İSTANBUL UNIVERSITY
– CERRAHPAŞA

ACT

CONFERENCE PROCEEDINGS

ICASE 2026

INTERNATIONAL CONFERENCE ON APPLIED SCIENCES AND ENGINEERING

Innovation and Intelligence for a Sustainable Future



9–10 NOVEMBER 2026

FULLY ONLINE • ISTANBUL, TÜRKİYE (UTC+3)

EDITED BY

GEBRAIL BEKDAŞ

SINAN MELIH NIGDELI

ÜMIT IŞIKDAĞ

Organized by Istanbul University-Cerrahpaşa, Mimar Sinan Fine Arts University and the Advanced Computational Techniques (ACT) Research Group

ISBN: 978-625-00-5623-3

CONFERENCE PROCEEDINGS

International Conference on Applied Sciences and Engineering: ICASE 2026

Innovation and Intelligence for a Sustainable Future

9–10 November 2026 · Online

Istanbul, Türkiye (UTC+3)

EDITED BY

Prof. Dr. Gebrail Bekdaş — Istanbul University-Cerrahpaşa, Türkiye

Prof. Dr. Sinan Melih Nigdeli — Istanbul University-Cerrahpaşa, Türkiye

Prof. Dr. Ümit Işıkdağ — Mimar Sinan Fine Arts University, Türkiye

ORGANIZED BY

Advanced Computational Techniques (ACT) Research Group
Istanbul University-Cerrahpaşa
Mimar Sinan Fine Arts University

ISBN: 978-625-00-5623-3

Conference website: https://icase.iuc.edu.tr/en/
ACT Research Group: https://act.iuc.edu.tr

SCAN TO VISIT

First published in November 2026.

The papers and abstracts in this volume are published as submitted by the authors;
the authors are responsible for the content of their contributions.

ICASE 2026

Organizing Committee

CONFERENCE CHAIRS AND EDITORS

Prof. Dr. Gebrail Bekdaş	Istanbul University-Cerrahpaşa, Türkiye
Prof. Dr. Sinan Melih Nigdeli	Istanbul University-Cerrahpaşa, Türkiye
Prof. Dr. Ümit Işıkdağ	Mimar Sinan Fine Arts University, Türkiye

SECRETARIAT

Dr. Ayla Ocak	İstanbul Gelişim University, Türkiye
Yaren Aydın	Istanbul University-Cerrahpaşa, Türkiye

Special Track Organizers

TRACK 1

Data-Driven Surrogate Models for Structural Applications

Javier Fernando Jiménez-Alonso	University of Seville, Spain
Javier Naranjo-Pérez	University of Seville, Spain

TRACK 2

Structural Engineering Innovations

Kada Abdelhak	Hassiba Benbouali University of Chlef, Algeria
Mohamed Ghrici	Hassiba Benbouali University of Chlef, Algeria

TRACK 3

Emerging Trends and Technologies in Naval Architecture, Marine Engineering and Maritime Sciences

Barış Barlas	Piri Reis University, Türkiye
Gülçin Çivi Bilir	Istanbul Technical University, Türkiye

TRACK 4

Data-Driven Circular Built Environment

Burcu Kısmet Conk	Özyeğin University, Türkiye

ICASE 2026

Scientific Committee

Abdul Samed Kazi
VTT Technical Research Centre, Finland

Ali Cemal Benim
Duesseldorf University, Germany

Alias Abdul Rahman
Universiti Teknologi Malaysia, Malaysia

Anupam Yadav
NIT Jalandhar, India

Aylin Ece Kayabekir
İstanbul Gelişim University, Türkiye

Barış Barlas
Piri Reis University, Türkiye

Camila Martins Saporetti
Rio de Janeiro State University, Brazil

Celal Cakiroglu
Eastern Michigan University, USA

Dario De Domenico
University of Messina, Italy

David Arditi
Illinois Institute of Technology, USA

Farzad Rahimian
Teesside University, UK

Gülçin Çivi Bilir
Istanbul Technical University, Türkiye

Ibrahim Aykut Olcer
World Maritime University, Sweden

Javier Fernando Jiménez Alonso
University of Seville, Spain

Joong Hoon Kim
Korea University, Republic of Korea

Marzia Bolpagni
Northumbria University and UCL, UK

Maziar Farzam
University of Maragheh, Iran

Melda Yücel
Istanbul Aydın University, Türkiye

Moacir Kripka
University of Passo Fundo, Brazil

Mohamed Ghrici
Hassiba Benbouali University of Chlef, Algeria

Mohammed Rashad Baker
University of Kirkuk, Iraq

Nassim Djedoui
Biskra University, Algeria

Paul Coates
University of Salford, UK

Said Elias Rahimi
Leibniz University Hannover, Germany

Salah Djerouni
University Mohamed Khider Biskra, Algeria

Salman Azhar
Auburn University, USA

Sanghun Kim
Temple University, USA

Sefer Anil Gunbeyaz
University of Strathclyde, UK

Serdar Ulusoy
Turkish-German University, Türkiye

Serdar Yıldız
Sharjah Maritime Academy, UAE

Seyedvahid Vakili
University of Southampton, UK

Xin-She Yang
Middlesex University, UK

Yasser E. Ibrahim
Prince Sultan University, Saudi Arabia

Zong Woo Geem
Gachon University, Republic of Korea

Zuzana Dimitrovová
Nova University of Lisbon, Portugal

35 members, listed in alphabetical order.

ICASE 2026

Preface

It is our great pleasure to present the Proceedings of the International Conference on Applied Sciences and Engineering (ICASE 2026), held fully online on 9–10 November 2026. The conference was organized by Istanbul University-Cerrahpaşa and Mimar Sinan Fine Arts University together with the Advanced Computational Techniques (ACT) Research Group. Under the theme “Innovation and Intelligence for a Sustainable Future”, ICASE 2026 brought together researchers, practitioners and students to present scientific and technological research and to discuss current needs and future trends in applied sciences and engineering.

The programme was organized in six parallel virtual rooms: four special tracks — Data-Driven Surrogate Models for Structural Applications (Track 1), Structural Engineering Innovations (Track 2), Emerging Trends and Technologies in Naval Architecture, Marine Engineering and Maritime Sciences (Track 3) and Data-Driven Circular Built Environment (Track 4) — and a General Session held in two parallel parts. Each special track opened with an invited keynote lecture, delivered by Javier Fernando Jiménez-Alonso (University of Seville), Kada Abdelhak (Hassiba Benbouali University of Chlef), Utku Cem Karabulut (Bandırma Onyedi Eylül University) and Hexin (Johnson) Zhang (Edinburgh Napier University), respectively.

This volume collects 59 contributions from authors in 16 countries: 47 full papers and 12 abstracts. Submissions were evaluated through the conference review process, which was managed with the Microsoft CMT system. The book is organized by track; each track section opens with its organizers and keynote speaker and is followed by the full papers and then the abstracts, in order of presentation.

In line with an open and inclusive approach to science, registration and publication costs of the conference were covered by the ACT Research Group. Selected extended papers will be considered for special issues of international journals.

We would like to express our sincere gratitude to all authors for their contributions, to the keynote speakers, the special track organizers and the session chairs, and to the members of the Scientific Committee and the reviewers for their valuable time and expertise. We hope that these proceedings will serve as a useful reference and will encourage further collaboration among researchers.

Warm regards,

Prof. Dr. Gebrail Bekdaş
Prof. Dr. Sinan Melih Nigdeli
Prof. Dr. Ümit Işıkdağ

Editors and Conference Chairs, ICASE 2026

Istanbul, November 2026

ICASE 2026

Table of Contents

TRACK 1 · DATA-DRIVEN SURROGATE MODELS FOR STRUCTURAL APPLICATIONS 10

FULL PAPERS

Nonlinear Dynamic Analysis of Stay Cables via Single-Mode Reduced-Order Modeling: Application to the Alamillo Bridge 11
María de la Torre Coca, Javier Naranjo-Pérez, Javier Fernando Jiménez-Alonso

Deep Learning-Based Surrogate Modeling for Multi-Objective Structural Size Optimization under Mixed-Integer Constraints 19
José Manuel Amuedo Jiménez, Javier Naranjo-Pérez, Luis Rodríguez-Tembleque, Javier Fernando Jiménez-Alonso

Neural Ordinary Differential Equations as Surrogate Models for Predictive Control of Non-Linear Dynamic Systems: Application to a Reaction Wheel Pendulum 28
Andy El Guarti Msaouri, Miguel Carrión Colchero, Javier Fernando Jiménez-Alonso

Proportional Control Tuning for Thermal Digital Twins of Asphalt Pavements: A Numerical Study 36
Javier Álvarez Pozo, Javier Fernando Jiménez-Alonso, Andrés Sáez

TRACK 2 · STRUCTURAL ENGINEERING INNOVATIONS 44

FULL PAPERS

An Interactive Tool for Temperature-Time Curve Prediction: Towards a Safety Strategy for Structures in Fire 45
Romaissa Hachana, Abdallah Benarous, Happy Sinkala, Selahaddin Orhan Akansu

Extremely Randomized Trees Approach for Predicting the Axial Capacity of FRP-Reinforced Concrete Compressive Members 53
Sarra Sendjasni, Klaus Holschemacher, Mohammed Berradia, Riad Benzaid, Asmaa Djilali Bouzina

Evaluation of thermal safety distances during fuel tank fires 62
Souhila Agred, Abdallah Benarous, Happy Sinkala, Selahaddin Orhan Akansu

Numerical Assessment of Steel Frame Connections Under Earthquake-Induced Post-Fire Loading 70
Ouissam Yessad, Abdelhak Kada

Free Vibration Analysis of Simply Supported FG-CNT Plates Resting on a Kerr Elastic Foundation 77
Asmaa Djilali Bouzina, El-Hadj Meziane, Riad Bennai, Sarra Sendjasni

Interpretable Machine Learning for Predicting the Compressive Strength of FRP-Confined Concrete: A Rigorous Data-Curation and Bayesian-Optimization Framework with TabPFN 88
Ali Benzaamia, Mohamed Ghrici

Numerical Investigation of the Behavior of Cold-Formed Steel Profiles with Web Openings under Fire Exposure 96
Sarra Taibi, Abdelhak Kada, Mohammed Hassoune, Mus'ab Salameh

Prediction of the Chloride Permeability of Fly Ash-Containing Self-Compacting Concrete Using an Optimized XGBoost Model 103
Fatima Kechroud, Ali Benzaamia, Mohamed Ghrici

Machine Learning-Based Prediction of Seismic Response and Fragility Curves for Reinforced Concrete Structures 111
Nesrine Atallah, Kawther Arbia Boudjelthia, Zamila Harichane, Halima Belbahi

Machine Learning-Based Prediction of the Nonlinear Seismic Response of Reinforced Concrete Frames 119
Yassamine Zahia Klouche, Zamila Harichane, Halima Belbahi

Multi-Layer Neural Network Classifier for Earthquake Building Damage Prediction 127
Halima Belbahi, Hichem Noura, Mohamed Abed

Table of Contents (continued)

ABSTRACTS

Structural Design of Mid-Rise RC Buildings in Abuja Subjected to Abuja-Specific Wind Loads 133
Ifeyinwa Ijeoma Obianyo, Fisayo Praise Olatunji, Anthony Muoka, Sanusi Abdulganiyu, Abdulhameed Danjuma Mambo

Assessing the Accuracy of the Bi-Objective Optimization-Based SonReb Method for Nondestructive In-Situ Concrete Strength Estimation: A Real Case Study 134
Faiza Neggaz, Khoudja Ali-Benyahia, Mohamed Ghrici

TRACK 3 · EMERGING TRENDS AND TECHNOLOGIES IN NAVAL ARCHITECTURE, MARINE ENGINEERING AND MARITIME SCIENCES 135

FULL PAPERS

Hydrodynamic Control and Diving Performance in the Conceptual Design of a Semi-submersible Naval Platform 136
Utku Cem Karabulut, Barış Barlas KEYNOTE

An Investigation of Forces and Moments Acting on Semi-Standardized and Non-Standardized Cargo Units Aboard Ships 151
Fatih Kolçak, Hakan Akyıldız

A Fuzzy DEMATEL Approach for Benefit–Risk Assessment in Sports and Recreational Activities 173
Ayhan Menteş, Hakan Akyıldız, Şebnem Helvacıoğlu

Evidence to Investigation Report: An IMO-Standards-Grounded Human-in-the-Loop Multi-Agent Framework for Marine Casualty Investigation 202
Serdar Yıldız, Özkan Uğurlu, Barış Barlas

A Multi-Objective Spatial Risk Optimization Framework for Safe, Sustainable and Efficient Shipyard Layout Planning 213
Tolga Aycı, Barış Barlas

Multi-Objective Optimization of Marine Current Turbine Blade Geometry for Enhanced Hydrodynamic Efficiency and Structural Strength 234
Tunahan Şık, Engin Kanun, İsmail Kahraman, Selahattin Özsayan, Mehmet Delikan, Burcu Erbaş Öz, Uğur Oral Ünal, Feyza Ertarhanacı

Orbital Welding of Stainless-Steel Pipes in the Nuclear Industry: History, Technological Development, Applications, and Future Perspectives 241
Gursel Yildiz, Barış Barlas

Securing Acoustic Communication for Underwater Drones Based on the Primitive Vector Cipher (PVC) 255
Gülçin Çivi Bilir, Volkan Ateş, Barış Barlas

ABSTRACTS

Full-Wave Design and Broadband Characterization of a Center-Fed Planar Elliptical Dipole Antenna for 3.1–10.6 GHz Ultra-Wideband Communication 268
Zakaria Mesmoudi

Design, Simulation and Physical Performance Analysis of an Ultra-Wideband Cavity-Backed Log-Periodic Folded-Slot Dipole Array Antenna for UHF/HF and EMC Applications 269
Zakaria Mesmoudi

TRACK 4 · DATA-DRIVEN CIRCULAR BUILT ENVIRONMENT 270

FULL PAPERS

Mapping Artificial Intelligence Applications for a Circular Economy in the Built Environment: A Bibliometric Analysis 271
Kudzai Mbowa, Clinton Aigbavboa

Table of Contents (continued)

ABSTRACTS

Bamboo–Timber Composite Gridshell: Design, Construction and Full-Scale Experimental Investigation 279
Yanghao Pei, Hexin Zhang

Bauhaus Earth – Towards a Systemic Transformation of the Built Environment 280
Patricia Jeglitsch

Circularity Across Scales and Disciplines. A New Education Programme for the Built Environment in the Netherlands 281
Bianca Andaloro, Poorya Eghtesadi, Cecilia Chiappini, Nurhan Abujidi

Adaptive Reuse for Urban Food Waste Collection: Hub Location and Routing Optimization 282
Esma Selen Aksoy

GENERAL SESSION 283

FULL PAPERS

The Effect of Isolator Mobility in Base Isolation Systems Optimized by Accounting for Soil-Structure Interaction 284
Ayla Ocak, Sinan Melih Nigdeli, Gebrail Bekdaş

Evolution of Artificial Intelligence Paradigms and Emerging Architectures for Geohazard Prediction 295
Hao Xu

