -th curve (see Fig. 3). First of all we have the deflection angle which is

$$\theta_j = \arccos \left(\frac{\langle \mathbf{V}_j - \mathbf{V}_{j-1}, \mathbf{V}_{j+1} - \mathbf{V}_j \rangle}{V_j V_{j-1} \cdot V_{j+1} V_j} \right) \in [0, \pi],$$

with $\mathbf{V}_0 = \mathbf{P}_a$ and $\mathbf{V}_{N+1} = \mathbf{P}_b$, where, for sake of simplicity, we name the distance between two points \mathbf{A} and $\mathbf{B} \in \mathbb{R}^2$ by AB , i.e.

$$AB = \|\mathbf{B} - \mathbf{A}\|.$$

Also, from now on, until the end of this Subsection, the subindex j are omitted in all the equations with the exception of the vertices for a better reading. The angle of the outgoing clothoid is

$$\mu^+ = \theta - \omega - \mu^-.$$

The tangency points between tangents and the ingoing and outgoing clothoids are, respectively:

$$\mathbf{T}^- = \mathbf{V}_j - \frac{d^-}{V_j V_{j-1}} (\mathbf{V}_j - \mathbf{V}_{j-1}) \quad \text{and} \quad \mathbf{T}^+ = \mathbf{V}_j + \frac{d^+}{V_{j+1} V_j} (\mathbf{V}_{j+1} - \mathbf{V}_j), \quad (2)$$

where

$$d^- = \Delta x^- + \Delta y^- \tan(\mu^-) + H^- H^+ \frac{\sin(\alpha^+)}{\sin(\beta)}, \quad (3)$$

$$d^+ = \Delta x^+ + \Delta y^+ \tan(\mu^+) + H^- H^+ \frac{\sin(\alpha^-)}{\sin(\beta)}. \quad (4)$$

In equations (3)–(4) the values for distances and angles are given by:

- Expressions for the distances:

$$\begin{aligned} \Delta x^- &= \int_0^{2R\mu^-} \cos\left(\frac{s^2}{4R^2\mu^-}\right) ds, & \Delta y^- &= \int_0^{2R\mu^-} \sin\left(\frac{s^2}{4R^2\mu^-}\right) ds, \\ \Delta x^+ &= \int_0^{2R\mu^+} \cos\left(\frac{s^2}{4R^2\mu^+}\right) ds, & \Delta y^+ &= \int_0^{2R\mu^+} \sin\left(\frac{s^2}{4R^2\mu^+}\right) ds, \end{aligned}$$

and

$$H^- H^+ = \sqrt{(CH^-)^2 + (CH^+)^2 - 2CH^- CH^+ \cos(\omega)},$$

with

$$CH^- = R + \frac{\Delta y^-}{\cos(\mu^-)}, \quad CH^+ = R + \frac{\Delta y^+}{\cos(\mu^+)}.$$

- Expressions for the angles:

$$\beta = \pi - \theta, \quad \alpha^- = \frac{\pi}{2} + \mu^- - \gamma^-, \quad \alpha^+ = \frac{\pi}{2} + \mu^+ - \gamma^+,$$

where those values γ^- and γ^+ are computed as

$$\gamma^- = \begin{cases} \arcsin\left(\frac{CH^-}{H^- H^+} \sin(\omega)\right) & \text{if } CH^- \cos(\omega) \leq CH^+, \\ \pi - \arcsin\left(\frac{CH^-}{H^- H^+} \sin(\omega)\right) & \text{otherwise.} \end{cases}$$

and

$$\gamma^+ = \begin{cases} \arcsin\left(\frac{CH^+}{H^- H^+} \sin(\omega)\right), & \text{if } CH^+ \cos(\omega) \leq CH^-, \\ \pi - \arcsin\left(\frac{CH^+}{H^- H^+} \sin(\omega)\right) & \text{otherwise.} \end{cases}$$

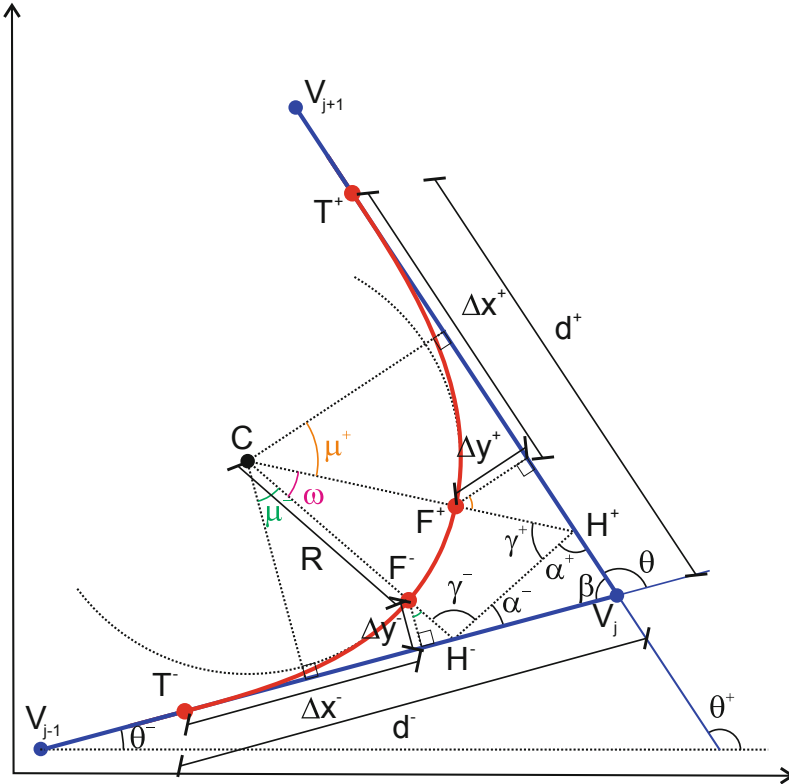


Fig. 3. Major elements in the j -th non symmetric transition curve. Only the subindex are shown in the vertex. In the other elements we avoid the use of subindex for sake of simplicity.

The tangency points between the clothoids and the circular curves are:

$$\mathbf{F}^- = \mathbf{H}^- + F^- H^- (\cos(\theta^- + \lambda^-(\alpha^- + \gamma^-)), \sin(\theta^- + \lambda^-(\alpha^- + \gamma^-))),$$

$$\mathbf{F}^+ = \mathbf{H}^+ + F^+ H^+ (\cos(\theta^+ + \lambda^+(\alpha^+ + \gamma^+)), \sin(\theta^+ + \lambda^+(\alpha^+ + \gamma^+))),$$

where

$$F^- H^- = \frac{\Delta y^-}{\cos(\mu^-)} \quad \text{and} \quad F^+ H^+ = \frac{\Delta y^+}{\cos(\mu^+)}.$$

The values $\theta^-, \theta^+ \in [0, 2\pi)$ are the the angles with respect to the positive part of the abscissa axis of the ingoing main tangent (going through \mathbf{V}_{j-1} and with direction $\mathbf{V}_j - \mathbf{V}_{j-1}$) and of the outgoing main tangent (going through \mathbf{V}_{j+1} and with direction $\mathbf{V}_{j+1} - \mathbf{V}_j$), respectively. In addition λ^- and λ^+ are the signs (-1 or 1) of the ingoing and outgoing clothoids, respectively, -1 if the clothoid goes clockwise 1 otherwise. Moreover

$$\mathbf{H}^- = \mathbf{V}_j - \frac{H^- V_j}{V_j V_{j-1}} (\mathbf{V}_j - \mathbf{V}_{j-1}) \quad \text{and} \quad \mathbf{H}^+ = \mathbf{V}_j + \frac{H^+ V_j}{V_{j+1} V_j} (\mathbf{V}_{j+1} - \mathbf{V}_j).$$

The last key point is the center of the circular curve which is given by:

$$\mathbf{C} = \mathbf{H}^- + CH^- (\cos(\theta^- + \lambda^-(\alpha^- + \gamma^-)), \sin(\theta^- + \lambda^-(\alpha^- + \gamma^-))). \quad (5)$$

From these key points, the arc length parametrization of the HA given by the decision vector \mathbf{x}^N can be computed in a similar way as proposed in [3]. We denote it by

$$\begin{aligned} \sigma_{\mathbf{x}^N} : [0, L(\mathbf{x}^N)] &\longmapsto \mathbb{R}^2 \\ s &\longmapsto \sigma_{\mathbf{x}^N}(s) = (\sigma_1(s), \sigma_2(s)), \end{aligned} \quad (6)$$

where $L(\mathbf{x}^N)$ the length of the track. In addition other important magnitudes can be obtained from the values of the tangency points such as:

- tangent lengths

$$L^{T_i} = T_{i-1}^+ T_i^-, \quad i = 1, \dots, N + 1,$$

where T_i names the i -th tangent, with $\mathbf{T}_0^+ = \mathbf{P}_a$ and $\mathbf{T}_{N+1}^- = \mathbf{P}_b$.

- circular curve lengths

$$L^{CC_j} = R_j \omega_j, \quad j = 1, \dots, N,$$

where CC_j names the j -th circular curve,

- clothoid lengths

$$L^{C_j^-} = 2R_j \mu_j^- \quad \text{and} \quad L^{C_j^+} = 2R_j \mu_j^+ \quad j = 1, \dots, N,$$

where C_j^- and C_j^+ names the j -th ingoing and outgoing transition curves, respectively

Remark 1. From now on, subindex i goes from $i = 1, \dots, N + 1$ and subindex j goes from $j = 1, \dots, N$. The first one makes reference to the number of tangents and the latter to the number of curves.

2.2 Constraints

In order to have a proper horizontal alignment the geometric model needs to fulfill certain conditions. Therefore, a series of constraints must be verified. In the following, a rigorous mathematical formulation of those constraints are stated.

2.2.1 Geometric Constraints In addition to the constraints previously stated ($R_j > 0, \omega_j \geq 0$) the decision variable has to fulfill the additional geometric constraints

- The angle of the ingoing clothoid and the circular curve ($\mu_j^- + \omega_j$) has to be smaller than the deflection angle at each curve, i.e.:

$$0 \leq \mu_j^- + \omega_j < \theta_j. \quad (7)$$

- A curve cannot begin before the preceding one ends. This is mathematically stated as:

$$\langle \mathbf{T}_i^- - \mathbf{T}_{i-1}^+, \mathbf{V}_i - \mathbf{V}_{i-1} \rangle \geq 0. \quad (8)$$

2.2.2 Technical Constraints The HA also has to fulfill other constraints given by the regulations of each country. These constraints are technical values to assure the safety as well as the comfort of the users. Usually, they set minimum values for the radii of circular curves R_j , and the lengths of tangents L^{T_i} , and clothoids $L^{C_j^-}, L^{C_j^+}$.

$$\begin{aligned}
 R_j &\geq R_{min}, \\
 L^{T_i}(\mathbf{x}^N) &\geq L^{T_{min}}, \\
 L^{C_j^-}(\mathbf{x}^N) &\geq L^{C_{min}}, \\
 L^{C_j^+}(\mathbf{x}^N) &\geq L^{C_{min}}.
 \end{aligned}
 \tag{9}$$

All railways are designed by civil engineers as HA complying with the technical restrictions set by the authorities. Those alignments are called admissible horizontal alignments (AHA).

2.3 Objectives and Optimization Problems

Railway lines, due to its use, over time, suffer fluctuations and what once was an AHA now is not anymore. To see the size those fluctuations, the center of the railway is measured with a set of points by track surveying (usually given by UTM coordinates). These track deviations must be amended periodically to guarantee the safety of train convoys as well as the comfort of passengers. The original HA is not rebuilt, what it is usually done is to obtain a new AHA as close as possible to those measured points from the deformed railway. This technique is called *recreation*. In this section a two stage algorithm trying to recreate a railway is stated obtaining a new AHA as close as possible to those points given by the center line of the deviated railway. Let us call

$$\{\mathbf{P}_1, \dots, \mathbf{P}_{N_{SP}}\},$$

to the set of points measured along the center line of the railway by track surveying $\mathbf{P}_l = (x_l, y_l)$ with $\mathbf{P}_1 = \mathbf{P}_a = (x_a, y_a)$ and $\mathbf{P}_{N_{SP}} = \mathbf{P}_b = (x_b, y_b)$, i.e. the initial and final measured points correspond to the terminals. Those terminal values must be elected within tangents placed at suitable locations; for example over a bridge, in stations or in halts. Also, the following hypothesis must be hold:

HP1 Backtracking is not allowed in the surveyed points, i.e.

$$\langle \mathbf{P}_b - \mathbf{P}_a, \mathbf{P}_l - \mathbf{P}_{l-1} \rangle > 0, \quad l = 2, \dots, N_{SP}.$$

HP2 Backtracking is also not allowed in the recreated alignment and all the vertices must be inside the $[x_a, x_b]$ strip, specifically

$$x_a \leq X_l \leq X_{l+1} \leq x_b, \quad l = 1, \dots, N - 1.$$

HP3 The initial and end points \mathbf{P}_a and \mathbf{P}_b , respectively, have the same y -coordinate:

$$y_a = y_b.$$

This is not an strong hypothesis since if this does not hold true, like in north to south alignments for example, a rotation can to be done and gyrate tangent joining the terminals into the abscissa axis as a previous step. The same rotation must also be done to all surveyed points.

Let us explain in detail both optimization problems.

2.3.1 Optimizing the y-Distances

In this first optimization problem (see [4,21]), a function f_{pch} is constructed by applying a piecewise cubic Hermite interpolating polynomial (monotonic cubic spline) along to all the available measured points. This function is well defined due to the hypothesis (HP1 to HP3) and provides for any coordinate $x \in [x_a, x_b]$ its corresponding coordinate y , allowing to easily compute the y -value on every point of the recreated alignment, not only on the surveyed data.

The next step is to define an objective function for evaluating the quality of the fitting between the recreated alignment and the surveyed data. In this case, the function $J_1(\mathbf{x}^N)$ given by the following equation is proposed.

$$J_1(\mathbf{x}^N) = \int_0^{L(\mathbf{x}^N)} (\sigma_2(s) - f_{pch}(\sigma_1(s)))^2 \sigma_1'(s) ds.$$

The optimization problem will consist on minimizing the function $J_1(\mathbf{x}^N)$ considering the geometric and technical constraints over the decision variable \mathbf{x}^N . Therefore, the optimization problem can be formulated as follows

$$\begin{aligned} \min_{\mathbf{x}^N \in \mathbb{R}^{5N}} \quad & J_1(\mathbf{x}^N). \\ \text{subject to} \quad & (7)-(9) \end{aligned} \tag{10}$$

This first step optimization problem can be seen as a sort of least square HA fitting.

2.3.2 Optimizing the Orthogonal Distance

In this second optimization problem we seek to minimize the orthogonal distance from an alignment to a set of points $\{\mathbf{P}_1, \dots, \mathbf{P}_{N_{SP}}\}$. The difficulty here, lies in identify with which section of the layout (tangents, clothoids or circular curves) the given points (measured by track surveying) intersect orthogonally with the alignment. We call *bind* to this process of finding the section(s) of the alignment with which to calculate the orthogonal distance. Without losing generality a direction is fixed, for example from starting to end point. Let us call $s_{T_i^+}$ the parametric value which gives the tangency point \mathbf{T}_i^+ when the parametrization of the layout is evaluated, i.e.

$$\sigma_{\mathbf{x}^N}(s_{T_i^+}) = \mathbf{T}_i^+,$$

and the same type of notation for the other tangency points.

Remark 2. The parametrization of $\sigma_{\mathbf{x}^N}$ consists of piecewise parametrization of all $4N + 1$ the sections of the HA. Those $4N + 1$ elements are tangents, ingoing clothoids, circular curves and outgoing clothoids; four sections for each curve and an additional tangent at the end.

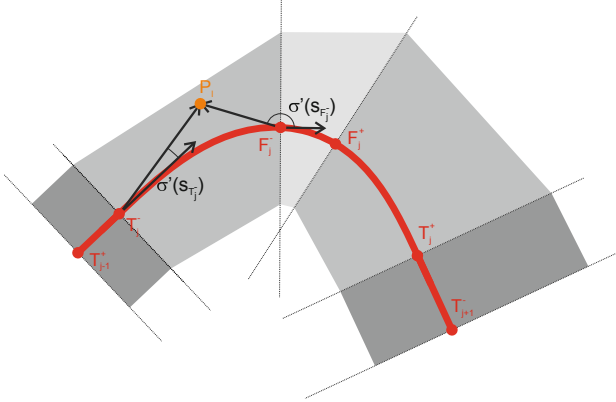


Fig. 4. Bind condition of point \mathbf{P}_l to section C_j^- . The angles between $\sigma'_{\mathbf{x}^N}(s_{T_j^-})$ and $\mathbf{P}_l - \mathbf{T}_j^-$ and between $\sigma'_{\mathbf{x}^N}(s_{F_j^-})$ and $\mathbf{P}_l - \mathbf{F}_j^-$ must change from acute to obtuse at both ends (condition (12)).

Let us see when a point \mathbf{P} is bind to a section of the HA and how to compute its orthogonal distance:

- If the section is the i -th tangent, i.e. $T_i = \{\sigma_{\mathbf{x}^N}(s) / s \in [s_{T_{i-1}^+}, s_{T_i^-}]\}$, a point \mathbf{P} is bind to i -th tangent if

$$\text{sign} \langle \sigma'_{\mathbf{x}^N}(s_{T_{i-1}^+}), \mathbf{P} - \mathbf{T}_{i-1}^+ \rangle \neq \text{sign} \langle \sigma'_{\mathbf{x}^N}(s_{T_i^-}), \mathbf{P} - \mathbf{T}_i^- \rangle. \quad (11)$$

The tangent equation ($a_i x + b_i y + c_i = 0$) can be easily computed by its end points \mathbf{T}_{i-1}^+ and \mathbf{T}_i^- and the distance to point $\mathbf{P} = (x, y)$ is

$$d^{T_i} = \text{dist}(\mathbf{P}, T_i) = \begin{cases} \left| \frac{a_i x + b_i y + c_i}{\sqrt{a_i^2 + b_i^2}} \right| & \text{if (11) holds} \\ \infty & \text{otherwise.} \end{cases}$$

- If the section is the j -th ingoing clothoid, i.e. $C_j^- = \{\sigma_{\mathbf{x}^N}(s) / s \in [s_{T_j^-}, s_{F_j^-}]\}$, a point \mathbf{P} is bind to the clothoid if

$$\text{sign} \langle \sigma'_{\mathbf{x}^N}(s_{T_j^-}), \mathbf{P} - \mathbf{T}_j^- \rangle \neq \text{sign} \langle \sigma'_{\mathbf{x}^N}(s_{F_j^-}), \mathbf{P} - \mathbf{F}_j^- \rangle. \quad (12)$$

Then

$$d^{C_j^-} = \text{dist}(\mathbf{P}, C_j^-) = \begin{cases} \|\mathbf{P} - \mathbf{Q}\| & \text{if (12) holds} \\ \infty & \text{otherwise.} \end{cases}$$

where $\mathbf{Q} = \sigma_{C_j^-}(\tilde{s})$ is the point of C_j^- with minimum distance to \mathbf{P} . This parametric value \tilde{s} , which provides \mathbf{Q} , can be computed by solving with, for example, Newton-Raphson, the non linear equation given by this dot product

$$\langle \sigma'_{C_j^-}(s), \sigma_{C_j^-}(s) - \mathbf{P} \rangle = 0, \quad s \in [0, L^{C_j^-}]$$

where $\sigma_{C_j^-}$ is the local parametrization of C_j^- given by

$$\sigma_{C_j^-}(s) = \mathbf{T}_j^- + \left(\int_0^s \cos \left(\frac{\lambda_j^- \tau^2}{4R_j^2 \mu_j^-} + \theta_j^- \right) d\tau, \int_0^s \sin \left(\frac{\lambda_j^- \tau^2}{4R_j^2 \mu_j^-} + \theta_j^- \right) d\tau \right).$$

Remark 3. Function $\sigma_{C_j^-}$ properly shifted provides the piecewise parametrization of C_j^- in $\sigma_{\mathbf{x}^N}$,

$$\sigma_{\mathbf{x}^N} \left(\sum_{k=1}^{j-1} [L^{T_k} + L^{C_k^-} + L^{CC_k} + L^{C_k^+}] + L^{T_j} + s \right) = \sigma_{C_j^-}(s)$$

- If the section is the j -th circular curve, i.e. $CC_j = \{\sigma_{\mathbf{x}^N}(s) / s \in [s_{F_j^-}, s_{F_j^+}]\}$, a point \mathbf{P} is bind to the circular curve if

$$\text{sign} \langle \sigma'_{\mathbf{x}^N}(s_{F_j^-}), \mathbf{P} - \mathbf{F}_j^- \rangle \neq \text{sign} \langle \sigma'_{\mathbf{x}^N}(s_{F_j^+}), \mathbf{P} - \mathbf{F}_j^+ \rangle. \quad (13)$$

Its key values, \mathbf{C}_j (center) and R_j (radius) are known by (5) and (1), respectively, and the distance of a circular curve to a point is obtained by straightforward computations as

$$d^{CC_j} = \text{dist}(\mathbf{P}, CC_j) = \begin{cases} |\|\mathbf{P} - \mathbf{C}_j\| - R_j| & \text{if (13) holds} \\ \infty & \text{otherwise.} \end{cases}$$

- If the section is the j -th outgoing clothoid, i.e. $C_j^+ = \{\sigma_{\mathbf{x}^N}(s) / s \in [s_{F_j^+}, s_{T_j^+}]\}$, a point \mathbf{P} is bind to the clothoid if

$$\text{sign} \langle \sigma'_{\mathbf{x}^N}(s_{F_j^+}), \mathbf{P} - \mathbf{F}_j^+ \rangle \neq \text{sign} \langle \sigma'_{\mathbf{x}^N}(s_{T_j^+}), \mathbf{P} - \mathbf{T}_j^+ \rangle, \quad (14)$$

When this condition holds true, then

$$d^{C_j^+} = \text{dist}(\mathbf{P}, C_j^+) = \begin{cases} \|\mathbf{P} - \mathbf{Q}\| & \text{if (14) holds} \\ \infty & \text{otherwise.} \end{cases}$$

where $\mathbf{Q} = \sigma_{C_j^+}(\tilde{s})$ is the point in C_j^+ with minimum distance to \mathbf{P} . This parametric value \tilde{s} , which provides \mathbf{Q} , can be computed by solving with, for example, Newton-Raphson the non linear equation given by this dot product

$$\langle \sigma'_{C_j^+}(s), \sigma_{C_j^+}(s) - \mathbf{P} \rangle = 0, \quad s \in [0, L^{C_j^+}],$$

where $\sigma_{C_j^+}$ is the local parametrization of C_j^+ given by

$$\sigma_{C_j^+}(s) = \mathbf{T}_j^+ + \left(\int_0^s \cos \left(\frac{\lambda_j^+ \tau^2}{4R_j^2 \mu_j^+} + \theta_j^+ \right) d\tau, \int_0^s \sin \left(\frac{\lambda_j^+ \tau^2}{4R_j^2 \mu_j^+} + \theta_j^+ \right) d\tau \right).$$

Remark 4. Function $\sigma_{C_j^+}$ adequately shifted provides the piecewise parametrization of C_j^+ in $\sigma_{\mathbf{x}^N}$:

$$\sigma_{\mathbf{x}^N} \left(\sum_{k=1}^j \left[L^{T_k} + L^{C_k^-} + L^{C_k} + L^{C_k^+} \right] - s \right) = \sigma_{C_j^+}(s) \tag{15}$$

The parametrization of the outgoing clothoid it is computed backwards from \mathbf{T}_j^+ to \mathbf{F}_j^+ . Therefore, the parameter in the left hand side of (15) is flipped. In addition $\lambda_j^+ = -\lambda_j^-$ as they have opposite orientations.