Microwave-Induced Fracturing and Drillability Bottleneck Breakthrough in Deep Hard Rock: A Multiphysics Coupling and Geomechanical Perspective 303
Jiawei Lv

AI-Based Prediction of Mechanical Properties of Microwave-Pretreated Rock and Optimization of Rock Breaking 313
Zhuo Pan

Research on a Dynamic Site Selection and Evaluation System for Underground Gas Storage Based on Hydro-Mechanical Coupling Dynamics 321
Miao Zhang

Exploring Correlation of Body Weight with Depth of 3D Footwear Impression for Forensic Investigations 336
Anwar Ama Salem, Loong Chuen Lee

Synthesis and Characterization of Fe3O4/rGO/Ag Nanocomposite in Extracting Opioids 344
Enas Amoori Hadi Hadi, Wijdan Shakir Khayoon, Loong Chuen Lee

Application of Image Processing in the Discrimination of Red Fabrics for Forensic Investigations 352
Wan Adam Harris Wan Azmirrudin, Loong Chuen Lee

Self-Attention-Enhanced 3D Gaussian Splatting for Deformable Endoscopic Surgical Scene Reconstruction 359
Veena K V, Supriya M H

A Critical Analysis of Artificial Intelligence Adoption in Higher Education 372
Samuel Adekunle, Solace Ameyibor, Simon Nyema, Temitope Olajide, Olajide Oluwatosin, Daniel Katu, Babatunde Ogunbayo

Integrated Urban Wind-Field Assessment and Predictive Control for Autonomous Systems under Multi-Directional Inflow: A Downtown Tulsa Case Study 380
Ahmed Ghareeb, Jamey Jacob, Kursat Kara

Trustworthy Geometry-to-Polar Learning for Unseen Airfoils: OOD Generalization, Uncertainty, and Physics-Aware Surrogates 388
Mohammed Alharbi, Ahmed Ghareeb, Nuha Alharbi

Data-Driven Grid-Aware Coordination of Residential PV–Battery Systems under Intermittent Electricity Supply 396
Abdulmajeed Mustafa, Fatimah Mustafa, Ahmed Ghareeb

Mapping the Evolution of Optimization Research on Seismic Isolation Systems through Bibliometric Analysis 405
Fatih Tarık Sarı

Table of Contents (continued)

Evaluation of the Performance and Generalizability of the TabPFN Under Different Data Splitting Strategies for Bond Strength Prediction 414
Yaren Aydın, Sinan Melih Nigdeli, Gebrail Bekdaş, Ümit Işıkdağ

Performance of the Competition-Amensalism Optimization (CAO) Algorithm in RC Beam Cost Optimization Under Different Cost Scenarios and Comparison with Other Metaheuristic Algorithms 422
Yaren Aydın, Sinan Melih Nigdeli, Gebrail Bekdaş

A Photo to 3D Workflow Using TRELLIS.2 for Generative Architectural Representation 430
Handan Aş Çemrek, Ümit Işıkdağ, Gebrail Bekdaş, Sinan Melih Nigdeli

Risk Management Maturity in Public Transportation Digitalisation 439
Kudzai Mbowa, Clinton Aigbavboa

Performance and robustness assessment of optimized tuned mass dampers in soft-story structures 447
Farah Alnayhoum, Sinan Melih Nigdeli, Gebrail Bekdaş

Physics-Informed Learning of Anisotropic Epidemic Dynamics from Chicago Bike-Share Mobility 457
Rasha Almshekhs, Suhaib Abdul Sattar Shukur

Different Generations of Metaheuristic Algorithms in Structural Optimization: Tubular Column and I-Beam Applications 465
Yaren Aydın, Gebrail Bekdaş, Sinan Melih Nigdeli

Machine Learning-Based Prediction of Concrete Compressive Strength Using a Multi-source Modified Concrete Dataset 473
Yaren Aydın, Gebrail Bekdaş, Sinan Melih Nigdeli, Ümit Işıkdağ

Particle-Tuned Mass Dampers for Structural Vibration Control: A Review of Modeling, Design and Applications 482
Aşan Şenlik, Yaren Aydın, Sinan Melih Nigdeli, Gebrail Bekdaş

ABSTRACTS

Challenges and Implications of Project Management Practices Implementation for Engineering Project Performance in the Liberian Construction Industry 492
Martins Randoulf, Samuel Adekunle

The Use of Plantain Leaf Ash as a Natural Pozzolan in Sandcrete Blocks for Sustainable Buildings 493
Ifeyinwa Ijeoma Obianyo, Chinonso Obika, Uwemedimo Nyong Wilson, Assia Aboubakar Mahamat, Azikiwe Peter Onwualu

Evaluating Stormwater Drainage Systems in Urban Roads: A Case Study of Abuja Ring Road 2 494
Ifeyinwa Ijeoma Obianyo, Muhammad Muazu Ibrahim, Abubakar Dayyabu, Assia Aboubakar Mahamat, Martin Agbili

Performance Evaluation of Interlocking Paving Blocks Produced with Ground Low-Density Polyethylene (LDPE) as a Partial Fine Aggregate Replacement 495
Ifeyinwa Ijeoma Obianyo, Muhammad Abu-Imran Ozeto, Abubakar Dayyabu, Akeem Amuda, Abdulhameed Danjuma Mambo

TRACK 1

Data-Driven Surrogate Models for Structural Applications

SPECIAL TRACK ORGANIZERS

Javier Fernando Jiménez-Alonso University of Seville, Spain
Javier Naranjo-Pérez University of Seville, Spain



KEYNOTE SPEAKER

Assoc. Prof. Javier Fernando Jiménez-Alonso

University of Seville, Spain

Reduced-Order Models for the Nonlinear Dynamics of Cables

Javier Fernando Jiménez-Alonso is an Associate Professor of Department of Continuum Mechanics and Structure at the Higher Technical School of Engineering, University of Seville, Spain. With over two decades of professional and academic experience in bridge engineering, his expertise spans vibration engineering, structural health monitoring, structural control, and structural optimization. He holds a PhD in Civil Engineering (2015) and an MSc in Applied Mathematics (2019) from the University of Seville, as well as an MSc in Civil Engineering (2000) from the University of Granada.

His research contributions focus on finite element model (FEM) updating, dynamic response optimization, and the control of structural vibrations. Notably, his recent work includes applying cooperative and evolutionary game theory to the numerical model updating of complex structures, including cable-supported systems such as moveable cable-stayed bridges. He has authored over 50 scientific contributions in high-impact international journals and has served as a key researcher in multiple projects funded by the Spanish Ministry of Science, Innovation, and Universities, focusing on dynamic loads, sustainability, and intelligent maintenance of transportation infrastructures. He has carried out international research stays at the University of Porto (Portugal), the University of Exeter (UK), and University College Dublin (Ireland).

TRACK TOPICS

- Physics-Informed Neural Networks (PINNs) in Structural Mechanics
- Reduced Order Modeling (ROM) and Model Order Reduction (MOR)
- Machine Learning and AI for Structural Health Monitoring (SHM)
- Digital Twins and Cyber-Physical Structural Systems
- Surrogate-Based Structural Optimization and Topology Optimization
- Uncertainty Quantification and Reliability Analysis in Mechanics
- Data-Driven Damage Detection, Localization, and Assessment
- Finite Element Model Updating via Machine Learning
- Deep Learning for Nonlinear Structural Dynamics and Vibration Control
- Multi-Fidelity and Multi-Scale Surrogate Modeling
- Gaussian Process Regression, Kriging, and Meta-Modeling Techniques
- Real-Time Structural Response Prediction and Hybrid Testing
- Operational Modal Analysis and Data-Driven System Identification
- Data Fusion and Smart Sensor Integration for Structural Models
- Generative AI, Transfer Learning, and Domain Adaptation in Structural Assessment
- Predictive Maintenance and Lifecycle Assessment of Civil & Mechanical Structures

Nonlinear Dynamic Analysis of Stay Cables via Single-Mode Reduced-Order Modeling: Application to the Alamillo Bridge

María de la Torre Coca [1], Javier Naranjo-Pérez[1[0000-0003-2239-7048]], and Javier Fernando Jiménez-Alonso[1[0000-0002-4592-0375]]

[1] Higher Technical School of Engineering, University of Seville, Seville 41092, Spain

Abstract. Stay cables in cable-stayed structures exhibit complex nonlinear dynamic behavior governed by sag-extensensibility effects, geometric nonlinearities, and large displacements under environmental excitations. While high-fidelity finite element models provide detailed predictions, their high computational cost limits their application in long-time integrations, uncertainty quantification, or real-time monitoring. This work presents the development and numerical validation of a low-dimensional reduced-order model tailored to efficiently capture the nonlinear dynamics of stay cables. The methodology is demonstrated on a representative stay cable of the Alamillo Bridge in Seville, spanning 292 m with an inclination angle of 26º. First, a benchmark high-fidelity finite element model is developed in commercial finite element analysis package to establish the static catenary equilibrium under pre-tension and gravity, accounting for geometric stiffness. Subsequently, a single-mode reduced-order model is formulated in a mathematical package by projecting the governing nonlinear equations onto the fundamental mode within the local Frenet frame of the equilibrium line. The resulting reduced system generalizes the classical Duffing oscillator by incorporating a quadratic nonlinear term to account for the gravitational asymmetry inherent to inclined cables, alongside cubic nonlinear stiffness terms. The reduced-order model performance is rigorously evaluated against the finite element benchmark across multiple regimes: static equilibrium deflection, natural frequency evaluation, modal assurance criterion, nonlinear free vibration, and forced response under turbulent wind loads. The results confirm that the proposed reduced-order model accurately reproduces nonlinear phenomena such as frequency shifts and phase portraits while reducing computational time drastically, demonstrating its effectiveness for structural engineering applications.

Keywords: Reduced-order modelling, stay cables, nonlinear vibrations, finite element method, geometric nonlinearity, turbulent wind loads.

1 Introduction

In current structural engineering, cable-stayed bridges represent essential infrastructure for spanning large distances efficiently. Within these structures, stay cables are key load-bearing components designed to operate under tension with minimal material usage. However, this structural efficiency comes at the expense of high flexibility, low inherent damping, and negligible flexural rigidity, rendering stay cables extremely vulnerable to adverse environmental excitations such as wind, rain-wind interactions, and support motions [1]. Under such operational conditions, stay cables exhibit highly complex nonlinear dynamic behavior driven by sag-extensensibility effects, geometric nonlinearities under large displacements, and strong nonlinear restoring forces. Accurately modeling these dynamic phenomena is crucial to evaluate structural fatigue, prevent catastrophic instabilities, and design mitigation or damping strategies.

To simulate the nonlinear dynamics of stay cables, high-fidelity finite element (FE) models are widely employed as numerical benchmarks [2]. FE formulations incorporate geometric stiffness, initial catenary static equilibrium under pretension and gravity, and nonlinear strain-displacement relationships. Despite their accuracy, high-fidelity FE models demand thousands of global degrees of freedom and require computationally intensive step-by-step implicit time-integration. Hence, their computational cost restricts their practical application in scenarios requiring thousands of evaluations, such as real-time structural health monitoring, long-time wind response integration, or multi-query uncertainty quantification.

To overcome these computational limitations, Reduced-Order Models (ROMs) and surrogate modeling techniques have emerged as efficient alternatives by projecting governing partial differential equations onto low-dimensional subspaces [3, 4]. In cable dynamics, projecting the nonlinear equations of motion onto the fundamental mode within the local Frenet frame of the static equilibrium catenary yields a low-dimensional oscillator [5]. Unlike the classical Duffing oscillator, which features only linear and cubic stiffness terms, the single-mode formulation for inclined cables under gravity intrinsically includes a quadratic nonlinear stiffness term [6]. This term breaks the symmetric restoring force of traditional models, capturing the positive-negative asymmetry induced by gravity-sag coupling.

While such ROM formulations offer significant speedups, rigorous numerical validation against high-fidelity FE benchmarks across diverse dynamic regimes remains essential to demonstrate their reliability in real-world civil infrastructure. To address this challenge, this paper presents the development and numerical validation of a single-mode reduced-order model intended to capture the nonlinear dynamics of stay cables. The methodology is applied to the longest stay cable of the Alamillo Bridge in Seville. A benchmark FE model developed in ANSYS [7] serves as the reference for static catenary equilibrium and modal features. Subsequently, the single mode ROM is formulated in MATLAB [8] via Galerkin projection onto the fundamental mode. The performance of the ROM is thoroughly evaluated against FE results across static deflection, natural frequencies, modal correlation (MAC), nonlinear free vibration, and forced response under turbulent wind loading.

The remainder of this paper is organized as follows: Section 2 outlines the theoretical foundations of stay cable dynamics; Section 3 presents the generation of the reduced-order model; Section 4 details the benchmark Alamillo Bridge stay cable and evaluates comparative FE-ROM results; and Section 5 presents concluding remarks.

2 Basics of nonlinear dynamics of stay cables

Stay cables are primary load-bearing components in cable-stayed bridges optimized to carry large axial tension loads. However, their high flexibility, low damping, and negligible flexural rigidity result in complex nonlinear dynamic responses governed by Irvine's elasto-geometric parameter λ^2, sag-extensibility effects, and geometric nonlinearities under large displacements.

The dynamic response of an inclined cable under pre-tension T_0 and self-weight is governed by the coupling between transverse motions and dynamic tension variations. Unlike ideal taut strings, inclined cables under gravity adopt a catenary equilibrium profile, introducing gravitational asymmetry. When vibrating around this equilibrium, the nonlinear strain-displacement relation yields both quadratic and cubic restoring forces. The quadratic terms account for positive-negative motion asymmetry caused by gravity-sag coupling, while cubic terms represent nonlinear stretching stiffness. This dynamic behavior involves amplitude-dependent natural frequencies, internal resonances, and jump bifurcations.

In structural design and multi-objective optimization (such as minimizing mass while maximizing fatigue life and damping), evaluating these nonlinear phenomena is critical. Operational constraint violations—such as local cable slackening ($T(S,t) < 0$), excessive deflection, and dynamic stress ranges—must be evaluated over long integration times across broad search spaces. Evaluating these constraints with high-fidelity FE models is computationally unfeasible for iterative optimization loops. Thus, low-dimensional ROMs are required to enable scalable optimization and near-instantaneous constraint evaluation [9, 10].

3 Reduced-order models of stay cables

To overcome the high computational cost of high-fidelity FE simulations for the optimum design of cables, low-dimensional ROMs provide an efficient mathematical surrogate. The primary objective of a ROM is to approximate the continuous spatial-temporal dynamics of the cable governed by nonlinear partial differential equations using a reduced set of ordinary differential equations with minimal degrees of freedom.

3.1 Kinematics in the local Frenet frame and single mode projection

The nonlinear motion of an inclined stay cable under pre-tension T_0 and self-weight is formulated along its curvilinear arc length $S \in [0,L]$. To represent the catenary curvature without introducing Cartesian displacement errors, the motion is projected onto the local Frenet triad consisting of the unit vectors tangent, normal, and binormal. In this work, the transverse dynamic displacement field perpendicular to the cable, $u(S,t)$, is expanded onto the fundamental spatial mode shape $\Phi(S)$:

$$u(S,t) = \Phi(S) \cdot q(t) \tag{1}$$

where $\Phi(S) = \sin(\pi S/L)$ represents the fundamental mode shape normalized to unit amplitude at mid-span, and $q(t)$ denotes the time-dependent modal coordinate. By multiplying the governing nonlinear continuum equations by the test function $\Phi(S)$ and integrating along the span length L, the spatial dependence is eliminated, yielding a single-degree-of-freedom (1-DOF) nonlinear ordinary differential equation:

$$M_n \ddot{q}_n(t) + C_n \dot{q}_n(t) + K_l q_n(t) + Q_{nnn} q_n^2(t) + C_{nnnn} q_n^3(t) = f_{gen}(t) \tag{2}$$

where M_n, C_n, and K_l are the linear modal mass, damping, and linear stiffness, respectively. The generalized force $f_{gen}(t)$ is obtained by projecting external dynamic loads (such as turbulent wind forces) onto the fundamental mode. Q_{nnn} and C_{nnnn} are the quadratic and cubic stiffness terms, respectively.

3.2 Nonlinear stiffness tensors and asymmetry mechanics

The governing ROM equation generalizes the classical Duffing oscillator. The cubic stiffness term C_{nnnn} captures the nonlinear geometric stretching stiffness of the longitudinal axis during large-amplitude vibrations and the quadratic stiffness term Q_{nnn} represents the asymmetry of the restoring force caused by the cable's initial sag and the nonlinear variation in static tension along its length, where downward displacements (driven by gravity) have a lower effective stiffness than upward displacements (which are opposed by gravity). These terms are calculated as follows:

$$\begin{aligned} Q_{nnn} &= \frac{EA}{2} \int_0^L \kappa(S) \left(\frac{d\Phi_n}{dS} \right)^2 \Phi_n(S)\, dS \\ C_{nnnn} &= \frac{EA}{2} \int_0^L \kappa(S) \left(\frac{d\Phi_n}{dS} \right)^4 dS \end{aligned} \tag{3}$$

where EA denotes axial stiffness and $\kappa(S)$ is the static catenary curvature induced by gravity.

The reduced 1-DOF equation can be integrated in fractions of a second. This substantial reduction in execution time makes single-mode ROM ideal for optimization applications. Specifically, the ROM enables near-instantaneous evaluations of state-dependent constraint violations (such as local tension loss $T(S,t) < 0$ or maximum deflection thresholds) across the design candidates. Consequently, it allows scalable exploration of multi-objective Pareto fronts in mixed-integer design spaces, providing a robust tool for structural design, uncertainty quantification, and real-time health monitoring.

4 Application example

The ROM was assessed against a nonlinear FE model of the longest stay cable of the Alamillo Bridge [1] in Seville (span $L = 292$ m, inclination $\theta = 26^{\circ}$, horizontal projection 263.6 m, vertical anchorage height 130 m). Properties: cross-sectional area $A = 8.38 \cdot 10^{-3}$ m^2, moment of inertia $I = 5.58 \cdot 10^{-4}$ m^4, mass per unit length $m = 60$ kg/m, Young's modulus $E = 1.9 \cdot 10^{11}$ N/m^2, and initial pretension $H = 4.13 \cdot 10^{6}$ N. The FE model was built in ANSYS [7] using 100 BEAM188 elements, restrained against translation/torsion with free in-plane bending. Geometrically nonlinear static analysis established static catenary equilibrium with axial tension $T_0 = 4.92 \cdot 10^{6}$ N.

4.1 Static and modal validation

Static validation was performed by reconstructing the catenary equilibrium shape using a 6-mode sinusoidal expansion. The ROM reproduced the FE model equilibrium profile with a maximum absolute error of 45.2 mm and a global relative error of only 1.128%, confirming that the reduced basis accurately captures the static catenary deformation.

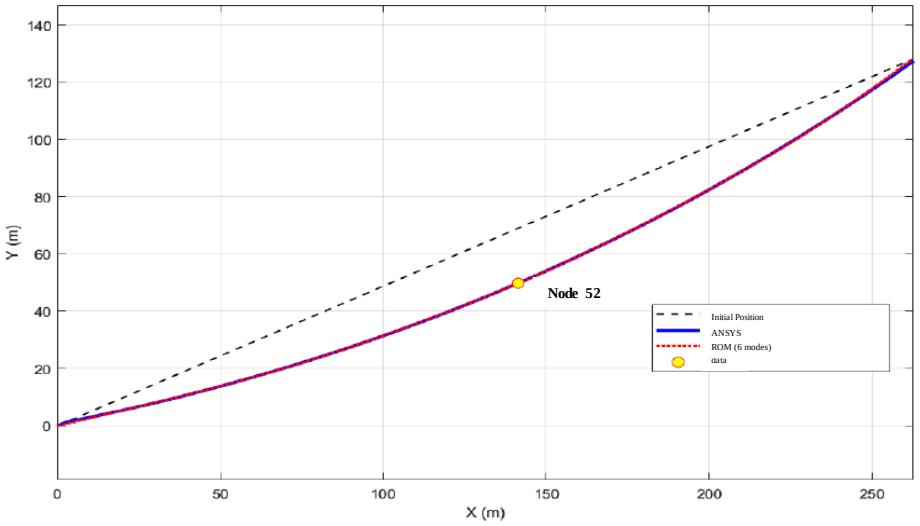


Fig. 1. Static catenary equilibrium profile: comparison between FEM reference (ANSYS) and 6-mode ROM reconstruction.

The dynamic fidelity of the ROM basis was evaluated by comparing natural frequencies and spatial mode shapes (Table 1).

4.2 Nonlinear dynamic response

Single mode ROM parameters obtained via numerical integration are M_n = 8744.68 kg, K_l = 83230.15 N/m, Q_{111} = 604.47 N/m^2, C_{1111} = 1131 N/m^3, and C_n = 161.49 N · s/m. Analytical linear frequency is 0.4910 Hz (FE: 0.4894 Hz). Rayleigh damping ζ = 0.3 % was applied. Three free-vibration scenarios were tested:

(i) **Case A (Displacement perturbation,** u_0 = 0.05165 m**,** v_0 = 0 m/s**):** Mid-span displacement perturbation showed strong displacement correlation (r = 0.9915, NRMSE = 7.59%; dominant frequencies: FE 0.4951 Hz, ROM 0.4937 Hz, 0.283% error). Velocity NRMSE (16.04%) reflects pointwise initial excitation of higher FE modes (f_2 = 2.94 Hz) absent in 1-DOF ROM.

(ii) **Case B (Velocity perturbation,** u_0 = 0 m**,** v_0 = −0.1589 m/s**):** Modally distributed velocity injection concentrated energy in the fundamental mode, yielding displacement NRMSE of 10.09%, velocity NRMSE of 5.51%, and phase-space error of 5.93%, validating the single mode ROM.

(iii) **Case C (Combined perturbation,** u_0 = 0.05165 m**,** v_0 = +0.847 m/s**):** Simultaneous modal perturbation achieved displacement NRMSE of 9.87% (r = 0.9501) and velocity NRMSE of 5.47% (r = 0.9943). Negative displacement peak matched closely (FE: −0.2723 m, ROM: −0.2741 m, 0.64% error); positive peak divergence (FE: +0.0297 m, ROM: +0.084 m) stems from intense gravitational sag-coupling clipping beyond single-mode quadratic capacity.

Table 1. Comparison of natural frequencies (FE model and ROM) and MAC values.

Vibration mode	FE [Hz]	ROM [Hz]	MAC
1	0.4956	0.4893	0.9991
2	0.9788	0.9788	0.9965
3	1.4689	1.4683	0.9922
4	1.9590	1.9579	0.9862
5	2.4500	2.4476	0.9785
6	2.9417	2.9376	0.9691

4.3 Response under turbulent wind excitation

Sustained aerodynamic loading was simulated using ten turbulent wind velocity time histories spatially interpolated across the 99 internal nodes. Quasi-steady aerodynamic forces were formulated assuming air density ρ_a = 1.23 kg/m^3, drag coefficient C_D = 1.2, lift coefficient C_L = 0.3, and a mean basic wind speed of $\overline{U} = 26$ m/s. Generalized forces were projected onto the fundamental mode to drive the ROM.

Over a 30 s forced-vibration simulation, the vertical displacement perturbational response exhibited exceptional tracking, achieving a Pearson correlation coefficient of r = 0.9898 (Fig.2a). Restricting phase-plane comparison to the active band of the

ROM ($f < 0.6$ Hz) yielded a 97.1% trajectory overlap and a trajectory RMS error of 3.82% (Fig.2b). These results prove that the single mode ROM accurately captures the dominant nonlinear forced dynamics of stay cables under operational turbulent wind.

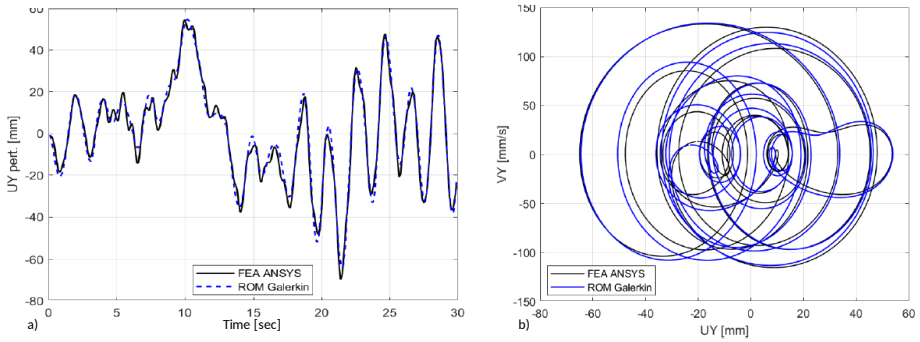


Fig.2. Forced vibration under turbulent wind: (a) 30-second displacement perturbation history, and (b) phase-plane trajectories.

Overall, the single mode ROM accurately reproduces dominant static and dynamic behavior with low global errors. Discrepancies stem from higher-mode contributions, out-of-plane dynamics, and large-amplitude sag clipping, which can be addressed in future multimodal extensions. relevant

5 Conclusions

This paper presented the development and numerical validation of a single-mode ROM for the nonlinear dynamic analysis of stay cables. Formulated via Ritz-Galerkin projection onto the fundamental mode within the local Frenet frame of the static catenary, the 1-DOF model generalizes the classical Duffing oscillator by introducing a quadratic stiffness term. This term accounts for the positive-negative motion asymmetry caused by gravity-sag coupling in inclined cables. The approach was compared with a high-fidelity nonlinear FE model of the longest stay cable of the Alamillo Bridge in Seville.

The static and linear modal validations confirmed high spatial and dynamic accuracy. The ROM reproduced the static equilibrium catenary with a global relative error of 1.128% and predicted the first six natural frequencies with a mean error of 0.2661% and MAC values exceeding 0.969. In nonlinear dynamic regimes, the ROM demonstrated strong agreement during free vibration, particularly under modally injected initial conditions (NRMSE = 5.51% for velocity). Under sustained turbulent wind loading, the ROM achieved exceptional displacement tracking ($r = 0.9898$, NRMSE = 2.80%) and a 97.1% trajectory overlap in the phase-plane within the dominant frequency band ($f < 0.6$ Hz).

The main discrepancies arise at large amplitudes, where severe gravitational sag-coupling causes response clipping exceeding single-mode representation, as well as from higher-mode velocity contributions and out-of-plane effects. Nevertheless, by reducing execution times from hours to fractions of a second, the ROM provides an efficient surrogate for near-instantaneous constraint verifications, real-time monitoring, long-time wind simulations, and multi-objective Pareto optimization across complex design spaces. Future work will focus on multimodal formulations and empirical bases.

Acknowledgements

This study has been funded by the Spanish Ministry of Science, Innovation and Universities through the project “Updated physics-informed models for the fatigue assessment of cables and external post-tensioning systems in bridges” under the grant agreement PID2024-157616OA-C22.

References

1. Naranjo-Pérez, J., Jiménez-Alonso, J.F., Díaz, I.M., Quaranta, G., Sáez, A.: Motion-based design of passive damping systems to reduce wind-induced vibrations of stay cables under uncertainty conditions. Applied Sciences 10(5): 1740 (2020).
2. Ereiz, S., Duvnjak, I., Jiménez-Alonso, J.F.: Review of finite element model updating methods for structural applications. Structures 41, 684-723 (2022).
3. Ruan, S., Class, A.G, Rozza, G.A.: Structured Review of Reduced Order Modeling for Domain Decomposition Problems: State of the Art and Perspectives. Archives of Computational Methods in Engineering https://doi.org/10.1007/s11831-026-10690-9 (2026).
4. Samadian, D., Muhit, I.B., Dawood, N.: Application of Data-Driven Surrogate Models in Structural Engineering: A Literature Review. Archives of Computational Methods in Engineering 32, 735–784 (2025).
5. Bertrand, C., Savadkoohi, A.T., Acary, V., Lamarque C.-H.: Reduced-Order Model for the Nonlinear Dynamics of Cables. Journal of Engineering Mechanics 148(9), 04022052-1-17 (2022).
6. Palazzolo, A.B., Shin, D., Falzarano, J.: Introduction to Engineering Nonlinear and Parametric Vibrations with MATLAB and Maple. 1st edn. John Wiley & Sons, Ltd, Chichester (U.K.) (2025).
7. ANSYS Homepage, http://www.ansys.com, last accessed 2026/10/15.
8. MATLAB Homepage, https://es.mathworks.com/products/matlab.html, last accessed 2026/10/15.
9. Forrester, A.I.J., Sóbester, A., Keane, A.J.: Engineering Design via Surrogate Modelling: A Practical Guide 1st edn. John Wiley & Sons, Ltd, Chichester (U.K.) (2008).
10. Kim, N-H.: Surrogate Modelling and Optimization. Theories, Applications and Limitations 1st edn. John Wiley & Sons, Ltd, Florida (U.S.) (2026).

Deep Learning-Based Surrogate Modeling for Multi-Objective Structural Size Optimization under Mixed-Integer Constraints

José Manuel Amuedo Jiménez[1], Javier Naranjo-Pérez[1[0000-0003-2239-7048]], Luis Rodriguez-Tembleque[1[0000-0003-2993-8361]] and Javier Fernando Jiménez-Alonso[1[0000-0002-4592-0375]]

[1] Higher Technical School of Engineering, University of Seville, Seville 41092, Spain

Abstract. Surrogate modeling has established itself as an indispensable framework in modern engineering design optimization, offering a computationally efficient alternative to repeated high-fidelity structural evaluations. Grounded in the practical surrogate-based design methods this paper investigates the application of deep learning-assisted surrogate models to the multi-objective size optimization of a benchmark cantilever beam. The optimization problem aims to simultaneously minimize two competing design criteria: (i) the cross-sectional area; and (ii) the maximum bending stress. A mixed-integer design variable formulation is implemented to account for realistic geometric aspect ratio bounds and discrete structural dimension constraints. Furthermore, de-sign feasibility is governed by strict failure criteria, including allowable bending and shear stress thresholds, alongside lateral-torsional buckling load limits. A multi-layer perceptron deep neural network is constructed and trained on initial design sampling points to approximate both the objective functions and structural constraint boundaries. The multi-objective optimization problem is solved using both the exact structural model and the trained deep learning surrogate. Performance is rigorously benchmarked by comparing optimal solution sets, overall computational efficiency, and, crucially, the geometric topology and spread of the resulting Pareto fronts. The findings reveal that the deep learning surrogate accurately reproduces the true non-linear Pareto front and constraint trade-offs with negligible error, while drastically reducing computational overhead. This demonstrates the robust applicability of deep learning surrogates for mixed-integer structural size optimization in engineering design.