After this bind process, every point \mathbf{P}_l has a distance to all sections of the alignment (finite value if is bind, infinite if not). Finally a point is only bind to the section with which has a minimum distance

$$d_l = \min \left\{ d_l^{T_1}, d_l^{C_1^-}, d_l^{C_1}, \dots, d_l^{C_N^+}, d_l^{T_{N+1}} \right\}.$$

The problem of optimizing orthogonal distances can be defined by the following objective function

$$J_2(\mathbf{x}^N) = \sum_{l=1}^{N_{SP}} d_l.$$

The optimization problem will consist on minimizing the function $J_2(\mathbf{x}^N)$ considering some constrains over the decision variable \mathbf{x}^N related to the definition of the geometrical model or some technical restrictions that must be taken into account. Therefore, this optimization problem can be formulated as follows

$$\begin{aligned} & \min_{\mathbf{x}^N \in \mathbb{R}^{5N}} J_2(\mathbf{x}^N). \tag{16} \\ & \text{subject to (7)–(9)} \end{aligned}$$

Remark 5. It should be noted about this of orthogonal distances optimization problem that:

- To do a properly bind and have a well posed problem (16), the initial alignment must be close enough to the set of points (all d_l must be finite). This is the main reason for creating as a first step the y -distances optimization (10).
- It is not necessary to previously identify the geometrical elements (tangents, circular curves a clothoids) of the existing alignment. The algorithm automatically binds all point to its closest element.

2.4 Proposed Algorithm

The proposed algorithm is a two-step method and consists of two consecutive optimization problems. In a first step, an AHA is computed to minimize a sort of y -distance to a cubic spline (Piecewise Cubic Hermite Interpolating Polynomial (PCHIP)) created with the N_{SP} points obtained in the deformed track centerline. The aim in this first step is to obtain a relatively close HA to the surveyed data to be used as initial point into the second optimization problem. The first step will give, with an easy and cheap cost function, a rough approximation of the HA. The second step consists of a more precise approach in which we try to obtain an AHA that minimize the sum of the orthogonal distances from the surveyed points. The proposed method for recreation can be summarized in the following two steps algorithm.

Input: Surveyed points $\{\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_{N_{SP}}\}$, $\mathbf{P}_a = \mathbf{P}_1, \mathbf{P}_b = \mathbf{P}_{N_{SP}}$, number of curves N , technical constraints $R_{min}, L^{T_{min}} \dots$, number of multi-starts M_S , threshold to go to step 2, J_{max} .

Output: Decision variable x_{sol} , Tangency points $\mathbf{T}_j^-, \mathbf{T}_j^+, \mathbf{F}_j^-, \mathbf{F}_j^+$ lengths of all alignment components L^{T_i}, L^{C_j}, \dots , mean vertical distance, mean orthogonal distance, max orthogonal distance

```

1 for  $k = 1 \dots M_S$  do
2   Create  $x_k^{inic}$  with the random AHA generator algorithm given in [20] ;
3   Step 1. Compute  $x_k^{Stp1}$  by solving (10) with starting point  $x_k^{inic}$  ;
4   if Step 1 converged to feasible solution .and.  $J_1(x_k^{Stp1}) < J_{max}$  then
5     | Step 2. Compute  $x_k^{Stp2}$  by solving (16) with starting point  $x_k^{Stp1}$ ;
6     | end
7   end
8  $x_{sol} = \min\{x_k^{Stp2}\}$  of all feasible solutions of Step. 2;

```

Algorithm 1: Algorithm for AHA recreation of railway alignments

3 Numerical Experiments. Checking the Algorithm

Two academic examples are conducted trying to show the good behavior of the presented model. We propose a two stage algorithm for automatically recreate railway sections from a given data $\{\mathbf{P}_1, \dots, \mathbf{P}_{N_{SP}}\}$ obtained from track surveying. In addition, technical constraints were provided to fulfill the operational requirements of the railway. In both experiments the following values were chosen $R_{min} = 100$ m., $L_{min}^T = 0$ m. and $L_{min}^{C^-} = L_{min}^{C^+} = 100$ m. and a threshold to go from step 1 to 2 of $J_{max} = 1$. In both experiments $M_S = 3000$ seeds were used provided by the algorithm proposed in [20] to solve NLP-problem (10). The solution of this first step it will be the the starting point to solve problem (16).

3.1 Experiment 1: Recovering a HA

In this first example we try to show the good behavior of the model. A known alignment of approx. 5.5km long is given, i.e. all the elements of the decision variable (1) are known. Their values can be seen in the column “Original” of Table 1. The layout it is formed by 3 curves with non symmetric clothoids at each side of the circular curves. From this HA we extract 378 points which are the equivalent of the measurements made by track surveying. The Algorithm shown in Sect. 2.4 is a two-step method. In the first step, the least square y -distance alignment to the PCHIP built with the 378 points is computed. The first optimization problem (10) is solved with a random multi-start and the sequential quadratic programming (SQP) method (see [14]). The seeds needed to start the SQP numerical resolution are obtained by using the algorithm for automatic generation of random admissible horizontal alignments described in [20]. In this first step the aim is to create alignments (using a very simple cost function) close to the points to easily bind them to the alignment sections in the second step.

In the second step the sum of the orthogonal distances to the measured given points is minimized (solve Problem (16)). The second step it only takes place if there was a satisfactory solution in the first step, i.e. if SQP goes to a feasible solution and if the value of the functional (10) is lower than a given threshold. To optimize this second step, the sequential quadratic programming (SQP) method is used once again.

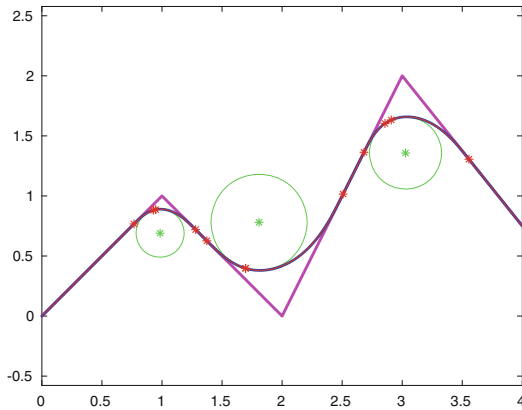


Fig. 5. Track surveying data, recreated alignment, IP’s and tangency points T^-, T^+, F^- and F^+ .

The results of this test can be seen in Fig. 5 and in the Table 1 in which the vertices of the original layout, the radii and the angles of the circular curves and ingoing clothoids are given. In the column “Recreated” the same values of the recreated alignment are given. The alignment is completely recovered from those 378 points. In addition extra data can be recovered from the decision variable.

This other data can be seen in Table 2. Furthermore, the algorithm can also provide the mean value of the orthogonal distance. In this first test problem this magnitude is 0.007537 cm and the maximum orthogonal distance and the alignment section in which it was found. In this first test the max. orthogonal distance is 0.076662 cm. and it was obtained in the 3rd section (CC_1) of the HA.

To get this solution a MATLAB (version 9.4 (R2018a)) code to solve (10) and (16) was implemented. This code was executed on a cluster with 8 PowerEdge R840 with Intel(R) Xeon(R) Gold 6126 processors, 12 cores per processor and 384GB of RAM. MATLAB proprietary parallelization was used with a number of 24 threads. The computation (wall clock) times for this first experiment was 9.070e+4 s.

Table 1. Original and recreated values of the decision variable data.

Vertices (IPs) [Km]			Radii [m]		
	Original	Recreated		Original	Recreated
V_1	(1,1)	(1.000000, 1.000000)	R_1	200	199.934965
V_2	(2,0)	(2.000000, 0.000000)	R_2	400	399.999782
V_3	(3,2)	(3.000000, 2.000000)	R_3	300	299.970329
Angles of circular curves [rad]			Angles of ingoing clothoid [rad]		
	Original	Recreated		Original	Recreated
ω_1	0.1	0.098969	μ_1^-	0.5	0.500375
ω_2	0	0.000000	μ_2^-	0.5	0.499998
ω_3	0.2	0.199585	μ_3^-	0.5	0.500150

Table 2. Length in meters of all the alignment sections (original and recreated). Column “A(m)” gives the clothoid parameter $\sqrt{R \cdot L}$.

No	Type	Original					Recreated				
		Inic. K.P	Final K.P	L(m)	R(m)	A(m)	Inic. K.P	Final K.P	L(m)	R(m)	A(m)
1	T_1	0+000.0	1+084.7	1084.7	-	-	0 + 000.0	1 + 084.7	1084.7	-	-
2	C_1^-	1 + 084.7	1 + 284.7	200.0	-	200.0	1 + 084.7	1 + 284.8	200.1	-	200.0
3	CC_1	1 + 284.7	1 + 304.7	20.0	200.0	-	1 + 284.8	1 + 304.6	19.8	199.9	-
4	C_1^+	1 + 304.7	1 + 693.0	388.3	-	278.6	1 + 304.6	1 + 693.1	388.5	-	278.7
5	T_2	1 + 693.0	1 + 824.7	131.7	-	-	1 + 693.1	1 + 824.8	131.7	-	-
6	C_2^-	1 + 824.7	2 + 224.7	400.0	-	400.0	1 + 824.8	2 + 224.8	400.0	-	400.0
7	CC_2	2 + 224.7	2 + 224.7	0.0	400.0	-	2 + 224.8	2 + 224.8	0.0	400.0	-
8	C_2^+	2 + 224.7	3 + 338.7	1114.0	-	667.5	2 + 224.8	3 + 338.8	1114.0	-	667.5
9	T_3	3 + 338.7	3 + 727.9	389.2	-	-	3 + 338.8	3 + 728.0	389.2	-	-
10	C_3^-	3 + 727.9	4 + 027.9	300.0	-	300.0	3 + 728.0	4 + 028.1	300.1	-	300.1
11	CC_3	4 + 027.9	4 + 087.9	60.0	300.0	-	4 + 028.1	4 + 088.0	59.9	300.0	-
12	C_3^+	4 + 087.9	4 + 869.8	781.9	-	484.3	4 + 088.0	4 + 870.0	782.0	-	484.4
13	T_4	4 + 869.8	5 + 581.3	711.5	-	-	4 + 870.0	5 + 581.5	711.5	-	-

In this test from the $M_s = 3000$ initial seeds in step 1, only 2064 reached a feasible valid solution for (10). From those 2064 only 1132 were valid feasible solutions for (16).