Keywords: Surrogate modeling, deep learning, multi-objective structural optimization, pareto front topology, size optimization, mixed-integer design variables.

1 Introduction

In current structural engineering, the design of load-bearing components requires a delicate balance between structural efficiency, mechanical performance, and manufacturing feasibility. Structural optimization has evolved from empirical trial-and-error procedures into a rigorous mathematical framework aimed at systematically identifying optimal configurations under demanding physical constraints. Among the various categories of structural optimization, structural size optimization plays a pivotal role during detailed engineering design. By optimizing cross-sectional dimensions while maintaining a fixed structural layout, size optimization directly controls structural stiffness, stress distribution, and overall material weight economy.

However, realistic engineering applications pose significant mathematical challenges. Structural sizing cannot be realistically formulated as a purely continuous optimization problem; commercial procurement and fabrication standards mandate discrete dimensions and standardized structural profiles. Consequently, the design domain must be modeled using mixed-integer design variables, which introduce non-smoothness, non-convexity, and severe combinatorial complexity.

Furthermore, structural safety is governed by strict multi-physics failure criteria, including allowable bending and shear stress thresholds, geometric aspect ratio limits, and global instability phenomena such as lateral-torsional buckling. When coupled with multi-objective formulations—such as the simultaneous minimization of cross-sectional area and maximum bending stress—determining the optimal trade-off set requires thousands of repeated high-fidelity finite element (FE) analysis evaluations. This iterative computational burden becomes prohibitively high, severely limiting the practical application of population-based metaheuristic optimizers in complex design spaces.

To overcome these computational bottlenecks, surrogate modeling (or metamodeling) has established itself as an indispensable framework in contemporary engineering design [1]. Surrogate models replace computationally expensive FE analysis calls with fast, data-driven mathematical approximations trained on a finite set of high-fidelity sampling points [2], [3]. In recent years, deep learning (DL) architectures, particularly multi-layer perceptrons (MLPs), have demonstrated exceptional capability in learning complex, non-linear mappings across high-dimensional spaces. By capturing the underlying structural response surfaces, deep neural networks can evaluate objective functions and constraint feasibility near-instantaneously, enabling rapid Pareto front exploration.

Despite the growing integration of DL in engineering design, a critical requirement remains: ensuring that surrogate-assisted frameworks accurately preserve the true non-linear trade-off boundaries without distorting the design space or introducing constraint violations under mixed-integer conditions. Specifically, the fidelity of DL surrogates in reproducing the exact geometric topology, continuity, and spread of non-linear Pareto fronts requires comprehensive benchmarking against exact structural solvers.

To address this need, this paper investigates the application of a DL-assisted surrogate modeling framework for the multi-objective structural size optimization of a

benchmark cantilever beam under mixed-integer constraints. The optimization aims to simultaneously minimize two competing design criteria: (i) the cross-sectional area, and (ii) the maximum bending stress. Design feasibility is strictly governed by allowable bending and shear stress limits, lateral-torsional buckling load thresholds, and discrete geometric constraints. Two MLP neural networks are constructed and trained on initial design samples to approximate both objective functions and constraint boundaries. The multi-objective optimization problem is solved using both the exact structural formulation and the trained DL surrogate. Performance is rigorously benchmarked by comparing optimal solution sets, overall computational efficiency, and the geometric topology and spread of the resulting Pareto fronts.

The remainder of this paper is organized as follows. Section 2 presents the theoretical foundations of structural optimization, mixed-integer formulations, and knee point selection using the angle method. Section 3 presents some basics about the generation of surrogate models based on DL architecture. Section 4 details the benchmark cantilever beam problem, including design variable definitions and structural failure criteria. Additionally, this section includes not only the deep neural network architecture, data sampling, and training methodology but also the discussion of the optimization results, comparing surrogate fidelity, Pareto front topology, and computational speedup. Finally, Section 5 provides concluding remarks to finish the manuscript.

2 Basics of Structural Optimization

Structural optimization seeks to systematically determine the design parameters of a structure, such that one or more performance criteria are optimized while strictly satisfying prescribed physical, mechanical, and geometric constraints. In its standard deterministic formulation, a structural multi-objective optimization problem is posed as a non-linear programming problem [4]:

$$\begin{aligned} &\text{Minimizing } \mathbf{f}(\theta) = [f_1(\theta) \quad \dots \quad f_k(\theta)] \\ &\text{Subjected to} \\ &g_j(\boldsymbol{\theta}) \le 0 \; j = 1,...,m \\ &h_l(\boldsymbol{\theta}) = 0 \; l = 1,...,p \\ &\theta_i^l \le \theta_i \le \theta_i^u \; i = 1,...,n \end{aligned} \tag{1}$$

where $\theta = [\theta_1 \quad \dots \quad \theta_n] \in D \subset \mathbb{R}^n$ denotes the vector of structural design variables; $\mathbf{f}(\boldsymbol{\theta})$ is the vector-valued objective function to be minimized; $g_j(\boldsymbol{\theta})$ and $h_l(\boldsymbol{\theta})$ represent the inequality and equality constraint functions respectively; and $[\theta_i^l, \theta_i^u]$ specify the lower and upper side bounds imposed on the design variables.

The structural response under static or dynamic load conditions is governed by the implicit state equilibrium equations derived from the FE method [5]:

$$\mathbf{K}(\theta)\,\mathbf{u}(\theta) = \mathbf{F} \tag{2}$$

where $\mathbf{K}(\boldsymbol{\theta})$ represents the global stiffness matrix (dependent on the design vector $\boldsymbol{\theta}$), $\mathbf{u}(\boldsymbol{\theta})$ is the structural displacement vector, and $\mathbf{F}$ is the vector of applied external loads. Inequality constraints $g_j(\boldsymbol{\theta})$ typically enforce limits on structural response

quantities, such as allowable element stresses (σ_{alw}), nodal displacements (u_{alw}), buckling load factors, or natural frequencies (for instance):

$$g_\sigma(\theta) = \frac{\sigma_s(\theta)}{\sigma_{alw}} - 1 \leq 0 \qquad (3)$$
$$g_u(\theta) = \frac{u_r(\theta)}{u_{alw}} - 1 \leq 0$$

where $\sigma_s(\boldsymbol{\theta})$ and $u_r(\boldsymbol{\theta})$ are respectively the value of the stress and displacement of the considered point of the structural system.

In accordance with classical structural optimization theory, structural optimization problems are categorized into three main structural domain levels according to the nature of the design variables [4]:

(i) **Size optimization:** The geometric domain and spatial topology of the structure remain fixed, while cross-sectional parameters—such as cross-sectional areas of truss members, second moments of inertia of frame elements, or thicknesses of shell plates—serve as design variables.

(ii) **Shape optimization:** The structural topology remains invariant, but the boundary contours, nodal coordinates, or surface profiles are continuously modified to alleviate local stress concentrations.

(iii) **Topology optimization:** The distribution of material within a predefined layout domain is optimized, determining the macro-structural connectivity and presence or absence of material.

In practical structural engineering (such as skeletal space frames, high-rise steel structures, and aerospace grids) structural size optimization constitutes the most critical phase during detailed design, as it directly dictates member selection from standard structural profiles/sheets and overall material weight economy. For this reason, this study is focused on this structural optimization domain.

While classical optimization literature often treats sizing variables as continuous quantities ($\theta_i \in R$), real-world manufacturing and structural design standards mandate discrete cross-sectional dimensions selected from commercial catalogs. Consequently, the design domain becomes mixed-discrete:

$$\boldsymbol{\theta_i} \in S_i = \{s_{i,1}, \ldots, s_{i,q}\},\ \forall i \in \{1,..,d\} \subseteq \{1,..,n\} \qquad (4)$$

where S_i denotes a finite non-equidistant set of standardized cross-sectional profiles or discrete values, transforming the formulation into a highly non-convex mixed-integer non-linear programming problem [6].

Furthermore, structural design intrinsically demands balancing conflicting requirements. Rather than optimizing a single criterion, structural sizing inherently involves trade-offs—such as minimizing total structural mass while simultaneously minimizing structural compliance or maximum deflection. Instead of a single optimal vector, the solution yields a set of Pareto-optimal solutions P^*:

$$P^* = \{\boldsymbol{\theta}^* \in \mathrm{D} \mid \nexists\ \boldsymbol{\theta} \in \mathrm{D} \text{ such that } \mathbf{f}(\boldsymbol{\theta}) \leq \mathbf{f}(\boldsymbol{\theta}^*) \text{ and } \mathbf{f}(\boldsymbol{\theta}) \neq \mathbf{f}(\boldsymbol{\theta}^*)\} \qquad (5)$$

After the determination of the Pareto front, a subsequent decision problem must be solved, the selection of a single representative solution from this set. The most desirable compromise is typically the knee point -a region on the Pareto front where a small improvement in one objective requires a disproportionately large sacrifice in

another. However, identifying the “knee point” objectively across a discrete, non-convex Pareto front presents a notable computational challenge, as trade-off surfaces can exhibit irregular density and complex curvature. To overcome this and automate the decision-making process, the angle method has been adopted herein [7]. This technique evaluates the geometric curvature of the Pareto front by calculating the bend angle formed between each non-dominated solution and its neighboring points in the normalized objective space. The point corresponding to the maximum interior angle represents the location of maximum trade-off efficiency [6].

Solving multi-objective mixed-integer structural sizing problems via traditional mathematical programming algorithms presents severe computational bottlenecks. Gradient-based techniques require continuous relaxation and are susceptible to trapping in sub-optimal local minima due to non-smoothness. Conversely, population-based metaheuristic algorithms can handle discrete variables but require tens to hundreds of thousands of high-fidelity FE analysis executions. For complex structural models, the cumulative computational cost of iterative FE analysis calls becomes prohibitively high. This critical limitation motivates the deployment of DL-based surrogate models. By training deep neural architectures to approximate the computationally expensive non-linear structural response mapping, $\theta \leftrightarrow \langle \mathbf{f}(\theta), \mathbf{g}(\theta) \rangle$, the structural optimizer can evaluate constraint violations and multi-objective trade-offs near-instantaneously, enabling efficient and scalable Pareto front exploration across mixed-integer search spaces.

3 DL-based Surrogate Modeling

Multi-objective structural optimization often requires repetitive, computationally intensive evaluations of objective functions and non-linear physical constraints. To alleviate this computational burden, a surrogate-assisted optimization framework based on DL can be implemented [1]. In this approach, the response of high-fidelity physical models can be approximated by MLP neural networks. The surrogate models act as fast, explicit function evaluators for the multi-objective genetic algorithm.

To ensure a uniform exploration of the decision space, D_d, defined by the structural dimensions (size optimization), a sampling method (such as a Latin Hypercube Sampling strategy) must be employed. A dataset of N_d design points can be generated:

$$D_d = \{(\boldsymbol{\theta}^i, \mathbf{f}^i, \mathbf{g}^i)\}_{i=1}^{N_d} \tag{6}$$

where is the $\boldsymbol{\theta}^i$ design vector i; $\mathbf{f}^i$ is the objective vector i; and $\mathbf{g}^i$ is the constraint vector i.

Several independent feedforward MLP neural networks can be designed for this purpose. Herein two independent ones were considered: (i) $\widehat{N}_{obj}(\boldsymbol{\theta})$ for the objectives and (ii) $\widehat{N}_{const}(\theta)$ for the non-linear constraints.

The following hyperparameters are normally considered for the design of these neural networks: (i) as network topology, an architecture with a minimum of four layers (1 input, 2 hidden and 1 output) are considered; (ii) activation functions (such as hyperbolic tangent sigmoid) must be selected for each layer; (iii) a training

algorithm (such as the Levenberg-Marquardt backpropagation algorithm) is chosen; and (iv) the sample dataset is randomly partitioned into 70% for training, 15% for validation (to prevent overfitting), and 15% for independent testing.
Mathematically, the forward mapping for a layer l is expressed as:

$$\mathbf{a}^{(l)} = \phi^{(l)}\left(\mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}\right) \tag{7}$$

where $\mathbf{W}^{(l)}$ denotes the weight matrix, $\mathbf{b}^{(l)}$ is the bias vector, and $\varphi^{(l)}$ is the layer activation function.

Finally, the accuracy of the trained surrogates can be evaluated on the unseen testing subset using the coefficient of determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^{N_{tes}} (y_i - \widehat{y}_i)^2}{\sum_{i=1}^{N_{test}} (y_i - \bar{y})^2} \tag{8}$$

where N_{tes} is the total number of test samples; y_i is the actual value for the i-th observation; $\widehat{y}_i$ is the predicted value from the surrogate model for the i-th observation; and $\bar{y}$ is the mean of the actual observed values.

4 Application example

The benchmark problem consists of the cross-sectional shape optimization of a cantilever beam subjected to a vertical point load F = 5000 N at its free end [2]. The length of the beam is about L = 1.5 m. The material specified is S275 structural steel with the following mechanical properties: (i) Young's Modulus, $2.1 \cdot 10^{11}$ N/m^2 and (ii) Poisson's ratio, 0.3. The design variables are the discrete cross-sectional dimensions: (i) width, $\theta_1 = b \in [0.01,0.05]$ m; and (ii) height $\theta_2 = h \in [0.02,0.25]$ m both with an increment step of 0.005 m (integer variables). The problem is formulated as a discrete multi-objective structural size optimization problem under mixed-integer constraints as follows:

$$\begin{gathered}
\text{Minimizing} \quad \mathbf{f}(\theta) = [f_1(\theta) = \theta_1 \bullet \theta_2 \quad f_2(\theta) = \tfrac{6FL}{(\theta_1 \bullet \theta_2^2)}] \\
\text{Subjected to} \\
g_1(\theta) = g_1(\theta) = \frac{FL^3}{3EI_y} - \delta_{\max} \le 0 \ (\delta_{\max} = 0.005 \text{ m}) \\
g_2(\theta) = \frac{6FL}{\theta_1 \bullet \theta_2^2} - \sigma_{\max} \le 0 \ \left(\sigma_{\max} = 2.75 \bullet 10^8 \text{ N}/m^2\right) \\
g_3(\theta) = \frac{1.5F}{\theta_1 \bullet \theta_2} - \tau_{\max} \le 0 \left(\tau_{\max} = \frac{2.75 \bullet 10^8}{\sqrt{3}} \text{ N}/m^2\right) \\
g_4(\theta) = \theta_2 - a_{\max} \cdot \theta_1 \le 0 \ (a_{\max} = 10) \\
g_5(\theta) = F - 0.5F_{cr} \le 0 \text{ where } F_{cr} = \frac{4.013}{L^2}\sqrt{EI_z GI_t} \\
I_y = \tfrac{1}{12}\theta_1\theta_2^3 \text{;} \ I_z = \tfrac{1}{12}\theta_2\theta_1^3 \text{;} I_t \approx \tfrac{1}{3}\theta_1^3\theta_2\left(1 - 0.63\tfrac{\theta_1}{\theta_2}\right)
\end{gathered} \tag{9}$$

To bypass repetitive evaluations of physical formulas and constraints during genetic algorithm iterations, a DL surrogate framework is implemented using two separate MLP neural networks. Decoupling the surrogates into two specialized networks yields significantly higher approximation accuracy compared to a single multi-task network: (i) network 1, $\widehat{N}_{obj}(\boldsymbol{\theta})$, which is trained specifically to map $[\theta_1,\theta_2] \rightarrow [f_1,f_2]$; and network 2, $\widehat{N}_{const}(\boldsymbol{\theta})$, which is trained specifically to map $[\theta_1,\theta_2] \rightarrow [g_1,g_2,g_3,g_4,g_5]$. A Design of Experiments (DoE) consisting of 400 points is generated using Latin Hypercube Sampling (LHS). Constraint values are standardized using Z-score normalization [8]. Both MLP networks feature 4 layers (1 input, 2 hidden layers with 12 and 8 neurons using hyperbolic tangent sigmoid activation functions, and 1 output layer). Training is carried out via Levenberg-Marquardt backpropagation with a 70/15/15% split for training, validation, and independent testing. As result of the training process, both surrogate models exhibit exceptional approximation performance on unseen test data, achieving coefficients of determination ($R^2 \geq 0.999$) across all objectives and non-linear physical constraints. The multi-objective optimization problem is solved under two evaluation models: (i) physics-based model (PBM); and (ii) surrogate-assisted model (SAM). The computation has been performed linking MATLAB's solver [9] with ANSYS FE formulations [10]. Integer constraints are strictly enforced on both design variables to map discrete index choices. Integer constraints are strictly enforced on both design variables to map discrete index choices. Thus, the primary objective is to evaluate whether the surrogate framework accurately preserves true solution fidelity without introducing artificial trade-offs. Fig. 1 illustrates the comparative Pareto fronts obtained by both methods.

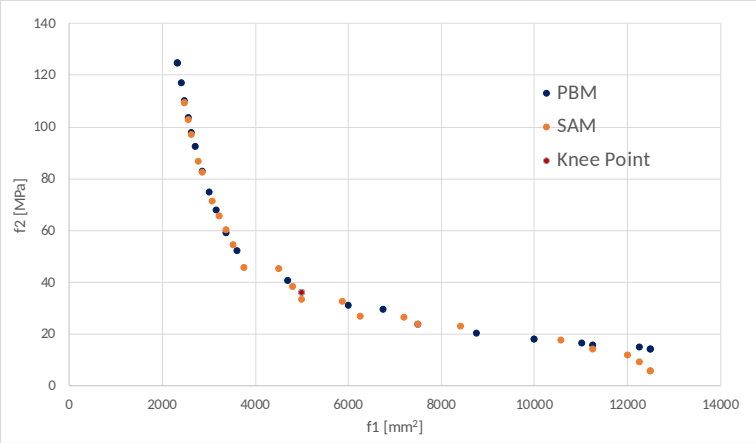


Fig. 1. Comparative analysis of the computed Pareto front according to the two approaches: (i) PBM and (ii) SAM.

Both PBM and SAM converge with the same set of non-dominated discrete design configurations. To quantify approximation accuracy across the non-dominated front, the mean absolute error (MAE) and root mean square error (RMSE), have been computed, obtaining the following values: (i) f_1 (Area), 153.53 mm^2 (MAE) and 230.11 mm^2 (RMSE) and (ii) f_2 (Stress) 3.39 MPa (MAE) and 5.19 MPA (RMSE). These values confirm that the surrogate model reproduces the physical Pareto front topology with exact fidelity. To solve the decision-making problem (the "knee point" selection), the geometric angle method is applied. The same structural design has been computed by both methods, $\theta_1^* = b^* = 20$ mm and $\theta_2^* = h^* = 250$ mm, achieving a cross-sectional area of 5000 mm^2 and a maximum bending stress of 36 MPa.

5 Conclusions

This paper presented a DL-assisted surrogate modeling framework applied to the multi-objective structural size optimization of a benchmark cantilever beam under mixed-integer constraints. Two methods have been considered to model the structural behavior: (i) PBM and (ii) SAM. From this comparative study, the following conclusions can be drawn: (i) the dual MLP network architecture demonstrated exceptional accuracy in capturing the complex non-linear mapping between mixed-integer design variables and structural responses ($R^2 \geq 0.999$); (ii) the surrogate model accurately reproduced the true non-linear Pareto front, preserving its exact geometric topology, continuity, and solution spread without introducing artificial discontinuities or false constraint violations; and (iii) the application of the angle method enabled an objective deterministic identification of the knee point on both Pareto fronts, identifying the identical optimal compromise structural design.

Acknowledgements

This study has been funded by the Spanish Ministry of Science, Innovation and Universities through the project "Updated physics-informed models for the fatigue assessment of cables and external post-tensioning systems in bridges" under the grant agreement PID2024-157616OA-C22.

References

1. Samadian, D., Muhit, I.B., Dawood, N.: Application of Data-Driven Surrogate Models in Structural Engineering: A Literature Review. Archives of Computational Methods in Engineering 32, 735–784 (2025).
2. Forrester, A.I.J., Sóbester, A., Keane, A.J.: Engineering Design via Surrogate Modelling: A Practical Guide 1st edn. John Wiley & Sons, Ltd, Chichester (U.K.) (2008).
3. Kim, N-H.: Surrogate Modelling and Optimization. Theories, Applications and Limitations 1st edn. John Wiley & Sons, Ltd, Florida (U.S.) (2026).

4. Arora, J.A.: Optimization of Structural and Mechanical Systems. 1st edn. World Scientific Publishing Co. Pte. Ltd., London (U.K.) (2007).
5. Ereiz, S., Duvnjak, I., Jiménez-Alonso, J.F.: Review of finite element model updating methods for structural applications. Structures 41, 684-723 (2022).
6. Nocedal, J., Wright, S.J.: Numerical Optimization. 2nd edn. Springer, New York (U.S.A.) (2006).
7. Jimenez-Alonso, J.F., Naranjo-Pérez, J., Pavic, A., Saez, A.: Maximum likelihood finite-element model updating of civil engineering structures using nature-inspired computational algorithms. Structural Engineering International 31(3), 326-338 (2021).
8. Prince, S.J.D.: Undestanding Deep Learning. 1str edn. Mit Press, Boston (U.S.A.) (2023).
9. ANSYS Homepage, http://www.ansys.com, last accessed 2026/10/15.
10. MATLAB Homepage, https://es.mathworks.com/products/matlab.html, last accessed 2026/10/15.

Neural Ordinary Differential Equations as Surrogate Models for Predictive Control of Non-Linear Dynamic Systems: Application to a Reaction Wheel Pendulum

Andy El Guarti Msaouri[1], Miguel Carrión Colchero[1] and Javier Fernando Jiménez-Alonso[1][0000-0002-4592-0375]

[1] Higher Technical School of Engineering, University of Seville, Seville 41092, Spain

Abstract. Model predictive control of non-linear dynamic systems relies heavily on ac-curate and computationally efficient internal prediction models. While physics-based models derived from first principles provide exact dynamic descriptions, their derivation can be complex or computationally demanding in real-time applications. Conversely, deep learning surrogates, specifically Neural Ordinary Differential Equations (Neural-ODE), offer a continuous-time learning framework naturally suited for dynamic system modeling. This paper presents a comparative performance evaluation of a Neural-ODE surrogate model embedded within a model predictive controller applied to a re-action wheel inverted pendulum - a benchmark non-linear mechanical system. First, the full non-linear dynamic equations of the pendulum are derived using the Euler-Lagrange formulation to establish a baseline physics-based model. Next, a Neural-ODE architecture is trained on state-space trajectory data to continuously capture the system's non-linear dynamics without explicit physical parameterization. Both models are subsequently integrated as predictive models into a control framework designed to execute swing-up and upright stabilization tasks. The performance of the Neural-ODE-based controller is benchmarked against the first-principles controller across multiple metrics, including trajectory tracking precision, closed-loop stability, robustness to unmodeled disturbances, and computational execution time per control iteration. The results show that the Neural-ODE surrogate accurately reproduces non-linear continuous dynamics within the predictive loop, achieving control performance comparable to the physical model while offering significant flexibility for data-driven modeling. This study validates Neural-ODEs as viable surrogate architectures for real-time predictive control of complex non-linear dynamic systems.

Keywords: Model predictive control, neural ordinary differential equations, reaction wheel inverted pendulum, dynamic surrogate modeling, non-linear control, dynamical systems.

1 Introduction

Model Predictive Control (MPC) complex non-linear dynamic systems relies heavily on accurate and computationally efficient internal prediction models [1, 2]. While physics-based models derived from first principles (such as Euler-Lagrange formulation) provide exact analytical descriptions, their derivation can become extremely complex in the presence of unmodeled dynamics, irregular geometries, or complex frictional phenomena. Furthermore, real-time evaluation of highly non-linear physical equations can be computationally demanding [3, 4].

To address these limitations, deep learning surrogate architectures have emerged as promising alternatives [5]. Specifically, Neural Ordinary Differential Equations (Neural-ODE) provide a continuous-time learning framework naturally suited for dynamic system modeling [6]. Rather than parameterizing discrete state transitions, a Neural-ODE parameterizes the continuous-time state derivative $\dot{x}(t) = f_\phi(x(t), u(t)$ using a neural network with parameters φ [7].

This paper presents a comparative performance evaluation of a Neural-ODE surrogate model embedded within a Non-linear MPC (NMPC) framework applied to a Reaction Wheel Inverted Pendulum (RWIP) - a classic benchmark non-linear mechanical system [3]. The primary contributions of this work are: (i) the derivation of the full non-linear dynamic model of the RWIP using the Euler-Lagrange framework as a baseline; (ii) training of a continuous Neural-ODE surrogate model using state-space trajectory data without explicit physical parameterization; (iii) the integration of both the physical and surrogate models into an NMPC framework to execute the underactuated swing-up maneuver and upright stabilization; and (iv) the comparative analysis of closed-loop performance using phase portrait trajectory tracking and Mean Squared Error (MSE) metrics.

The remainder of this paper is organized as follows. Section 2 presents the theoretical foundations of MPC and its implementation to control nonlinear dynamic systems. Section 3 presents some basics about the generation of reduced-order models to simulate the dynamic response of nonlinear systems based on deep learning neural networks, in particular Neural-ODEs. Section 4 details the benchmark problem considered to assess the performance of the proposal. For this purpose, a reaction wheel inverted pendulum has been modelled, and its dynamic behavior has been controlled using MPC and modelling the plant considering two different approaches: (i) a physics-based model and (ii) a surrogate model based on a Neural-ODE network. Finally, Section 5 provides concluding remarks to finish the manuscript.

2 Basics of MPC

MPC is an advanced multivariable control strategy governed by the receding horizon control principle. At each discrete sampling time *k*, the controller utilizes an internal dynamic model of the plant to predict future state trajectories across a finite

prediction horizon of N steps [1]. An optimal control problem (OCP) is solved online to determine a sequence of future control inputs $U_k = \{u_{k|k},...,u_{k+N-1|k}\}$ that minimizes a predefined performance index while satisfying state and input constraints.

The fundamental optimization problem balances tracking accuracy against control effort over the prediction horizon [1]:

$$\min_{U_k} \sum_{j=0}^{N-1} \left(\|x_{k+j|k} - x_{\text{ref}}\|_Q^2 + \|u_{k+j|k}\|_R^2 \right) + \|x_{k+N|k} - x_{\text{ref}}\|_P^2 \quad \text{subjected to} \quad u_{\min} \le u_{k+j|k} \le u_{\max} \quad j = 0, \ldots, N-1 \tag{1}$$

where $x_{k+j|k} \in \mathrm{R}^n$ represents the predicted state vector at horizon step $k + j$, computed using state information available at time step k, where n denotes the state space dimension; $x_{\text{ref}} \in \mathrm{R}^n$ is the target equilibrium reference state vector; $u_{k+j|k} \in \mathrm{R}^m$ is the predicted control input vector at horizon step $k + j$ evaluated at time k, where m denotes the control input dimension; $u_{\min}, u_{\max} \in \mathrm{R}^m$ represent the lower and upper physical actuation bounds, respectively; $Q \in \mathrm{R}^{n\times n} \succeq 0$ is the positive semi-definite state tracking error weighting matrix; $R \in \mathrm{R}^{m\times m} \succ 0$ is the positive definite control input effort weighting matrix; $P \in \mathrm{R}^{n\times n} \succeq 0$ is the positive semi-definite terminal state error weighting matrix; and $\|v\|_M^2 = v^T M v$ denotes the weighted squared Euclidean norm for any vector v and symmetric weighting matrix M.

Following the receding horizon paradigm, only the first calculated control action $u_{k|k}^* \in U_k^*$ is executed on the physical system. At the subsequent sampling interval $k + 1$, the true state x_{k+1} is re-evaluated, the prediction horizon shifts forward by one step ($k \to k + 1$), and the optimization is repeated. This online re-optimization establishes a closed-loop feedback mechanism that provides inherent robustness against bounded disturbances and minor modeling inaccuracies.

Standard linear MPC formulations rely on linear time-invariant state-space representations. However, local linearization fails when dynamic systems undergo large operational transitions, severe non-linear motions, or underactuated maneuvers —such as the full swing-up of an inverted pendulum. NMPC addresses these limitations by embedding continuous-time non-linear differential equations directly into the prediction model [1]:

$$\frac{dx(t)}{dt} = f(x(t), u(t)) \tag{2}$$

where $t \in \mathrm{R}_{\ge 0}$ denotes the continuous time variable; $x(t) \in \mathrm{R}^n$ is the continuous-time state vector; $u(t) \in \mathrm{R}^m$ is the continuous-time control input vector; f:$\mathrm{R}^n \times \mathrm{R}^m \to \mathrm{R}^n$ represents the continuous, non-linear vector field governing the physical dynamics.

Because continuous differential equations cannot be processed directly by discrete numerical optimization algorithms, the continuous field $f(x(t),u(t))$ must be discretized using numerical integration schemes over a constant sampling period $\Delta t \in \mathrm{R}^+$.

The state transition across a sampling step is expressed as [1]:

$$x_{k+j+1|k} = \blacksquare\left(x_{k+j|k}, u_{k+j|k}, \Delta t\right) = x_{k+j|k} + \int_{t_j}^{t_j+\Delta t} f(x(t), u(t))\, dt \tag{3}$$

where $t_j = j \cdot \Delta t$ is the discrete time instance at horizon step j, and $\mathbf{\Phi}$:$R^n \times R^m \times R^+ \rightarrow R^n$ denotes the discrete-time numerical state transition operator (such as 4th-order Runge-Kutta integration).