3.2 Experiment 2: Recreating a Deviated HA

In this second example a case is presented trying to verify the good behavior of the Algorithm with a deviated HA. The same way as it was done in Sect. 3.1 a known alignment grater than 4 km long is given. The layout is formed by 4 curves with non symmetric clothoids at each side of the circular curves. A set of 329 points were extracted from the original layout. These points were orthogonally shifted from the original alignment a random value between [-10 cm,10 cm] trying to emulate the displacements suffered by the railway track due to the passage of trains over time. The new alignment should move the track as little as possible from its deformed state. In this experiment the aim is to find this new AHA with the methodology presented in Sect. 2. The most relevant data of the alignment can be seen in the “Original” column of Table 3. The methodology and stages are the same as in Subsect. 3.2. A first optimization step with least squares to solve (10) and create HA close enough to the points and those solutions are used as as initial value for the second of the steps (solve (16)) to easily bind the points to each alignment section. Numerical results were obtained with the same MATLAB code as Experiment 1. That code was also run in the same cluster.

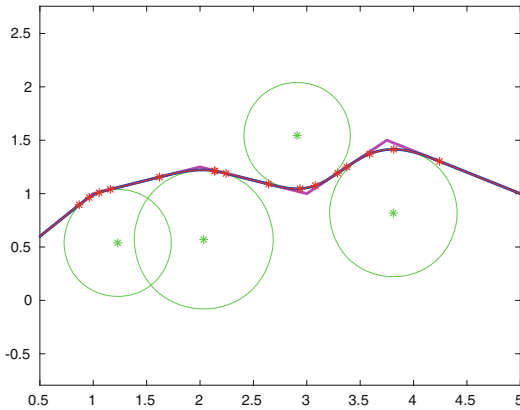


Fig. 6. Track surveying data, recreated alignment, IP’s and tangency points T^- and T^+ and F^- and F^+ .

The results of this test can be seen in Fig. 6 and in Table 3 in which the vertices of the original layout, the radii and the lengths of each alignment sections are given. In the column “Recreated” the same values of the recreated alignment are given. A new alignment is obtained from those 329 points. From the decision variable we can extract other important data which is shown in Table 4. The

Table 3. Original and recreated values of the decision variable data.

Vertices [Km]			Radii [m]		
	Original	Recreated		Original	Recreated
V_1	(1,1)	(0.999911, 0.999959)	R_1	500	502.360420
V_2	(2,1.25)	(1.999847, 1.249979)	R_2	650	651.049244
V_3	(3,1)	(3.000327, 0.999999)	R_3	500	498.184734
V_4	(3.75,1.5)	(3.749879, 1.500099)	R_4	600	597.251374
Angles of circular curves [rad]			Angles of ingoing clothoid [rad]		
	Original	Recreated		Original	Recreated
ω_1	0.2	0.204833	μ_1^-	0.12	0.116253
ω_2	0	0.007118	μ_2^-	0.4	0.400905
ω_3	0.3	0.291977	μ_3^-	0.3	0.300816
ω_4	0.4	0.387914	μ_4^-	0.2	0.207484

Table 4. Length in meters of all the alignment sections (original and recreated).

No	Typo	Original					Recreated				
		Inic. K.P	Final K.P	L(m)	R(m)	A(m)	Inic. K.P	Final K.P	L(m)	R(m)	A(m)
1	T_1	0 + 000,0	0 + 471,4	471.4	-	-	0 + 000,0	0 + 472,3	472.3	-	-
2	C_1^-	0 + 471,4	0 + 591,4	120.0	-	244.9	0 + 472,3	0 + 589,1	116.8	-	242.24
3	CC_1	0 + 591,4	0 + 691,4	100.0	500,0	-	0+589,1	0 + 692,0	102.9	502,4	-
4	C_1^+	0 + 691,4	0 + 801,2	109.8	-	234.3	0 + 692,0	0 + 801,2	109.2	-	234.2
5	T_2	0 + 801,2	1 + 276,6	475.4	-	-	0 + 801,2	1 + 275,8	474.6	-	-
6	C_2^-	1 + 276,6	1 + 796,6	520.0	-	581.4	1 + 275,8	1 + 797,8	522.0	-	582.9
7	CC_2	1 + 796,6	1 + 796,6	0.0	650,0	-	1 + 797,8	1 + 802,4	4.6	651,0	-
8	C_2^+	1 + 796,6	1 + 913,5	116.9	-	275.7	1 + 802,4	1 + 909,0	106.6	-	263.4
9	T_3	1 + 913,5	2 + 318,3	404.8	-	-	1 + 909,0	2 + 319,2	410.2	-	-
10	C_3^-	2 + 318,3	2 + 618,3	300.0	-	387.2	2 + 319,2	2 + 618,9	299.7	-	386.4
11	CC_3	2 + 618,3	2 + 768,3	150.0	500,0	-	2 + 618,9	2 + 764,4	145.5	498,2	-
12	C_3^+	2 + 768,3	3 + 001,3	233.0	-	341.3	2 + 764,4	3 + 004,0	239.6	-	345.5
13	T_4	3 + 001,3	3 + 110,1	108.8	-	-	3 + 004,0	3 + 107,4	103.4	-	-
14	C_4^-	3 + 110,1	3 + 350,1	240.0	-	379.4	3 + 107,4	3 + 355,2	247.8	-	384.7
15	CC_4	3 + 350,1	3 + 590,1	240.0	600,0	-	3 + 355,2	3 + 586,9	231.7	597,3	-
16	C_4^+	3 + 590,1	4 + 032,3	442.2	-	515.1	3 + 586,9	4 + 033,1	446.2	-	516.3
17	T_5	4 + 032,3	4 + 847,3	815.0	-	-	4 + 033,1	4 + 847,5	814.4	-	-

mean value of the orthogonal distance of the recreated HA is 4.839518 cm and the maximum orthogonal distance is 14.690249 cm. The worst measurement was obtained in the 14th section (C_4^-) of the HA. In contrast, The mean value of the orthogonal distance of the original HA is 5.126904 cm and the maximum orthogonal distance is 9.999455 cm. The worst measurement was obtained in the 16th section of the HA.

It should be noted that the mean value of the orthogonal distances of the recreated HA is decreased by a 5.61% with respect to the original, i.e. the new

alignment with this two stage method fits the surveyed data better than the original alignment.

The computational (wall clock) time for this experiment was: $7.0775e + 4$ s. In this test from the $M_s = 3000$ initial seeds in step 1, only 2861 reached a feasible valid solution for (10). From those 2861 only 978 were valid feasible solutions for (16). It is also worth to point out that the minimum value was reached around 14 times, which means that 3000 seeds may be an oversize value. Good HA might probably also be obtained in less time by properly adjusting the number of initial number of seeds.

4 Conclusions and Future Work

A new mathematical approach of non-symmetric HA is stated and the expressions for its tangency points are also given. Based on this new formulation, this new HA model is used to recreate a railway layout from a set of points fulfilling all technical constraints. This recreation is carried out by means of a two stage optimization problem. In the first stage a least square HA is obtained by minimizing the y -distance, while in the second stage a more accurate optimization problem is used to compute a HA that minimizes the orthogonal distance to the surveyed points. It is worth to point out that there is no need to prior identification of the existing alignment by means of its geometrical elements. The only additional information needed is to know the number of curves N that the recreation section has. Although the number of curves can also be deduced with an exhaustive search varying the number of curves N in the first optimization step, minimizing the y -distance with the algorithm proposed in [20]. The two stage algorithm was tested against two academic examples showing the good behavior of the model. From what has been seen in this work it will be very useful to upgrade it in the following aspects:

1. To completeness the good behavior of the method, this should be tested with a real case study and compare the obtained solutions with a recreation given by a civil engineer.
2. It has been notice during the development of the numerical experiments that working with this high degree of precision (cm) in a section of several km, the computation of the optimal solution might affected by accuracy with which the SQP computes the numerical gradient, compromising sometimes the convergence of the algorithm.

These two aspects will be studied in a forthcoming work.

Acknowledgments. This research was funded by Ministerio de Ciencia e Innovación (Spain) grant number TED2021-129324B-I00, and by the collaboration agreement between Xunta de Galicia (Spain) and Universidade de Santiago de Compostela (Spain) which regulates the Specialization Campus “Campus Terra”. Finally, second authors thanks the support given by Xunta de Galicia (Spain) under research projects ref. ED341D R2016/023 and third author has received financial support from the Xunta de Galicia (2021 GRC GI-1563-ED431C2021/15).

References

1. Ben-Arieh, D., Chang, S., Rys, M., Zhang, G.: Geometric modeling of highways using global positioning system data and B-spline approximation, *J. Transp. Eng.*, **130**(5), 632–636 (2004) [https://doi.org/10.1061/\(ASCE\)0733-947X\(2004\)130:5\(632\)](https://doi.org/10.1061/(ASCE)0733-947X(2004)130:5(632))
2. Camacho-Torregrosa, F. J., Pérez-Zuriaga, A. M., Campoy-Ungría, J. M., García, A., Tarko, A. P.: Use of heading direction for recreating the horizontal alignment of an existing road. *Comput.-Aided Civ. Inf.*, **30**(4), 282–299 (2015) <https://doi.org/10.1111/mice.12094>
3. Casal, G., Santamarina, D., Vázquez-Méndez, M.E.: Optimization of horizontal alignment geometry in road design and reconstruction. *Transpot. Res. C-Emer.* **74**, 261–274 (2017) <https://doi.org/10.1016/j.trc.2016.11.019>
4. Castro, A., Casal, G., Santamarina, D., Vázquez-Méndez, M.E.: A simple method for automatic recreation of railway horizontal alignments. Manuscript submitted for publication.
5. Castro, M., Iglesias, L., Rodríguez-Solano, R., Sánchez, J. A.: Geometric modeling of highways using global positioning system (GPS) data and spline approximation, *Transport. Res.C-Emer.* **14**(4), 233–243 (2006) <https://doi.org/10.1016/j.trc.2006.06.004>
6. Dong, H., Easa, S., Li, J.: Approximate extraction of spiralled horizontal curves from satellite imagery, *J. Surv. Eng.*, **133**(1), 36–40 (2007) [https://doi.org/10.1061/\(ASCE\)0733-9453\(2007\)133:1\(36\)](https://doi.org/10.1061/(ASCE)0733-9453(2007)133:1(36))
7. Easa, S., Dong, H., Li, J.: Use of satellite imagery for establishing road horizontal alignments, *J. Surv. Eng.*, **133**(1), 29–35 (2007) [https://doi.org/10.1061/\(ASCE\)0733-9453\(2007\)133:1\(29\)](https://doi.org/10.1061/(ASCE)0733-9453(2007)133:1(29))
8. Hans, Z., Souleyrette, R., Bogenreif, C.: *Horizontal Curve Identification and Evaluation*, Iowa State University, Ames, IA. (2012)
9. Hummer, J. E., Rasdorf, W. J., Findley, D. J., Zegeer, C. V., Sundstrom, C. A.: *Procedure for Curve Warning Signing, Delineation, and Advisory Speeds for Horizontal Curves*, Technical Report, FHWA/NC/2009-07, Department of Civil, Construction and Environmental Engineering, North Carolina (2010)
10. Imran, M., Hassan, Y., Patterson, D.: GPS-GIS-based procedure for tracking vehicle path on horizontal alignments, *Comput.-Aided Civ. Inf.*, **21** 382–394 (2006) <https://doi.org/10.1111/j.1467-8667.2006.00444.x>
11. Li, W., Pu, H., Schonfeld, P., Song, Z., Zhang, H., Wang, L., Wang, J., Peng, X., Peng, L.: A method for automatically recreating the horizontal alignment geometry of existing railways, *Comput.-Aided Civ. Inf.*, **34**, 71–94 (2018) <https://doi.org/10.1111/mice.12392>
12. Li, W., Zhen, S., Schonfeld, P., Pu, H., Zhang, Z., Wang, L., Zhao, L., Qui, X., Wei, F., Yan, W.: Recreating Existing Railway Horizontal Alignments Automatically Using Overall Swing Iteration, *J. Transp. Eng. A-Syst.*, **148**(8) (2022) <https://doi.org/10.1061/JTEPBS.0000691>
13. Li, Z., Chitturi, M. V., Bill, A. R., Noyce, D. A.: Automated identification and extraction of horizontal curve information from geographic information system roadway maps, *Transport. Res. Rec.*, **2291**, 80–92 (2012) <https://doi.org/10.3141/2291-10>
14. Nocedal, J., Wright, S.J.: *Numerical Optimization*, Springer Series in Operations Research and Financial Engineering, Springer Science+Business Media, New York (2006)

15. Othman, S., Thomson, R., Lannér, G.: Using naturalistic field operational test data to identify horizontal curves, *J. Transp. Eng.*, **138**(9), 1151–1160 (2012) [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000408](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000408)
16. Tsai, Y., Wu, J., Wang, Z., Hu, Z.: Horizontal roadway curvature computation algorithm using vision technology, *Comput.-Aided Civ. Inf.*, **25**, 78–88 (2010) <https://doi.org/10.1111/j.1467-8667.2009.00622.x>
17. Vázquez-Méndez, M.E., Casal, G.: The clothoid computation: a simple and efficient numerical algorithm. *J. Surv. Eng.* **142**(3), 04016005 (2016) [https://doi.org/10.1061/\(ASCE\)SU.1943-5428.0000177](https://doi.org/10.1061/(ASCE)SU.1943-5428.0000177)
18. Vázquez-Méndez, M.E., Casal, G., Castro, A., Santamarina, D.: A 3D model for optimizing infrastructure costs in road design. *Comput.-Aided Civ. Inf.* **33**, 423–439 (2018) <https://doi.org/10.1111/mice.12350>
19. Vázquez-Méndez, M.E., Casal, G., Castro, A., Santamarina, D.: Optimization of an urban railway bypass. A case study in A Coruña-Lugo line, Northwest of Spain, *Comput. Ind. Eng.*, **151**, 106935 (2021) <https://doi.org/10.1016/j.cie.2020.106935>
20. Vázquez-Méndez, M.E., Casal, G., Castro, A., Santamarina, D.: An algorithm for random generation of admissible horizontal alignments for optimum layout design, *Comput.-Aided Civ. Inf.*, **36**, 1056–1072 (2021) <https://doi.org/10.1111/mice.12682>
21. Vázquez-Méndez, M.E., Casal, G., Castro, A., Santamarina, D.: An automatic method for generating multiple alignment alternatives for a railway bypass, *Comput. Oper. Res.*, **154**, 106217 (2023) <https://doi.org/10.1016/j.cor.2023.106217>



Analytical Estimation of Natural Frequencies of Offshore Monopile Wind Turbines

Hadi Pezeshki^(✉), Dimitrios Pavlou, and Sudath C. Siriwardane

University of Stavanger, 4021 Stavanger, Norway
Hadi.Pezeshki@uis.no

Abstract. A comprehensive analytical solution for the estimation of natural frequencies of monopile supported Offshore Wind Turbines (OWTs) is proposed in this paper. The solution is based on the non-dimensional equations established using the Euler-Bernoulli beam theory and relevant boundary conditions. The equations of motion for above water and underwater are separated to compensate for the effect of added mass. Therefore, three equations of motion are established to simulate the motion of the monopile underwater, above water, and tower. The tower is considered a tapered tubular structure that linearly varies in diameter from bottom to top. Four sets of boundary conditions are also established at the seabed, seawater, platform, and nacelle levels. The seabed boundary conditions are to model the effect of the monopile section under the seabed by using the coupled springs. The seawater and platform level boundary conditions are the continuity conditions to link the monopile underwater, above water, and tower. The translational and rotational masses of the nacelle-rotor assembly are also included in the boundary conditions at the nacelle level. The non-dimensional equations are solved for a DTU 10 MW OWT case, and the results are compared to similar works. Besides, the 1st natural frequency of the system is evaluated for different values of monopile length and cross-section. The proposed method can easily be programmed to estimate the natural frequencies of the OWT structures.

1 Introduction

Accurate estimation of the natural frequencies is a key to opening the structural properties of every vibration system. This estimation is essential in the case of OWT, where structural health monitoring (SHM) is intended. SHM is a technique by which the damage severity can be evaluated. Several approaches have been proposed to establish SHM including model-based [1, 2], vibration-based [3] and vision-based [4–6]. Establishing every SHM method requires taking three main steps:

Structural data measurement via several types of sources such as sensors [7–9], Laser/LIDAR [10, 11], ground penetration radar [12], Unmanned Aerial Vehicle (UAV) [13].

Data cleansing and signal processing [14, 15].

Interpretation by the aid of the machine learning algorithms [16–23].

SHM has been successfully applied for the vast variety of the infrastructures in several studies [7, 24, 25]. For the case of offshore and marine structures, SHM application

is not as mature as onshore infrastructures. Pezeshki et al. [1] reviewed the challenges and approaches for establishing SHM in the offshore and marine structures. They stated that the offshore environment creates challenges in the establishing SHM in these structures. The main reason is due to the presence of environment-structure interactions such as soil-structure and fluid-structure interactions. One of the major effects of these interactions is the natural variation of the modal parameters. These variations may cause misunderstanding in damage detection via vibration-based SHM in which any changes in modal parameters are concerned [26, 27]. Studying these variations requires deep understanding of interactions via mathematical modeling which can be ended up to the model-based SHM [2].

The variation of the natural frequencies of OWT structure has been reported based on the in-situ measurements [28–31]. Variation of the natural frequencies of the OWTs is due to the variation in their structural properties. Most of the works focused on the soil and foundation situation, seeking an explanation for this variation. Damgaard et al. [28] have reported the time dependency of the natural frequencies of an OWT by investigating its cross-wind modal properties. They introduced soil erosion as a responsible factor for this variation. Later, Prendergat et al. [29, 30] investigated the scouring effect on the natural frequencies. They pointed out that scouring is not the only factor changing the structural properties of OWTs. Other factors, such as added mass, were recognized to influence this phenomenon. In a study conducted by Cosgriff et al. [31] the measured 1st natural frequency of an OWT was plotted against the wave height and wind speed. They reported that the 1st natural frequency was the lowest when the wave's crest passed the monopile. Further, the frequency immediately increases when the water level decreases to the trough. The reason for this phenomenon can be explained by the effect of the added mass on the system.

Studying the variation of the natural frequencies of OWT structures requires establishing an accurate model, including all the important parameters. Commercial software packages or numerical platforms were the popular tools among many researchers [32–36] in simulating these structures. Only a few researchers have attempted to use the analytical approach in this field. The analytical approach requires solving partial differential equations (PDEs) obtained by beam theories. Solutions for these PDEs have been attempted by several scientists such as Graff [37] and Meirovitch [38] in the 1960s and 1970s by utilizing classical methods such as Finite Fourier transform, Expansion in the natural modes (in the spatial domain), and Laplace transform. Pavlou [39] developed an analytical solution for a bottom fixed OWT loaded by the wave load by utilizing Laplace Transforms. Recently, Pezeshki et al. [2] developed an analytical solution by including the soil-structure interaction, wave-structure interaction, and translational and rotational mass inertia in their model. Their proposed method calculated the natural frequencies analytically by considering the geometrical differences between the monopile and the tower. Their results reveal approximately 11–13% inaccuracy of the 1st natural frequency of their model and similar works. This inaccuracy is mainly due to their simplification in modeling the tower as a uniform cross-section.

This paper aims to address the weakness of the model developed by Pezeshki et al. [2] by considering the tower as a tapered shape in which the cross-sectional diameter linearly decreases in height. The governing equations of motion for the tower and the

monopile are established. Four sets of boundary conditions are formulated in the sea level, water level, platform level, and the nacelle level. The method for calculation of the natural frequencies is introduced. In the end, the proposed method was implemented in solving an example, and the results were compared with similar works. The proposed method consists of two novelty, (1) the equations of motions are nondimensionalized, (2) the linear variation of the tower diameter is considered. Therefore, the proposed method can effectively estimate the system's natural frequencies.

2 The Structural Model

A bottom fixed OWT structure consists of a tower carrying a nacelle-rotor assembly at the top and a foundation supporting the tower. The foundation for typical shallow to intermediate water depth is primarily a monopile driven in the seabed. This system can be modeled as a cantilever beam shown in Fig. 1.

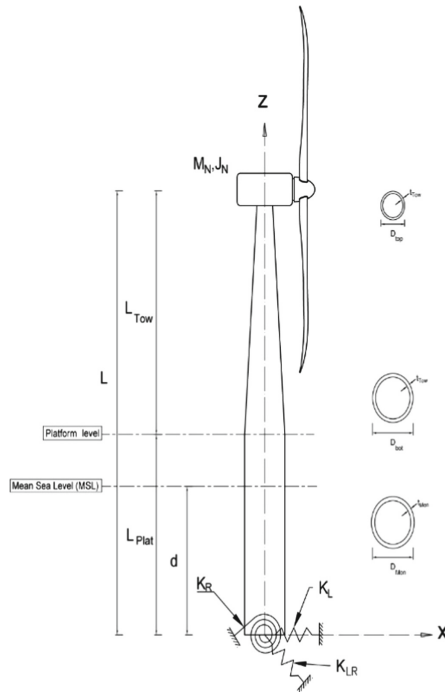


Fig. 1. The configuration of the model of the OWT structure

For the Euler-Bernoulli beam, the equation of motion can be written as

$$\partial^2 / (\partial z^2) (EI(z) (\partial^2 x(z, t)) / (\partial x^2)) + m(z) (\partial^2 x(z, t)) / (\partial t^2) = 0 \tag{1}$$

In the above equation, only a homogenous form is established since the natural frequencies of the system are of interest. Introducing a form of solution for Eq. (1) as

$$x(z, t) = X(z) e^{-i\omega t} \tag{2}$$

the above equation will simplify to

$$\partial^2 / (\partial z^2) \left(EI(z) \frac{\partial^2 X(z)}{\partial x^2} \right) - m(z) \omega^2 X(z) = 0 \tag{3}$$

The diameter of the cross-section can be expressed as a function of z as

$$D(z) = \begin{cases} D_{bot}(1 - a(z - L_{plat})) & \text{for } L_{plat} \leq z \leq L \\ D_{Mon}, & \text{for } 0 < z \leq L_{plat} \end{cases} \tag{4}$$

where

$$a = \frac{1 - \frac{D_{top}}{D_{bot}}}{L_{Tow}} \tag{5}$$

Using Eq. (4), $EI(z)$ and $m(z)$ can be derived as

$$EI(z) = \begin{cases} EI_{bot}(b - az)^3 & \text{for } L_{plat} \leq z \leq L \\ EI_{Mon}, & \text{for } 0 < z \leq L_{plat} \end{cases} \tag{6}$$

where

$$I_{bot} = \frac{\pi}{8} D_{bot}^3 t_{tow}, \quad I_{Mon} = \frac{\pi}{8} D_{Mon}^3 t_{Mon} \tag{7}$$

$$b = 1 + aL_{plat} \tag{8}$$

And

$$m(z) = \begin{cases} m_{bot}(b - az)^3 & \text{for } L_{plat} \leq z \leq L \\ m_{Mon}, & \text{for } 0 < z \leq L_{plat} \\ m_{Mon} + m_{Add}, & \text{for } 0 < z \leq d \end{cases} \tag{9}$$

where

$$m_{bot} = \pi D_{bot} t_{tow} \rho_s, \quad m_{Mon} = \pi D_{Mon} t_{Mon} \rho_s \tag{10}$$