Embedding continuous non-linear dynamics $\mathbf{\Phi}$ into the prediction model transforms the objective function into a non-convex non-linear program, introducing three primary operational challenges [6]: (i) non-convexity, generating that he search space contains multiple local minima, rendering convergence sensitive to initial conditions and optimization parameters; (ii) computational complexity, numerical solvers must repeatedly compute multi-step numerical integration steps ($\mathbf{\Phi}$) and evaluate dynamic Jacobians $\nabla_x f(x,u)$ and $\nabla_u f(x,u)$ at every iteration; and (iii) real-time bottlenecks, for fast-responding mechanical systems requiring high sampling frequencies (50 − −100 Hz), real-time evaluation of complex closed-form analytical equations inside the optimization loop can exceed the sampling time limit Δt.

These constraints motivate the deployment of data-driven surrogate models specifically - continuous-time Neural ODE - which replace analytical physics formulations with a learned vector field $f_\varphi(x(t),u(t))$, maintaining smooth differentiability and enabling efficient real-time optimal control [8].

3 Basics of Neural ODE

Traditional discrete deep learning architectures model dynamic system transitions across fixed discrete-time intervals. In residual networks, hidden state updates follow a discrete transformation sequence $h_{k+1} = h_k + f(h_k,\theta)$. Taking the limit as the step size approaches zero converts discrete residual skip-connections into continuous-time differential equations. Neural-ODEs capitalize on this continuum limit by parameterizing the continuous derivative of the state vector using a neural network. Because physical dynamic systems are inherently continuous processes governed by physical laws, Neural-ODEs provide a natural mathematical framework for modeling non-linear physical dynamics. Unlike discrete surrogate models bound to a single fixed sampling time, Neural-ODEs represent the continuous vector field of the system, enabling evaluation over arbitrary continuous-time prediction intervals [9].

Let the dynamic system state vector be $x(t) \in R^n$ and the control input vector be $u(t) \in R^m$. The continuous-time rate of state change is modeled by a Multi-Layer Perceptron (MLP) f_φ [6]:

$$\frac{dx(t)}{dt} = f_\phi\left(x(t), u(t)\right) \tag{4}$$

where f_φ:$R^n \times R^m \rightarrow R^n$ is a continuously differentiable neural network mapping state-input pairs to time derivatives $\dot{x}(t)$ and $\varphi \in R^p$ represents the vector of trainable neural network parameters (weights and biases), with p denoting the total number of network parameters.

To compute the predicted state trajectory $x(t_0 + \tau)$ starting from an initial condition $x(t_0)$ over an integration step $\tau \in R^+$, the continuous vector field f_φ is integrated numerically [6,7]:

$$x(t_0+\tau) = x(t_0) + \int_{t_0}^{t_0+\tau} f_\phi\left(x(t), u(t)\right) dt = \text{ODESolve}\left(f_\phi, x(t_0), u, t_0, t_0+\tau\right) \tag{5}$$

where ODESolve represents a numerical differential equation integration operator (such as 4th-order Runge-Kutta).

The neural network parameters φ are optimized using state-space trajectory data collected from simulations or physical experimentation. The dataset D consists of $M \in \mathrm{N}^+$ sampled tuples containing state instances, applied control inputs, and state time-derivatives [9]:

$$\mathscr{D} = \left\{ \left(x^{(i)}, u^{(i)}, \dot{x}^{(i)} \right) \right\}_{i=1}^{M} \tag{6}$$

The parameters φ^* are obtained by minimizing the Mean Squared Error (MSE) loss function over the dataset:

$$\min_{\phi} \mathscr{L}(\phi) = \frac{1}{M} \sum_{i=1}^{M} \left\| \dot{x}^{(i)} - f_\phi\left(x^{(i)}, u^{(i)} \right) \right\|_2^2 \tag{7}$$

where $\|\|_2^2$ represents the standard Euclidean L_2 norm.

For trajectory-based training across time intervals, gradients $\frac{\partial \mathscr{L}}{\partial \phi}$ can be evaluated efficiently using the Adjoint Sensitivity Method [6]. By solving an augmented continuous-time adjoint ODE backward in time, gradients are computed with O(1) memory complexity relative to integration depth, avoiding the memory overhead of backpropagating through unrolled numerical solver steps.

In the context of NMPC, the trained Neural-ODE continuous field $f_\varphi(x,u)$ directly replaces the analytical dynamic equations derived from first principles. Thus, the discrete state transition operator $\mathbf{\Phi}_\varphi$ used inside the NMPC optimization loop becomes:

$$x_{k+j+1|k} = \blacksquare_\phi\left(x_{k+j|k}, u_{k+j|k}, \Delta t \right) = x_{k+j|k} + \int_{t_j}^{t_j+\Delta t} f_\phi\left(x(t), u_{k+j|k} \right) dt \tag{8}$$

where $t_j = j \cdot \Delta t$, and $\Delta t \in \mathrm{R}^+$ is the NMPC sampling period.

The use of this surrogate model has the following operational advantages in MPC: (i) Neural-ODEs learn complex, non-linear continuous vector fields directly from input-output trajectory data, bypassing the need to derive complex analytical parameters; (ii) MLP built with smooth activation functions yield continuous dynamic Jacobians $\nabla_x f_\varphi(x,u)$ and $\nabla_u f_\varphi(x,u)$, accelerating gradient convergence in Non-Linear Programming solvers; and (iii) because the surrogate model represents a continuous differential equation $\dot{x} = f_\phi(x,u)$ rather than a discrete map, the same trained network can be evaluated across varying integration steps or changing control sampling frequencies Δt without retraining [6, 8].

4 Application example

The RWIP is a classic underactuated non-linear mechanical system. It consists of a pendulum arm hinged at a fixed pivot and a reaction wheel attached to the free end of

the pendulum, driven by an electric motor [3]. The system is underactuated because control torque is applied directly to the reaction wheel, while the pendulum arm angle is controlled indirectly via coupling reaction torques. To derive the continuous-time equations of motion, we define two generalized coordinates: (i) $\theta(t) \in [-\pi,\pi]$, angular position of the pendulum arm relative to the unstable upright position ($\theta = 0$ rad top upright equilibrium, while $\theta = \pi$ rad denotes the downward hanging position); and (ii) $\varphi_w(t) \in \mathrm{R}$, relative angular position of the reaction wheel with respect to the pendulum body (being the relative angular velocity $\omega_w(t) = \dot{\phi}_w(t)$).

Defining the continuous state vector $x(t) = [x_1(t), x_2(t), x_3(t)]^T = [\theta(t), \dot{\theta}(t), \omega_w(t)]^T \in \mathbb{R}^3$ and the control input $u(t) = \tau(t) \in \mathrm{R}$, the system dynamics are expressed in non-linear state-space form $\dot{x}(t) = f(x(t), u(t))$ (derived from Euler-Lagrange formulation):

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} x_2(t) \\ \frac{m_p g l \sin(x_1(t)) - c_p x_2(t) - u(t)}{I_p} \\ \frac{u(t) - c_w x_3(t)}{I_w} - \frac{m_p g l \sin(x_1(t)) - c_p x_2(t) - u(t)}{I_p} \end{bmatrix} \tag{9}$$

where m_p is the pendulum mass (0.3 kg); l is the center of mass length (0.2 m); g is the gravity (9.81 m/s^2); I_p pendulum arm inertia (0.006 kgm^2); I_w reaction wheel inertia (0.001 kgm^2); c_p pendulum damping (0.01 sN/m); c_w wheel damping (0.005 sN/m); and $u(t)$ is the torque actuation ([-0.8,0.8] Nm).

This non-linear vector field $f(x,u)$ serves as the ground-truth physical dynamic model for both trajectory data generation and the physical baseline NMPC controller. The primary control objective is to perform a full swing-up maneuver combined with upright stabilization: (i) initial state, $x(0) = [\pi \quad 0 \quad 0]^T$ rad; and (ii) target state, $x_{ref} = [0 \quad 0 \quad 0]^T$ rad. This task requires the controller to build up kinetic energy via reaction wheel torque swings until the pendulum reaches the upright position ($\theta \approx 0$), where the controller must immediately transition to stabilization mode while respecting strict motor torque bounds $u(t) \in [-0.8,0.8]$ Nm. To construct the surrogate continuous vector field $f_\varphi(x,u)$, Neural-ODE architecture is configured. The dataset was generated by combining 5000 uniform random samples across the state-input bounds $\theta \in [-\pi,\pi]$ rad, $\dot{\theta} \in [-12, 12]$ rad/s , $\omega_w \in [-120,120]$ rad/s, and $u \in [-0.8,0.8]$ N · m, with continuous dynamic trajectory samples under sinusoidal excitation inputs. The network consists of a MLP with two hidden layers of 32 neurons utilizing hyperbolic tangent activation functions, trained using Bayesian Regularization. Open-loop continuous integration validation confirmed high model fidelity, achieving trajectory goodness-of-fit coefficients of determination $R^2 > 0.999$ across all state channels (as it is shown by Fig.1).

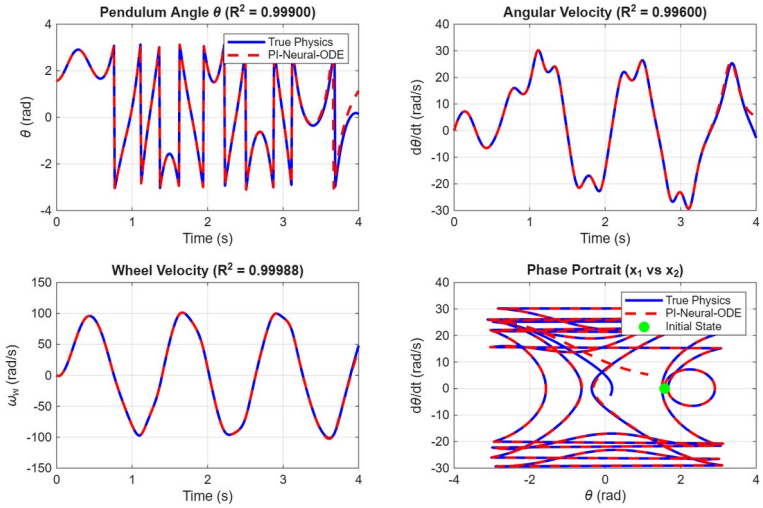


Fig. 1. Comparative study of the dynamic response of the physics-based model versus the surrogate model (Neural-ODE).

To eliminate computational overhead from high-level neural network toolboxes during online NMPC evaluation, raw weight matrices and bias vectors were extracted directly into a vectorized matrix integration solver [10]. Both the physis-based model and the Neural-ODE surrogate were integrated into an NMPC framework evaluated at a sampling step $\Delta t = 0.01$ s over a prediction horizon $N = 10$ steps (0.10 s) [1]. As it is illustrated in Fig. 2, both control formulations successfully perform the underactuated swing-up maneuver from the downward hanging position ($\theta = \pi$ rad) and maintain upright stabilization ($\theta = 0$ rad). The Neural-ODE NMPC matches the baseline trajectory dynamics while significantly accelerating numerical optimization iterations, confirming its efficacy for real-time optimal control loops.

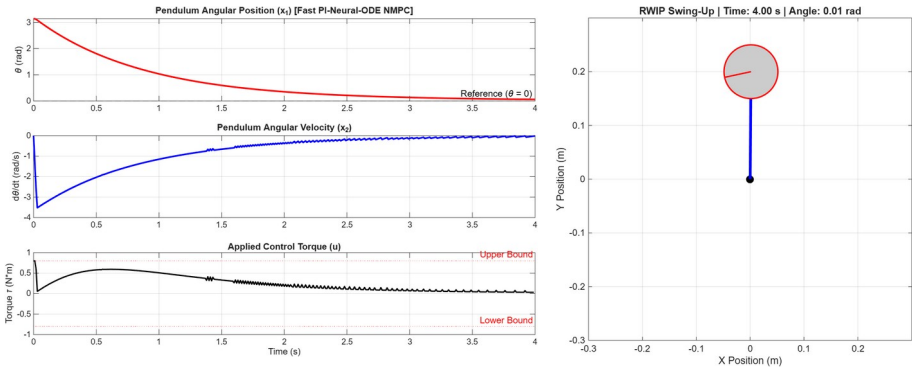


Fig. 2. Evolution of the states during the swing-up maneuver considering the surrogate model.

5 Conclusions

This paper presented a continuous-time Neural-ODE surrogate modeling framework embedded within NMPC architecture for non-linear underactuated dynamic systems. Applied to a Reaction Wheel Inverted Pendulum benchmark, the proposed methodology integrates exact kinematic constraints while learning complex continuous-time rotational accelerations directly from trajectory data without physical parameterization. By extracting trained weight matrices into a vectorized dynamic solver, the surrogate model enables fast gradient-based numerical optimization within the receding horizon loop. Closed-loop simulation results demonstrate that the surrogate model achieves trajectory tracking precision, closed-loop stability, and control effort performance comparable to the first-principles Euler-Lagrange baseline during both non-linear swing-up and upright stabilization.

Acknowledgements

This study has been funded by the Spanish Ministry of Science, Innovation and Universities through the project “Updated physics-informed models for the fatigue assessment of cables and external post-tensioning systems in bridges” under the grant agreement PID2024-157616OA-C22.

References

1. Khaled, N., Pattel, B.: Practical Design and Application of Model Predictive Control MPC for MATLAB® and Simulink® Users. 1st edn. Cambridge (U.S.A.) (2018).
2. Samadian, D., Muhit, I.B., Dawood, N.: Application of Data-Driven Surrogate Models in Structural Engineering: A Literature Review. Archives of Computational Methods in Engineering 32, 735–784 (2025).
3. Block, D.J., Åström,K.J., Spong, M.W.: The Reaction Wheel Pendulum. 1str edn. Springer Cham (Switzerland) (2007).
4. Ereiz, S., Duvnjak, I., Jiménez-Alonso, J.F.: Review of finite element model updating methods for structural applications. Structures 41, 684-723 (2022).
5. Forrester, A.I.J., Sóbester, A., Keane, A.J.: Engineering Design via Surrogate Modelling: A Practical Guide. 1st edn. John Wiley & Sons, Ltd, Chichester (U.K.) (2008).
6. Chen, R.T.Q., Rubanova, Y., Bettencourt, J., Duvenau, D.: Neural Ordinary Differential Equations. Advances in Neural Information Processing Systems 31, 6572-6583 (2018).
7. Worsham, J.M., Kalita, J.K. A guide to neural ordinary differential equations: Machine learning for data-driven digital engineering. Digital Engineering 6, 100060 (2025).
8. Aghili, J., Mula, O. An optimal control framework for adaptive neural ODEs. Advances in Computational Mathematics 50, 52 (2024)
9. Prince, S.J.D.: Undestanding Deep Learning. 1str edn. Mit Press, Boston (U.S.A.) (2023).
10. MATLAB Homepage, https://es.mathworks.com/products/matlab.html, last accessed 2026/10/15.