It should be noted that the thickness of the tower and monopile is assumed to be constant along the length. Besides the added mass, m_{Add} , is included in the system's

total mass underwater in Eq. (9). Introducing Eqs. (6) and (9) into Eq. (3), the equations of motions will be formed as

$$(b - az)^2 X_{Tow}^{(iv)}(z) - 6a(b - az)X_{Tow}'''(z) + 6a^2 X_{Tow}''(z) - \frac{m_{bot}\omega^2}{EI_{bot}} X_{Tow} = 0, \text{ for } L_{Plat} \leq z \leq L \quad (11)$$

$$X_{MA}^{(iv)}(z) - \frac{m_{Mon}\omega^2}{EI_{Mon}} X_{MA}(z) = 0, \text{ for } d < z \leq L_{Plat} \quad (12)$$

$$X_{MU}^{(iv)}(z) - \frac{(m_{Mon} + m_{Add})\omega^2}{EI_{Mon}} X_{MU}(z) = 0, \text{ for } 0 < z \leq d \quad (13)$$

3 The Boundary Conditions

Four sets of boundary conditions are considered for the model. The first set is at the seabed, where the coupled springs simulate the soil-structure interaction [40]. Therefore, the seabed boundary conditions can be written as

$$-EI_{Mon}X_{MU}'''(0) = K_L X_{MU}(0) + K_{LR}X_{MU}'(0) \quad (14)$$

$$EI_{Mon}X_{MU}''(0) = K_{LR}X_{MU}(0) + K_R X_{MU}'(0) \quad (15)$$

The second sets of boundary conditions are at sea level, which are

$$X_{MU}(d) = X_{MA}(d) \quad (16)$$

$$X_{MU}'(d) = X_{MA}'(d) \quad (17)$$

$$X_{MU}''(d) = X_{MA}''(d) \quad (18)$$

$$X_{MU}'''(d) = X_{MA}'''(d) \quad (19)$$

The third sets are for the platform level linking the tower and the monopile. They are

$$X_{MA}(L_{Plat}) = X_{Tow}(L_{Plat}) \quad (20)$$

$$X_{MA}'(L_{Plat}) = X_{Tow}'(L_{Plat}) \quad (21)$$

$$EI_{Mon}X_{MA}''(L_{Plat}) = EI_{Tow}X_{Tow}''(L_{Plat}) \quad (22)$$

$$EI_{Mon}X'''_{MA}(L_{Plat}) = \frac{\partial}{\partial z} \left(EI(z) \frac{\partial^2 X(z)}{\partial x^2} \right) \Bigg|_{z=L_{Plat}} \quad (23)$$

And Finally, the fourth sets of boundary conditions are at the nacelle level. They are

$$\frac{\partial}{\partial z} \left(EI(z) \frac{\partial^2 X(z)}{\partial x^2} \right) \Bigg|_{z=L} = -M_N \omega^2 X_{Tow}(L) \quad (24)$$

$$-EI_{top}X''_{Tow}(L) = -J_P \omega^2 X'_{Tow}(L) \quad (25)$$

Substituting Eq. (6) into Eqs. (23) and (25), they will be

$$EI_{Mon}X'''_{MA}(L_{Plat}) = -3EI_{bot}bX''_{Tow}(L_{Plat}) + EI_{bot}X'''_{Tow}(L_{Plat}) \quad (26)$$

$$-3EI_{bot}a(b-aL)^2X''_{Tow}(L) + EI_{bot}(b-aL)^3X'''_{Tow}(L) = -M_N \omega^2 X_{Tow}(L) \quad (27)$$

4 Nondimensionalization

The equations of motion and their associated boundary conditions can be expressed in the non-dimensional form. Introducing the following non-dimensional coefficients

$$\bar{z} = \frac{z}{L}, \bar{L}_{Plat} = \frac{L_{Plat}}{L}, \bar{L}_{Tow} = \frac{L_{Tow}}{L}, \bar{d} = \frac{d}{L}, \bar{D} = \frac{D_{top}}{D_{bot}} \quad (28)$$

Equations (5) and (8) yield in

$$\bar{a} = aL = \frac{1 - \bar{D}}{\bar{L}} \quad (29)$$

$$b = 1 + (1 - \bar{D}) \frac{\bar{L}_{Plat}}{\bar{L}_{Tow}} \quad (30)$$

Introducing the non-dimensional nacelle-rotor translational and rotational mass inertia as

$$\bar{M} = \frac{M_N}{M_{Tow}}, \bar{J} = \frac{J_N}{J_{Tow}} \quad (31)$$

where M_{Tow} is the total mass of the tower and J_{Tow} is the mass moment of inertia of the tower with respect to the top section. They can calculate with the following equations as

$$M_{Tow} = \int_{L_{Plat}}^L m_{bot}(b-az)dz = m_{bot} \frac{L_{Tow}}{2} (1 + \bar{D}) = m_{bot} L \frac{\bar{L}_{Tow}}{2} (1 + \bar{D}) \quad (32)$$

$$J_{Tow} = \int_{L_{Plat}}^L m_{bot}(L-z)^2(b-az)dz = m_{bot} \frac{L_{Tow}^3}{12} (3 + \bar{D}) = m_{bot} L^3 \frac{\bar{L}_{Tow}^3}{12} (3 + \bar{D}) \quad (33)$$

Moreover, the equations of motion in Eqs. (11)–(13) will become non-dimensional by using Eqs. (28)–(30). They are

$$(b - \bar{az})^2 X_{Tow}^{(iv)}(\bar{z}) - 6\bar{a}(b - \bar{az}) X_{Tow}'''(\bar{z}) + 6\bar{a}^2 X_{Tow}'(\bar{z}) - \beta_{Tow}^4 X_{Tow}(\bar{z}) = 0, \text{ for } \bar{L}_{Plat} \leq z \leq 1 \quad (34)$$

$$X_{MA}^{(iv)}(\bar{z}) - \beta_{MA}^4 X_{MA}(\bar{z}) = 0, \text{ for } \bar{d} < z \leq \bar{L}_{Plat} \quad (35)$$

$$X_{MU}^{(iv)}(\bar{z}) - \beta_{MU}^4 X_{MU}(\bar{z}) = 0, \text{ for } 0 < z \leq \bar{d} \quad (36)$$

where

$$\beta_{Tow}^4 = \frac{m_{bot} L^4 \omega^2}{EI_{bot}} \quad (37)$$

$$\beta_{MA}^4 = \frac{m_{Mon} L^4 \omega^2}{EI_{Mon}} \quad (38)$$

$$\beta_{MU}^4 = \frac{(m_{Mon} + m_{Add}) L^4 \omega^2}{EI_{Mon}} \quad (39)$$

The non-dimensional form of the boundary conditions can also be obtained by using the non-dimensional coefficients in Eqs. (28)–(30). They are

$$X_{MU}'''(0) + \bar{K}_L X_{MU}(0) + \bar{K}_{LR} X_{MU}'(0) = 0, \bar{K}_L = \frac{K_L L^3}{EI_{Mon}}, \bar{K}_{LR} = \frac{K_{LR} L^2}{EI_{Mon}} \quad (40)$$

$$X_{MU}''(0) - \bar{K}_{LR} X_{MU}(0) - \bar{K}_R X_{MU}'(0) = 0, \bar{K}_R = \frac{K_R L}{EI_{Mon}} \quad (41)$$

$$X_{MU}(\bar{d}) - X_{MA}(\bar{d}) = 0 \quad (42)$$

$$X_{MU}'(\bar{d}) - X_{MA}'(\bar{d}) = 0 \quad (43)$$

$$X_{MU}''(\bar{d}) - X_{MA}''(\bar{d}) = 0 \quad (44)$$

$$X_{MU}'''(\bar{d}) - X_{MA}'''(\bar{d}) = 0 \quad (45)$$

$$X_{MA}(\bar{L}_{Plat}) - X_{Tow}(\bar{L}_{Plat}) = 0 \quad (46)$$

$$X'_{MA}(\bar{L}_{Plat}) - X'_{Tow}(\bar{L}_{Plat}) = 0 \tag{47}$$

$$X''_{MA}(\bar{L}_{Plat}) - \bar{E}I X''_{Tow}(\bar{L}_{Plat}) = 0, \bar{E}I = \frac{EI_{bot}}{EI_{Mon}} \tag{48}$$

$$X'''_{MA}(\bar{L}_{Plat}) + 3\bar{E}I\bar{a}X''_{Tow}(\bar{L}_{Plat}) - \bar{E}I X'''_{Tow}(\bar{L}_{Plat}) = 0 \tag{49}$$

$$X''_{Tow}(1) - \bar{J}\eta_1\beta_{Tow}^4 X_{Tn'}(1) = 0, \eta_1 = \frac{\bar{L}_{Tow}^3(3 + \bar{D})}{12\bar{D}^3} \tag{50}$$

$$X'''_{Tow}(1) - 3\frac{\bar{a}}{D}X''_{Tow}(1) + \bar{M}\eta_2\beta_{Tow}^4 X_{Tow}(1), \eta_2 = \frac{\bar{L}_{Tow}(1 + \bar{D})}{2\bar{D}^3} \tag{51}$$

Solution of the equations of motion and boundary conditions.

The solution of Eq. (34) may be written in terms of Bessel Functions [41, 42] namely

$$X_{Tow}(\bar{z}) = \frac{1}{\sqrt{b - \bar{a}\bar{z}}} \left(T_1 J_1 \left(2\frac{\beta_{Tow}}{\bar{a}} \sqrt{b - \bar{a}\bar{z}} \right) + T_2 Y_1 \left(2\frac{\beta_{Tow}}{\bar{a}} \sqrt{b - \bar{a}\bar{z}} \right) \right. \\ \left. + T_3 I_1 \left(2\frac{\beta_{Tow}}{\bar{a}} \sqrt{b - \bar{a}\bar{z}} \right) + T_4 K_1 \left(2\frac{\beta_{Tow}}{\bar{a}} \sqrt{b - \bar{a}\bar{z}} \right) \right) \tag{52}$$

where $J_1, Y_1, I_1,$ and K_1 are the first-order Bessel functions and $T_1 \dots T_4$ are constants. For the Eqs. (35) and (36), the solution can be found by their characteristic equations as

$$X_{MA}(\bar{z}) = A_1 \cos(\beta_{MA}\bar{z}) + A_2 \cos h(\beta_{MA}\bar{z}) + A_3 \sin(\beta_{MA}\bar{z}) + A_4 \sin h(\beta_{MA}\bar{z}) \tag{53}$$

$$X_{MU}(\bar{z}) = U_1 \cos(\beta_{MU}\bar{z}) + U_2 \cos h(\beta_{MU}\bar{z}) + U_3 \sin(\beta_{MU}\bar{z}) + U_4 \sin h(\beta_{MU}\bar{z}) \tag{54}$$

where $A_1 \dots A_4$ and $U_1 \dots U_4$ are constants that can be found from the boundary conditions.

To write the frequency equation for the system, it is required to find relationships between $\beta_{Tow}, \beta_{MA},$ and β_{MU} . Using Eqs. (37)–(39), these relationships can be derived as

$$\beta_{MA} = \gamma_{AT} \beta_{Tow}, \gamma_{AT} = \sqrt[4]{\frac{EI}{m_{bot}} \frac{m_{Mon}}{m_{bot}}} \tag{55}$$

$$\beta_{MU} = \gamma_{UT} \beta_{Tow}, \gamma_{UT} = \sqrt[4]{\frac{EI}{m_{bot}} \frac{m_{Mon} + m_{Add}}{m_{bot}}} \tag{56}$$

Substituting the solutions in Eqs. (52)–(54) into the boundary conditions in Eqs. (40)–(51) and using Eqs. (56) and (57), the frequency equation can be derived.