Proportional Control Tuning for Thermal Digital Twins of Asphalt Pavements: A Numerical Study

Javier Álvarez Pozo[1], Javier Fernando Jiménez-Alonso[1[0000-0002-4592-0375]], and Andrés Sáez[1[/0000-0001-5734-6238]]

[1] Higher Technical School of Engineering, University of Seville, Seville 41092, Spain

Abstract. Thermal digital twins offer a promising approach to monitoring pavement behavior by combining physics-based models, environmental measurements, and feedback mechanisms. However, their performance depends partly on the selection of controller parameters. In a proportional controller, the gain determines the corrective action associated with a temperature error, affecting tracking accuracy and control effort. This work proposes a numerical parametric study of proportional gain tuning for a reduced-order thermal digital twin for asphalt pavements. The framework couples a one-dimensional transient heat-transfer model, solved using the finite difference method, with a proportional controller that adjusts an effective convective boundary parameter. Temperature tracking error, cumulative control effort, transient response, and controller saturation are considered as complementary performance indicators. The study is designed to characterize the trade-off between temperature correction and corrective effort and to establish a reproducible basis for selecting controller gains. Quantitative results and the preferred gain range must be established from the proposed simulations before submission.

Keywords: Digital twins, asphalt pavements, thermal modelling, proportional control, finite difference method, controller tuning.

Introduction

Digital twins (DTs) integrate physical assets, numerical models, and measured data within a monitoring framework. In civil infrastructure, they can support condition assessment, predictive maintenance, and the interpretation of asset behaviour under changing environmental conditions. For pavement systems, temperature is a relevant variable because it influences the thermal state of asphalt layers and their mechanical response [1].

Physics-based thermal models can reproduce pavement temperature evolution using material properties, initial conditions, and environmental boundary conditions. Nevertheless, a numerical model operating independently of measurements cannot continuously account for discrepancies between its predictions and the behavior of the physical pavement. A feedback mechanism can address this limitation by adjusting a model parameter in response to temperature errors [2]. The effectiveness of this strategy depends partly on controller tuning. In a proportional controller, the gain determines the magnitude of the corrective action associated with a given temperature error. A low gain may produce slow correction, whereas a high gain may lead to excessive corrective actions, oscillatory behavior, or frequent activation of imposed control limits. Controller tuning is therefore an important step in developing a reliable thermal DT.

This work investigates the influence of proportional gain tuning on a reduced-order thermal DT for asphalt pavements. Unlike an assessment focused on detecting model–reality discrepancies, the present study focuses on the behavior of the feedback mechanism itself. Its contributions are: (i) a systematic numerical procedure for assessing proportional gain selection; (ii) complementary indicators for evaluating temperature tracking and controller effort; and (iii) an assessment framework for examining the trade-off between correction intensity and stable operation.

The remainder of the paper is organized as follows. Section 2 presents the thermal model and feedback controller. Section 3 describes the numerical experiment and performance indicators. Section 4 defines the application example and the results to be reported after simulation. Section 5 presents the conclusions.

1 Thermal modelling and control framework

1.1 One-dimensional thermal model

The virtual component of the DT is based on a one-dimensional (1-D)transient heat-transfer model representing temperature evolution through the pavement depth. Assuming heat conduction in the vertical direction, the governing equation is [3]:

$$\rho c_p \frac{\partial T(z,t)}{\partial t} = \frac{\partial}{\partial z}\left[\frac{k \partial T(z,t)}{\partial z}\right] \quad (1)$$

where T(z,t) is temperature at depth z and time t, ρ is material density, c_p is specific heat capacity, and k is thermal conductivity. The model is discretized in space using

the finite difference method (FDM). The resulting system calculates the transient temperature distribution from the initial thermal state, material properties, and boundary conditions. At the pavement surface, the thermal balance incorporates relevant environmental exchanges, including solar radiation, atmospheric longwave radiation, and convective heat transfer. The 1-D formulation reduces computational requirements and makes repeated simulations feasible. This is useful in feedback applications where the model must be updated as new environmental or temperature measurements become available.

1.2 Proportional feedback controller

The control mechanism compares the temperature predicted by the virtual model with the corresponding reference temperature. The tracking error is defined as [4]:

$$e(t) = T_{ref}(t) - T_{DT}(t) \tag{2}$$

where $T_{ref}(t)$ denotes the reference temperature and $T_{DT}(t)$ is the temperature calculated by the thermal model at the monitored depth. The proportional controller generates a corrective action according to [4]:

$$u(t) = K_p\, e(t) \tag{3}$$

where K_p is the proportional gain and $u(t)$ is the control signal. In the framework considered, the control signal modifies an effective convective boundary parameter $h_c(t)$.

A generic bounded implementation can be written as:

$$h_c(t) = sat\,[u(t)] \tag{4}$$

where sat[·] represents the imposed saturation limits. The precise implementation must remain consistent with the numerical model used in the simulations. The adjusted parameter is a lumped correction variable; it should not be interpreted as a direct measurement of the physical convective heat-transfer coefficient.

The gain K_p controls the sensitivity of the correction to the tracking error. Its effect depends on the thermal model dynamics, the time step, boundary conditions, and control limits. Therefore, its performance must be evaluated through the closed-loop response rather than inferred from the gain value alone.

At each time step, environmental conditions and the reference temperature are read; the thermal model calculates the pavement temperature; the tracking error is evaluated; and the controller output updates the effective boundary parameter for the next calculation. Repeating this cycle produces temperature and control-signal histories that can be analyzed together.

2 Numerical experimental design

2.1 Parametric study

The same thermal simulation will be executed for a set of proportional gain values. Initial conditions, thermal properties, environmental forcing, temporal discretization, reference temperature, and controller limits will remain constant so that the effect of gain selection can be isolated [5]. The gain range and increment should be selected from preliminary runs to cover weak correction and more aggressive responses, while avoiding a range that is too narrow to reveal meaningful differences. For each gain, the simulation will store the reference and model temperature histories, the tracking error, and the controller output. The reference temperature history must be defined independently of the controller being evaluated, allowing all configurations to be assessed against the same trajectory.

2.2 Performance indicators

Mean absolute error (MAE) measures the average magnitude of the tracking error:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|e_i| \tag{5}$$

Root mean square error RMSE) gives greater weight to larger deviations:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}e_i{}^2} \tag{6}$$

The integral of squared control signal quantifies cumulative corrective effort:

$$J_u = \sum_{i=1}^{N}u_i{}^2\Delta t \tag{7}$$

where Δt is the simulation time step. This is a numerical measure of control effort and should not be interpreted as physical energy consumption unless a separate physical relationship has been established. The saturation ratio is defined as:

$$R_{sat} = \frac{N_{sat}}{N} \tag{8}$$

where N_{sat} is the number of time steps in which the controller reaches an imposed limit. In addition to these indicators, time histories of the temperature error and control signal will be inspected for oscillations, delayed correction, and transient behavior.

The gain configurations will be compared using individual indicators and their combined interpretation. A reduction in temperature error will not automatically be considered an improvement if it is accompanied by a disproportionate increase in control effort or persistent saturation. The aim is to identify a region of acceptable performance rather than assume that the highest gain is necessarily the best choice.

3 Application example and results

3.1 Synthetic case study and simulation setup

A synthetic pavement case study was considered to investigate the influence of proportional gain tuning on the temperature tracking performance of the thermal DT. The experiment was designed to isolate the effect of the controller gain under a common set of thermal properties, initial conditions, and environmental forcing. Accordingly, all simulations were performed using the same reference temperature series and the same model configuration, while only the proportional gain K_p was varied.

The thermal model represents a homogeneous pavement domain in which heat transfer is assumed to occur predominantly in the vertical direction. The transient heat-conduction equation is solved using the finite difference method (FDM). The computational domain extends to a depth of 10 m, with a spatial discretization of Δz=0.01m. The temporal integration uses a fixed time step of Δt =60 s.

The physical parameters used in the numerical model are summarized in Table 1.

Table 1. Physical parameters

Properties	Value
ρ Density (kg/m^3)	*2200*
c Specific heat capacity J/kg/K	*4000*
σ Stefan- Boltzmann constant ($\frac{W}{m\,K^4}$)	*$5.670 \cdot 10^{-8}$*
ε_s surface emittance (non dimensional)	*1.0*

The environmental forcing was obtained from a meteorological station dataset containing shortwave radiation, downwelling longwave radiation, air temperature, and wind speed. The reference temperature series was generated synthetically using the thermal-model framework and the meteorological forcing. Therefore, the reference temperatures used in this experiment should not be interpreted as independent measurements from a physical pavement. This synthetic configuration enables a controlled comparison of controller gains under identical environmental conditions [5].

The simulation comprises 164,338 evaluated observations at a temporal resolution of 60 s, corresponding to approximately 114.1 days. The period of simulation was set from March 2020 to July 2020. For each proportional gain, the controller output was constrained between 0 and 100. The gain values were selected to span several orders of magnitude, allowing the response to be examined from weak correction to a regime dominated by control saturation.

4.2. Influence of proportional gain on temperature tracking

Table 2 summarizes the performance obtained for each tested proportional gain. The temperature tracking error decreases substantially as K_p increases from 0.01 to 100. For the lowest gain, the MAE and RMSE were 12.314 °C and 12.520 °C, respectively. Increasing the gain to 1 reduced the MAE to 3.597 °C, while (K_p =10) yielded an MAE of 0.592 °C.

Table 1. Performance obtained for each tested proportional gain

K_p	MAE	RMSE	J_{hc}	Saturation (%)
0.01	12.31389	12.51998	154558.9	0.004259514
0.03	11.52712	11.70574	1215981	0.004259514
0.1	9.569851	9.691603	9261432	0.004259514
0.3	6.743822	6.808817	41140817	0.004259514
1	3.597032	3.628312	$1.3 \cdot 10^8$	0.004259514
3	1.64575	1.66729	$2.47 \cdot 10^8$	0.004259514
5	1.084456	1.103167	$3 \cdot 10^8$	0.004259514
10	0.591817	0.606564	$3.63 \cdot 10^8$	0.004259514
30	0.212764	0.221778	$4.36 \cdot 10^8$	0.01521255
50	0.130163	0.137376	$4.65 \cdot 10^8$	0.35475666
100	0.067263	0.073393	$5.21 \cdot 10^8$	2.962796188
1000	0.155496	0.239992	$2.93 \cdot 10^9$	77.83835753

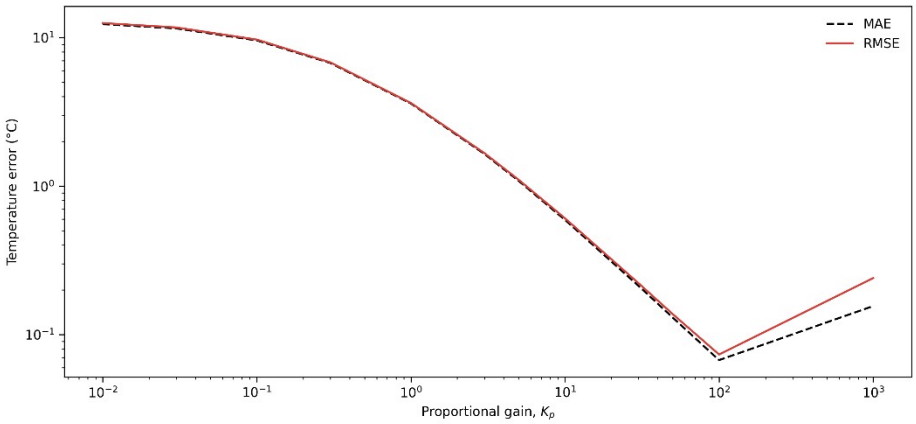


Fig. 1. MAE and RMSE obtained for each proportional gain performed

The best tracking performance among the tested configurations was obtained at (K_p =100), with an MAE of 0.067 °C and an RMSE of 0.073 °C. Compared with K_p =0.01), this represents an approximately 99.45% reduction in MAE. The improvement was not maintained at (K_p =1000), for which the MAE increased to 0.155 °C and the RMSE to 0.240 °C.

These results indicate that increasing the proportional gain improves tracking over most of the investigated range, but that the relationship is not monotonically beneficial. The deterioration observed at the highest gain coincides with a substantial increase in saturation, as discussed in the following subsection.

4.3. Control effort and saturation

The tracking improvement was accompanied by an increase in the magnitude of the controlled coefficient. The mean value of h_c increased from 0.123 for (K_p=0.01) to 6.605 for (K_p =100). Over the same range, the integral of the squared controlled coefficient increased from ($1.55 \cdot 10^5$) to ($5.21 \cdot 10^8$), using the time step in seconds.

Saturation was negligible for low gains, with a reported saturation ratio of approximately 0.0043% for gains up to (K_p =10). It increased to 0.3548% for (K_p =50) and 2.9628% for (K_p =100). Thus, the gain yielding the lowest tracking error also produced some saturation, although it remained limited to a small proportion of evaluated time steps.

A markedly different response was obtained at (K_p =1000). The controller reached its upper limit of (100 W,m^{-2},K^{-1}) during 77.84% of the evaluated steps. This configuration also exhibited higher MAE and RMSE than the (K_p =100) case. The results suggest that excessive gain can drive the controller into a regime dominated by saturation, limiting its ability to adjust the effective boundary parameter in response to the temperature error.

The integral of the squared controlled coefficient is used here as a numerical measure of control magnitude, rather than as a measure of physical energy consumption. Moreover, because the controller directly computes ($h_c = K_p$) before saturation, the resulting coefficient should be interpreted as an effective model-correction variable rather than an independently measured physical convection coefficient.

4.4. Discussion

The synthetic experiment demonstrates the importance of proportional gain selection for temperature tracking in the investigated thermal DT. Low gains produced substantial tracking errors, whereas intermediate and high gains reduced the discrepancy between the model and the synthetic reference series. The best tracking accuracy among the tested cases was obtained for (K_p = 100). However, the comparison also reveals that the highest gain does not necessarily provide the best performance. At (K_p = 1000), the controller was saturated for most of the evaluated period, and both error indicators increased relative to the (K_p = 100) configuration. This result supports evaluating controller performance through a combination of tracking accuracy, control magnitude, and saturation, rather than relying exclusively on MAE or RMSE.