5 Numerical Example

DTU 10 MW three-bladed OWT properties are selected as a numerical example. The structural and foundation properties of this OWT are obtained from Bak et al. [43] and Alkhoury et al. [36], respectively, and represented in Tables 1 and 2. In Table 3, the

Table 1. DTU 10 MW OWT structural properties [43]

	Symbol	Value
Tower length (m)	L_{Tow}	119
Tower top diameter (m)	D_{top}	5.5
Tower bottom diameter (m)	D_{bot}	8.3
Tower average thickness (m)	t_{Tow}	0.0314
Tower Young's modulus (GPa)	E_{Tow}	210
Tower total mass (kg)	M_{Tow}	682442
Tower Material density (kg/m^3)	ρ_{sT}	8500
Nacelle-Rotor assembly mass (kg)	M_N	676723
Nacelle-Rotor assembly rotational inertia (kg.m^2)	J_N	1.27×10^8
Platform level from the seabed (m)	L_{Plat}	45
Monopile average diameter (m)	D_{Mon}	8.3
Monopile average thickness (m)	t_{Mon}	0.09
Monopile Young's modulus (GPa)	E_{Mon}	210
Material density (kg/m^3)	ρ_{sT}	7850

Table 2. The values of the coupled springs [36]

	Symbol	Value
Lateral stiffness (GN/m)	K_L	2.48
Cross stiffness (GN)	K_{LR}	-20.7
Rotational stiffness (GN.m)	K_R	412

Table 3. Hydrodynamic loading properties

	Symbol	Value
Water depth (m)	d	35
Added mass coefficient	C_A	1
Sea water density (kg/m^3)	ρ_w	1030

hydrodynamic loading properties are listed by the recommendation provided by DNV-RP-C205 [44]. It should be noted that the effect of the added mass is considered from Morison's formula [45], which is $\rho_w C_{AA} A_{Mon}$.

The natural frequencies are calculated by the proposed method and listed in Table 4. These results are also compared with recently published work by Pezeshki et al. [2]. In their model, the uniform cross-section of the tower was considered, while the tapered form is modeled in this paper. Table 4 reveals a significant improvement in the 1st natural frequency estimation by only considering the tower as a tapered beam reaching an 11.2% deviation. Alkhoury et al. [36] also reported this deviation which was modeling the OWT using the commercial numerical software program Abaqus.

Table 4. The comparison of the results with the previously published work

Mode number	Pezeshki et al. [2]	Proposed model	Deviation (%)
1	0.166	0.185	11.2
2	1.032	1.001	-2.6
3	1.984	1.911	-3.7
4	3.817	3.484	-8.7
5	6.593	6.371	-3.4
6	9.890	9.899	0.1

Besides, the results of the proposed method are compared with the detailed numerical model created in Abaqus for DTU 10 MW OWT performed by Alkhoury et al. [36]. The embedded part of the monopile and surrounding soil were included in their model to create a full 3D model. Using this model, they performed a parametric study on the different variables of the monopile. The same parametric study was performed using the proposed model, and the results are compared with the full 3D model in Table 5. The proposed method underestimates the 1st natural frequency by approximately 4.1% with respect to the full 3D model. The only source of this deviation can be the variation of the tower thickness in length, which is fully modeled in the numerical model while it was kept constant in this study. Differences in the nature of the methods of modeling of the structure at the boundary conditions are also sources of this deviation.

The model proposed in this paper is an improvement of the model proposed by Pezeshki et al. [2] by considering the tower variation of the tower cross-section. They also performed the parametric study in Table 5. The average deviation of their work was -14.7% with respect to the results of Alkhoury et al. [36], while the value for the proposed method is -4.1%. Therefore, an almost 11% improvement can be seen by considering the variation of the tower cross-section.

Table 5. The comparison of the 1st natural frequency of the proposed method with the full 3D model

Water depth (m)	Platform level (m)	Monopile outer diameter (m)	Monopile thickness (cm)	1st Natural frequency (Hz)		Deviation (%)
				Alkhoury et al. [36]	Proposed model	
25	35	8.3	9	0.200	0.194	-3.0
			10	0.202	0.196	-2.5
			12	0.205	0.200	-2.7
		9	10	0.207	0.201	-3.0
			12	0.211	0.203	-3.5
			14	0.215	0.205	-4.6
		10	11	0.217	0.206	-5.1
			13	0.218	0.208	-4.7
			15	0.220	0.209	-5.0
35	45	8.3	9	0.190	0.185	-3.1
			10	0.192	0.187	-2.7
			12	0.197	0.191	-2.8
		9	10	0.199	0.192	-3.3
			12	0.203	0.196	-3.8
			14	0.209	0.198	-5.1
		10	11	0.209	0.199	-4.8
			13	0.212	0.201	-5.2
			15	0.215	0.203	-5.6
45	55	8.3	9	0.181	0.175	-3.2
			10	0.183	0.178	-2.9
			12	0.189	0.182	-3.3
		9	10	0.190	0.184	-3.5
			12	0.196	0.188	-4.2
			14	0.202	0.190	-5.6
		10	11	0.203	0.192	-5.3
			13	0.207	0.194	-5.9
			15	0.210	0.196	-6.3
Average deviation						-4.1

6 Conclusion

In this study, an analytical model is presented for calculating the natural frequencies of OWT structures. The structure's motion is formulated into three distinct parts, namely the tower, the monopile underwater, and the monopile above water. There are four sets of boundary conditions formulated for the nacelle, platform, sea level, and seabed levels. An important feature of the proposed method is that (1) all equations of motion and associated boundary conditions are non-dimensional. (2) A tapered cantilever beam is considered as the equation of motion of the tower. The DTU 10 MW OWT's natural frequencies are solved by the proposed method as a numerical example. Based on the results, considering the linear variation of the tower cross-section will improve estimation accuracy by approximately 11% compared to the case in which a uniform cross-section is considered. Nevertheless, the proposed method lacks around -4% accuracy in comparison to the full 3D numerical model. This method is analytical in nature and can be programmed in order to estimate the natural frequencies of OWT structures quickly and cost-effectively.

Appendix 1: Nomenclature

Symbol	Structural properties
L	Nacelle level from the seabed (m)
L_{Tow}	Tower length (m)
L_{Plat}	Platform level from the seabed (m)
d	Water depth (m)
D_{Top}	Tower top cross-section diameter (m)
D_{bot}	Tower bottom cross-section diameter (m)
t_{Tow}	Tower average thickness (m)
m_{bot}	Tower mass of unit length at the bottom cross-section (Kg/m)
E_{Tow}	Tower Young's modulus (GPa)
EI_{Tow}	Flexural rigidity of the tower, i.e., $E_{Tow}I_{Tow}$ ($GPa.m^4$)
M_N	Nacelle-Rotor assembly mass (Kg)
J_N	Nacelle-Rotor assembly rotational inertia ($Kg.m^2$)
D_{Mon}	Monopile average diameter (m)
t_{Mon}	Monopile average thickness (m)
A_{Mon}	Monopile cross-sectional area (m^2)
m_{Mon}	Monopile mass of unit length (Kg/m)
E_{Mon}	Monopile Young's modulus (GPa)
EI_{Mon}	Flexural rigidity of the monopile, i.e., $E_{Mon}I_{Mon}$ ($GPa.m^4$)
ρ_s	Material Density (Kg/m^3)

(continued)

(continued)

Symbol	Structural properties
K_L	Lateral stiffness (GN/m)
K_{LR}	Cross stiffness (GN)
K_R	Rotational stiffness ($GN.m$)
C_A	Added mass coefficient
ρ_w	Sea water density (Kg/m^3)

References

1. Pezeshki, H., Adeli, H., Pavlou, D.G., and Siriwardane, S. C. State of the art in structural health monitoring of offshore and marine structures, Proceedings of the Institution of Civil Engineers-Maritime Engineering, Volume 0: 1–8, (2022), <https://doi.org/10.1680/jmaen.2022.027>.
2. Pezeshki, H., Pavlou, D.G., Adeli, H., and Siriwardane, S. C. Modal analysis of offshore monopile wind turbine: An analytical solution, Journal of Offshore Mechanics and Arctic Engineering 145(1): 010907, (2022), doi: <https://doi.org/10.1115/1.4055402>.
3. Sajedi, S., and Liang, X. Dual Bayesian Inference for Risk-Informed Vibration-Based Damage Diagnosis, Computer-Aided Civil and Infrastructure Engineering 36 (9): 1168–84, (2021). <https://doi.org/10.1111/mice.12642>.
4. Sajedi, S.M. and Liang, X., Deep Generative Bayesian Optimization for Sensor Placement in Structural Health Monitoring, Computer-Aided Civil and Infrastructure Engineering, 37:9, 1109-1127, (2022). <https://doi.org/10.1111/mice.12799>.
5. Zhang, Y. and Lin, W., Computer Vision-based Differential Remeshing for Updating the Geometry of Finite Element Model, Computer-Aided Civil and Infrastructure Engineering, 37:2, 185-203, (2022). <https://doi.org/10.1111/mice.12708>.
6. Zhao, J., Hu, F., Xu, Y., Zuo, W., Zhong, J., and Li, H., Structure-PoseNet for identification of dense dynamic displacement and 3D poses of structures using a monocular camera, Computer-Aided Civil and Infrastructure Engineering, 37:6, 704-725, (2022). <https://doi.org/10.1111/mice.12761>.
7. Zhang, Q. and Zhang, J., Internal Force Monitoring and Estimation of a Long-Span Reinforced Concrete Ring Beam using Long-gauge Strain Sensing, Computer-Aided Civil and Infrastructure Engineering, 36:1, 109-124, (2021). <https://doi.org/10.1111/mice.12569>.
8. Civera, M., Pecorelli, M.L., Ceravolo, R., Surace, C., and Fragonara, L.Z., A Multi-objective Genetic Algorithm Strategy for Robust Optimal Sensor Placement, Computer-Aided Civil and Infrastructure Engineering, 36:9, 1185-1202, (2021). <https://doi.org/10.1111/mice.12646>.
9. Jauhainen, J., Pour-Ghaz, M., Valkonen, T., and Seppänen, A., Non-planar sensing skins for structural health monitoring based on electrical resistance tomography, Computer-Aided Civil and Infrastructure Engineering, 36:12, (2021). <https://doi.org/10.1111/mice.12689>.
10. Nasimi, R. and Moreu, F., A methodology for measuring the total displacements of structures using a laser-camera system, Computer-Aided Civil and Infrastructure Engineering, 36:4, 421-437, (2021). <https://doi.org/10.1111/mice.12652>.
11. Shen, J., Yan, W., Li, P., and Xiong, X., Deep Learning-based Object Identification with Instance Segmentation and Pseudo-LiDAR Point Cloud for Work Zone Safety, Computer-Aided Civil and Infrastructure Engineering, 36:12, 1549-1567, (2021). <https://doi.org/10.1111/mice.12749>.

12. Li, X., Liu, H., Zhou, F., Chen, Z., Giannakis, I., and Slob, E., Deep learning-based nondestructive evaluation of reinforcement bars using ground-penetrating radar and electromagnetic induction data, *Computer-Aided Civil and Infrastructure Engineering*, 37:14, 1834-1853, (2022). <https://doi.org/10.1111/mice.12798>.
13. Tian, Y., Zhang, C., Jiang, S., Zhang, J., and Duan, W., Noncontact Cable Force Estimation with Unmanned Aerial Vehicle and Computer Vision, *Computer-Aided Civil and Infrastructure Engineering*, 36:1, 73-88, (2021). <https://doi.org/10.1111/mice.12567>.
14. Qarib, H. and Adeli, H., A Comparative Study of Signal Processing Methods for Exponentially Damped Signals, *Journal of Vibroengineering*, 18:4, NoP (2027-2079), pp. 2186-2204, (2016). <https://doi.org/10.1111/mice.12567>.
15. Sajedi, S., and Liang, X., Uncertainty-Assisted Deep Vision Structural Health Monitoring, *Computer-Aided Civil and Infrastructure Engineering* 36 (2): 126–42, (2021). <https://doi.org/10.1111/mice.12580>.
16. Adeli, H., & Hung, S. L.. Machine learning–neural networks, genetic algorithms, and fuzzy systems. John Wiley and Sons, New York, (1995).
17. Ahmadlou, M., & Adeli, H., Enhanced probabilistic neural network with local decision circles: A robust classifier. *Integrated Computer-Aided Engineering*, 17(3), 197–210, (2010). DOI: <https://doi.org/10.3233/ICA-2010-0345>.
18. Rafiei, M. H., & Adeli, H., A new neural dynamic classification algorithm. *IEEE Transactions on Neural Networks and Learning Systems*, 28(12), 3074–3083, (2017). <https://doi.org/10.1109/TNNLS.2017.2682102>.
19. Sørensen, R.A., Michael N., and Henrik K., Routing in Congested Baggage Handling Systems Using Deep Reinforcement Learning, *Integrated Computer-Aided Engineering* 27 (2): 139–52, (2020). <https://doi.org/10.3233/ICA-190613>.
20. Gil-Gala, F.J., Carlos M., María R.S., and Ramiro V., Learning Ensembles of Priority Rules for Online Scheduling by Hybrid Evolutionary Algorithms, *Integrated Computer-Aided Engineering* 28: 65–80 (2021). DOI:<https://doi.org/10.3233/ICA-200634>.
21. Gao, Y., Zhai, P., and Mosalam, K.M. Balansced Semi-Supervised Generative Adversarial Network for Damage Assessment from Low-Data Imbalanced-Class Regime, *Computer-Aided Civil and Infrastructure Engineering*, 36:9, 1094-1113, (2021). DOI: <https://doi.org/10.1111/mice.12741>.
22. Sarmadi, H. and Yuen, K.V. Early damage detection by an innovative unsupervised learning method based on kernel null space and peak-over-threshold, *Computer-Aided Civil and Infrastructure Engineering*, 36:9, 1150-1167, (2021). DOI: <https://doi.org/10.1111/mice.12635>.
23. Maeda, H., Kashiyama, T., Sekimoto, Y., Seto, T., Omata, H., Generative Adversarial Networks for Road Damage Detection, *Computer-Aided Civil and Infrastructure Engineering*, 36:47-60, (2021). DOI: <https://doi.org/10.1111/mice.12561>.
24. Chun, P.J., Yamane, T., and Maemura, Y., A deep learning based image captioning method to automatically generate comprehensive explanations of bridge damage, *Computer-Aided Civil and Infrastructure Engineering*, 37:11, 1387-1401, (2022). <https://doi.org/10.1111/mice.12793>.
25. Rafiei, M. H., & Adeli, H., A novel machine learning based algorithm to detect damage in highrise building structures. *The Structural Design of Tall and Special Buildings*, 26, 18, (2017). <https://doi.org/10.1002/tal.1400>.
26. Amezquita-Sanchez, J. P., Park, H.P., and Adeli, H., A Novel Methodology for Modal Parameters Identification of Large Smart Structures Using MUSIC, Empirical Wavelet Transform, and Hilbert Transform. *Engineering Structures* 147 (September): 148–59, (2017). <https://doi.org/10.1016/j.engstruct.2017.05.054>.

27. Lin, Z., Dubravka Pokrajac, D., Guo, Y., Jeng, D.S., Tang, T., Rey, N., Zheng, J., and Zhang, J., Investigation of Nonlinear Wave-Induced Seabed Response around Mono-Pile Foundation. *Coastal Engineering* 121: 197–211, (2017). <https://doi.org/10.1016/j.coastaleng.2017.01.002>.
28. Damgaard, M., Ibsen, L. B., Andersen, L. V., and Andersen, J. K.F., Cross-Wind Modal Properties of Offshore Wind Turbines Identified by Full Scale Testing, *Journal of Wind Engineering and Industrial Aerodynamics* 116: 94–108, (2013). <https://doi.org/10.1016/j.jweia.2013.03.003>.
29. Prendergast, L. J., Gavin, K., and Doherty, P., An Investigation into the Effect of Scour on the Natural Frequency of an Offshore Wind Turbine, *Ocean Engineering* 101: 1–11, (2015). <https://doi.org/10.1016/j.oceaneng.2015.04.017>.
30. Prendergast, L. J., Reale, C., and Gavin, K., Probabilistic Examination of the Change in Eigenfrequencies of an Offshore Wind Turbine under Progressive Scour Incorporating Soil Spatial Variability. *Marine Structures* 57: 87–104, (2018). <https://doi.org/10.1016/j.marstruc.2017.09.009>.
31. Norén-Cosgriff, K., and Kaynia, A.M. Estimation of Natural Frequencies and Damping Using Dynamic Field Data from an Offshore Wind Turbine, *Marine Structures* 76: 102915, (2021). <https://doi.org/10.1016/j.marstruc.2020.102915>.
32. Luan, M., Qu, P., Jeng, D.S., Guo, Y., and Yang, Q., Dynamic Response of a Porous Seabed–Pipeline Interaction under Wave Loading: Soil–Pipeline Contact Effects and Inertial Effects. *Computers and Geotechnics* 35 (2): 173–86, (2008). <https://doi.org/10.1016/j.compageo.2007.05.004>.
33. Chang, K.T., and Jeng, D.S., Numerical Study for Wave-Induced Seabed Response around Offshore Wind Turbine Foundation in Donghai Offshore Wind Farm, Shanghai, China. *Ocean Engineering* 85: 32–43, (2014). <https://doi.org/10.1016/j.oceaneng.2014.04.020>.
34. Sui, T., Zhang, C., Guo, Y., Zheng, J., Jeng, D., Zhang, J., and Zhang, W., Three-Dimensional Numerical Model for Wave-Induced Seabed Response around Mono-Pile. *Ships and Offshore Structures* 11 (6): 667–78, (2015). <https://doi.org/10.1080/17445302.2015.1051312>.
35. Zhang, C., Zhang, Q., Wu, Z., Zhang, J., Sui, T., and Wen, Y., Numerical Study on Effects of the Embedded Monopile Foundation on Local Wave-Induced Porous Seabed Response. *Mathematical Problems in Engineering* 2015: 1–13, (2015). <https://doi.org/10.1155/2015/184621>.
36. Alkhoury, P., Soubra, A.H., Rey, V., and Ait-Ahmed, M., A Full Three-Dimensional Model for the Estimation of the Natural Frequencies of an Offshore Wind Turbine in Sand. *Wind Energy* 24 (7): 699–719, (2021). <https://doi.org/10.1002/we.2598>.
37. Graff, K F., *Wave Motion in Elastic Solids*. Dover Publications, New York, 2012.
38. Meirovitch, L., *Analytical Methods in Vibrations*. Macmillan, New York.
39. Pavlou, Dimitrios G. “Soil–Structure–Wave Interaction of Gravity-Based Offshore Wind Turbines: An Analytical Model.” *Journal of Offshore Mechanics and Arctic Engineering* 143 (3), 2021. <https://doi.org/10.1115/1.4048997>.
40. Darvishi-Alamouti, S., Bahaari, M.R., and Moradi, M., Natural Frequency of Offshore Wind Turbines on Rigid and Flexible Monopiles in Cohesionless Soils with Linear Stiffness Distribution. *Applied Ocean Research* 68: 91–102 (2017). <https://doi.org/10.1016/j.apor.2017.07.009>.
41. Polyanin A.D, Zaitsev. *Handbook of Exact Solutions for Ordinary Differential Equations*. Chapman & Hall/CRC, NewYork, 2003.
42. Sanger, D. J., Transvers vibration of a class of non-uniform beams, *Journal of Mechanical Engineering science*, Vol 10 (2): 111-120, 1967.
43. Bak, Christian, Frederik Zahle, Robert Bitsche, Taeseong Kim, Anderes Yde, Lars Christian Henriksen, Anand Natarajan, and Morten Hansen., Description of the DTU 10 MW Reference Wind Turbine, 2013. <https://dtu-10mw-rwt.vindenergi.dtu.dk>.

44. Veritas, Det Norske. Environmental Conditions and Environmental Loads.” NNV no. October: 9–123, 2010.
45. Morison, J R, J W Johnson, and S A Schaaf. The Force Exerted by Surface Waves on Piles.” *Journal of Petroleum Technology* 2 (05): 149–54, (1950). <https://doi.org/10.2118/950149-g>.

Author Index

A

Adugna, Yosef W. 143, 254
Aktosun, Erdem 187, 225

B

Bai, Jiang-Bo 111
Bao, Yuequan 18
Benniu, Zhang 58
Bracke, Jona 327
Brüggemann, Heinrich 143

C

Casal, Gerardo 401
Castro, Alberte 401
Chen, Peng 349

D

De Backer, Hans 327
De Waele, Wim 327
Djuv, Ane 143
Drakoulas, George 28

E

EL Farragui, Meryieme 385

F

Fantuzzi, Nicholas 111
Fragkoulis, Dimitrios G. 201, 213

G

Gang, Zheng 58
Georgiou, Georgios C. 377, 385
Ghorbani, Maryam 280
Giljarhus, Knut Erik Teigen 241, 280
Ginigaddara, Thusitha 93
Gortsas, Theodore 28

H

Hafnor, Erik 143
Hågbo, Trond-Ola 241

Hagland, Hanne R. 143
Hectors, Kris 327
Huang, Xiaoping 298
Hussen, Kedir 155

J

Jayasinghe, Thushara 93
Jigang, Yu 58
Jung, Yunho 172

K

Kato, Tsuyoshi 80
Katsaros, Konstantinos P. 3
Kim, Joo-Hyung 172
Koumboulis, Fotis N. 201, 213
Kurukkal, Navaneethan S. 254
Kuwabara, Yuto 80

L

Lamsikine, Hind 377
Lemu, Hirpa G. 143, 155, 254, 268
Li, Hui 18, 124
Liu, Tian-Wei 111

M

Maruyama, Taizo 359
Mendis, Priyan 93
Miki, Akihiro 359

N

Nakahata, Kazuyuki 359
Nikolakopoulos, Pantelis G. 3
Nikolakopoulos, Pantelis 311

P

Papadopoulos, Dimitrios 213
Paulsen, Aksel 143
Pavlou, Dimitrios 298, 421
Pezeshki, Hadi 421
Polyzos, Demosthenes 28

Q

Qian, Wenliang 124

R

Regassa, Yohannes 268

S

Saelens, Lien 327

Saitoh, Takahiro 80

Santamarina, Duarte 401

Shin, Kyeong-ho 172

Siriwardane, Sudath C. 421

Skadsem, Hans Joakim 280

Souhar, Otmane 377, 385

Sun, Ceshi 349

T

Totomis, Panayotis G. 201

Tselios, Ioannis 311

Tsinopoulos, Stephanos 28

Tura, Amanuel Diriba 155

Tzamtzi, Maria P. 201, 213

V

Vázquez-Méndez, Miguel E. 401

Vinogradov, Alexey 45

W

Wang, Qingbo 45

Wang, Yu 45

X

Xiros, Nikolaos I. 187, 225

Xu, Yang 18

Y

Yu, Jigang 349

Z

Zeng, Chuan 349

Zheng, Gang 349

Zhuangzhuang, Gu 58