4 Conclusions

This study investigated the influence of proportional gain tuning on the temperature tracking performance of a reduced-order thermal DT for asphalt pavements. A parametric numerical study was conducted using a 1-D FD thermal model driven by meteorological data and a proportional controller adjusting an effective convective boundary parameter within prescribed limits. The results demonstrated that the proportional gain has a substantial influence on temperature tracking accuracy. As K_p increased from 0.01 to 100, the MAE decreased from 12.314 °C to 0.067 °C, corresponding to a reduction of approximately 99.45%. The RMSE

also decreased from 12.520 °C to 0.073 °C, providing the best tracking performance among all tested configurations. However, the findings also showed that increasing the proportional gain indefinitely does not necessarily improve system response. At K_p = 1000, tracking performance degraded, with MAE and RMSE increasing to 0.155 °C and 0.240 °C, respectively. This performance degradation coincided with severe saturation: the controller reached its imposed upper limit during 77.84% of the evaluated time steps at K_p = 1000, compared to a saturation ratio of only 2.96% at K_p = 100 and 0.004% at low gains ($K_p \leq 10$). These results indicate that excessive gain can drive the controller into a regime dominated by saturation, diminishing the effectiveness of further corrective action. The analysis highlights a clear trade-off between tracking accuracy and control effort. The cumulative control effort (measured by the integral of the squared controlled coefficient, J_{hc}) increased continuously with proportional gain, rising from $1.55 \cdot 10^5$ at K_p = 0.01 to $5.21 \cdot 10^8$ at K_p = 100, and reaching $1.20 \cdot 10^9$ at K_p = 1000. Consequently, controller selection must jointly consider temperature error, control effort, and saturation rather than relying on a single tracking metric. Under the evaluated conditions, K_p = 100 provided the best observed balance between accuracy and control constraints without extensive saturation.

Several limitations should be noted. The study relied on a synthetic reference temperature series generated from meteorological forcing and fixed material properties. Future work will evaluate gain robustness under diverse weather conditions, variable thermal parameters, and independent field measurements from operational instrumented pavements.

Acknowledgements

This study has been funded by the Spanish Ministry of Science, Innovation and Universities through the project “Updated physics-informed models for the fatigue assessment of cables and external post-tensioning systems in bridges” under the grant agreement PID2024-157616OA-C22.

References

1. Grieves, M., Vickers, J.: Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In: Kahlen, F.-J., Flumerfelt, S., Alves, A. (eds.) Transdisciplinary Perspectives on Complex Systems, pp. 85–113. Springer, Cham (2017).
2. Incropera, F.P., DeWitt, D.P., Bergman, T.L., Lavine, A.S.: Fundamentals of Heat and Mass Transfer. 7th edn. John Wiley & Sons, Hoboken (2011).
3. Carslaw, H.S., Jaeger, J.C.: Conduction of Heat in Solids. 2nd edn. Oxford University Press, Oxford (1959).
4. Ogata, K.: Modern Control Engineering. 5th edn. Prentice Hall, Upper Saddle River (2010).
5. ASHRAE: ASHRAE Handbook—Fundamentals. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta (2021).

TRACK 2

Structural Engineering Innovations

SPECIAL TRACK ORGANIZERS

Kada Abdelhak Hassiba Benbouali University of Chlef, Algeria
Mohamed Ghrici Hassiba Benbouali University of Chlef, Algeria



KEYNOTE SPEAKER

Prof. Kada Abdelhak

Hassiba Benbouali University of Chlef, Algeria

Endeavour to Implement Performance Fire Design in Algeria: Case of Light Weight Cold Formed Structures

Kada Abdelhak is a Professor in the Department of Civil Engineering at the University Hassiba Benbouali of Chlef, Algeria. He obtained his Ingéniorat degree (B.Eng) in Civil Engineering from the National Polytechnic School of Algiers, Algeria, in 1985, followed by an MSc by research from Heriot-Watt University, Scotland, UK, in 1988, and a doctorate from the University Hassiba Benbouali of Chlef, Algeria, in 2016. He began his academic career as a Lecturer at the Faculty of Civil Engineering and Architecture, UHBC of Chlef, from 1990 to 2000, before advancing to Senior Lecturer from 2000 to 2019, Associate Professor from 2019 to 2024, and Professor from 2024 to the present, while also serving as Director of the LISICPE Laboratory since 2025.

His research interests span steel and composite structural design, fire engineering, earthquake engineering, and sustainable construction materials, with particular focus on multi-hazard assessment for structural analysis, 3D steel, composite steel-concrete and steel-timber systems, the behavior of structures and materials in fire, and sustainable construction materials.

TRACK TOPICS

- Advanced Structural Analysis
- Artificial Intelligence in Structural Engineering
- Structural Health Monitoring
- Earthquake-Resistant Structures
- Fire-Resistant Structures
- Innovative Construction Materials
- Recycled and Sustainable Materials
- Fiber-Reinforced Polymer (FRP)
- Advanced Steel Design
- Building Information Modeling (BIM)
- Nonlinear Structural Analysis
- Structural Dynamics
- Structural Strengthening and Retrofitting
- Structural Durability

An Interactive Tool for Temperature-Time Curve Prediction: Towards a safety strategy for structures in fire

Romaissa Hachana[1[0009-0002-1384-3333]], Abdallah Benarous[1,2*[0009-0000-4433-3569]]

Happy Sinkala [3[0009-0006-8871-6813]], Selahaddin Orhan Akansu [3 [0000-0002-0085-7915]]

[1]Department of Mechanics, Faculty of Technology, University Blida 1, Algeria
[2]LISICPE Laboratory, University H.B of Chlef, Algeria
[3]Erciyes University, Faculty of Engineering, Department of Mechanical Engineering, Kayseri, Turkey.

*a.benarous@univ-blida.dz

Abstract. Fire resistance of structures is evaluated by applying thermal action to the elements as a temperature-time (T-t) curve. Accordingly, the thermos-mechanical response is strongly related to the fire dynamics and boundary properties. The T-t curve is, therefore, the main and controlling input for any structural fire safety design. However, the traditional curves typically used for this purpose assume a spatially uniform gas temperature and do not account for ventilation and fuel loading effects. This work presents an interactive MATLAB®-based software tool, featuring several graphical user inter-faces (GUIs), designed to thermally characterize both localized and traveling fires in naturally ventilated compartment, allowing fire safety engineers to generate scenarios and face consistent T-t curves alongside the classical regulatory ones. The software provides a rapid estimation of gas temperature, velocity and smoke dynamics, while supporting fire safety evaluation during both design and operational phases of a building.

Keywords: Compartment fire, Travelling fire, Temperature-time curve, MATLAB App Designer.

1 Introduction

The fire resistance of structural elements is evaluated by applying a thermal action to the members expressed as a temperature-time (T-t) curve. This thermal load is then propagated through heat transfer and mechanical analysis [1,2]. The T-t curve is, therefore, the main and controlling input for any structural fire design. Fire safety engineering is concerned with the protection of life, property, and the environment against fire hazards [3]. The “compartment fire” is one of the concepts that has most influenced the development of the built environment and is particularly difficult to characterize due to its confined geometry and ventilation-driven dynamics [4].

The traditional curves typically used for this purpose assume a spatially uniform gas temperature and do not capture the ventilation, fuel load, or geometry of the real compartment [5]. the assumption of spatial uniformity becomes increasingly questionable in the large, open-plan, naturally ventilated compartments of modern buildings (offices, warehouses, shopping centers) where accidental fires are now recognized to burn locally and to spread over the floor plate as fuel is consumed rather than simultaneously engulfing the whole compartment [6, 7].

In this work, a software tool developed in the MATLAB® App Designer environment around a user-friendly, interactive graphical interface is presented, aiming to combine scientific accuracy with practicality by coupling the plume correlation, the ceiling-jet correlation and a Traveling Fire Methodology (TFM) alongside the classical fire curves for naturally ventilated compartments. The tool guides the user through the compartment geometry definition and fire scenario selection and automatically generates regulatory and parametric T-t curves. The objective is to offer a faster and representative thermal action input upstream of structural fire resistance assessment, which is usable during the design and operational phases of a building, as well as for teaching and research. The other section of this paper is organized as follows: the models and software architecture, the case study for demonstration, the resulting curves and conclusions.

2 Software Description and Methodology

The tool is built upon several families of models in a single MATLAB® GUI, selected to provide the thermal boundary conditions needed for structural fire analysis.

2.1 Plume/Ceiling-Jet Correlations

The correlations developed from laboratory tests on models or numerical simulations are generally specific to stationary and one-dimensional situations. They mainly characterize the dynamics and thermal properties of the smoke plume (including the flame) or the stratified flow at ceiling level (ceiling jet).
Since the plume is assumed to be axisymmetric, the correlations are therefore not suitable for situations where the plume is strongly influenced by ventilation. Furthermore, these correlations do not depend on the type of fuel, but only on the heat flow released by the fire source, sometimes taking into account only the convective part of this flow. It is important to note that this is a limitation of the correlations, because radiation losses vary by fuel, and therefore flame temperatures also vary [8].

In the Heskestad model, the heat release rate of the fire is replaced by its convective value to account for the energy lost through radiation, which does not contribute to the increase in temperatures and velocities within the plume. Additionally, the point source assumption is modified by introducing a virtual origin, inducing a rise in temperature and a flow of smoke, expressed as:

$$\Delta T = T_0 - T_\infty = 9.1 \left(\frac{T_\infty}{g C_p^2 \rho_\infty^2} \right)^{\frac{1}{3}} \dot{Q}_c^{\frac{2}{3}} (z - z_0)^{\frac{-5}{3}} \quad (1)$$

where T_0 is the centerline temperature, $\dot{Q}_c$ the convective heat release rate, z the elevation above the fire source, and z_0 the elevation of the virtual origin above the source.

For elements that are below the ceiling and outside the plume, the tool also uses the Alpert ceiling-jet correlation. This correlation provides the maximum gas temperature and velocity below an unconfined ceiling as a function of the radial distance from the plume axis and the height of the ceiling above the fire source [9].

2.2 Temperature time curves

The explicit formulations describing a 'zero-parameter' fire and commonly used in thermo-mechanical analysis, do not account for fire development phases and are written for ISO 834 and the hydrocarbon curve as [10]:

$$T = T_0 + 345log(8t + 1) \tag{2}$$

$$T = T_0 + 1080\,[1 - 0.325\exp\exp\,(-0.167t)\, - 0.675exp(-2.5t)] \tag{3}$$

where T is the temperature in °C and t is the time in minutes. For the so-called 'parametric fire', several expressions exist depending on whether the decay phase is considered; the Eurocode Parametric Fire Curve (EPFC) is:

$$T_g(t) = 20 + 1325(1 - 0.324e^{-0.2t^*} - 0.204e^{-1.7t^*} - 0.472e^{-19t^*}) \tag{4}$$

where $t^* = t.\Gamma(O,b)$is the scaling time, related to the opening factor O and the thermal inertia b of the bounding walls [11]. Because O and b depend on the specific compartment (openings, wall material), the EPFC is, unlike ISO 834, ventilation- and construction-consistent – a property the tool exploits by computing O and b directly from the user-entered compartment and wall data. As implemented on the Fire Curves tab, the tool limits these outputs to the heating phase of the fire (no burn-out or decay branch), which is the phase most relevant to peak thermal demand on structural members. The heating phase ends at the time of maximum gas temperature, t_{max} = max(0.2×10^{-3} $q_{t,d}/O$; t_{lim}), after which the cooling branch would begin.

2.3 Travelling Fire Methodology (TFM)

For large open naturally ventilated compartments where a single uniform curve is not representative, the tool employs a Near-field/Far-field Traveling Fire Methodology (TFM) [6,7]. This method is based on a fire of limited size (relative to the total floor area) burning over a portion of the floor plate and propagating at a defined rate so that any particular location on the floor is subjected to a short duration near-field temperature as the fire front passes overhead or in close proximity, followed by the longer duration, lower far-field temperature once the fire has moved on. The tool computes (from the heat release rate per unit area, the fuel load density, a reference flame flapping angle, and the assumed fire size) the local burning time, the total fire duration, the fire spread rate, the flapping (flame) length, and the resulting near-field and peak fire temperatures, then reconstructs the gas temperature-time history both as a function of time at fixed locations and as a function of location at fixed times, the two complementary views used for design. These generated gas temperature-time curves

subsequently serve as the boundary conditions for the heat-transfer analysis of structural members described.

2.4 Software Architecture

The tool is built in MATLAB® App Designer and presents its inputs and outputs in successive tabs: compartment geometry, pool fire, fuel loading, heat release rate (HRR), flame characteristics, plume characteristics, ceiling jet, fire curves and traveling fire. Each tab automates the calculations for the corresponding quantity and returns the relevant intermediate quantities (compartment surface, volume, opening factor, equivalent height, thermal inertia, time-scaling factor, etc.) together with the predicted temperature curves so that the most representative T-t curve for a given compartment and fire scenario can be selected and exported before passing to a structural fire model. The data-input/results-output structure and tabbed layout are designed to be usable by academic researchers and engineering practitioners (civil, architectural and fire-safety offices) without the need of MATLAB® programming skills. The modular structure allows the tool to be used as a teaching, research and field-analysis platform.

3 Case study

The tool is demonstrated in a case study of a carpentry workshop (32 m × 21 m floor area, 5 m ceiling height) according to the work of Santarpia et al. [12]. The compartment is ventilated by natural air flow through 5 windows (1.2 m x 2 m each) and 2 gates (3.2 m x 3.5 m). The solid boundaries are constructed of standard concrete with the physical properties as specified in Table 1. Figure 1 shows the Compartment Geometry tab of the tool. Here, the dimensions, openings and corresponding geometric and ventilation quantities are entered and calculated.

Table 1. Wall material physical properties used in calculations.

Material property	Value
Density ($kg.m^{-3}$)	2300
Specific heat ($J.kg^{-1}.K^{-1}$)	913
Thermal conductivity ($W.m^{-1}.K^{-1}$)	1.4

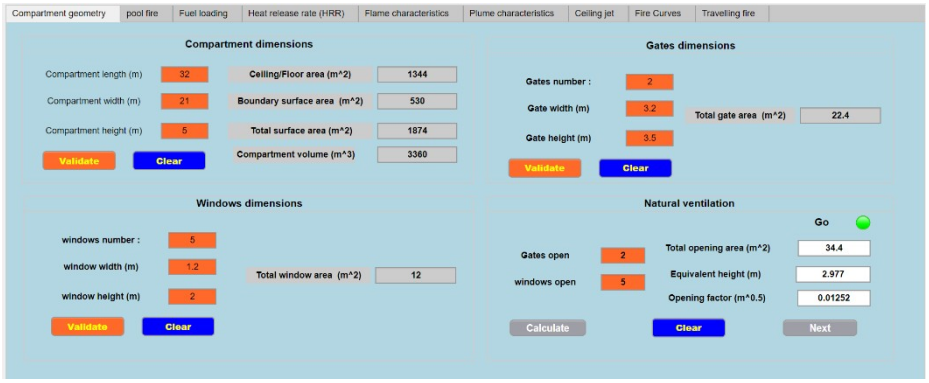


Fig. 1. Main graphical user interface: Compartment geometry tab

4 Results and Discussion

4.1 Fire Curves

Figure 2 shows the Fire Curves tab, in which the wall boundary properties from Table 1 and standard ambient conditions are used to compute the ISO 834, hydrocarbon and EPFC curves over a 90-minute window (for the EPFC, the heating phase only). The three curves exhibit distinct growth patterns and a clear ordering over the 90-minute window: the hydrocarbon curve rises the fastest, plateauing near 1100°C; ISO 834 climbs more gradually, without plateauing; and the EPF curve shows a slower heating rate than the ISO 834 and hydrocarbon curves, indicating the influence of natural ventilation on fire advancement.

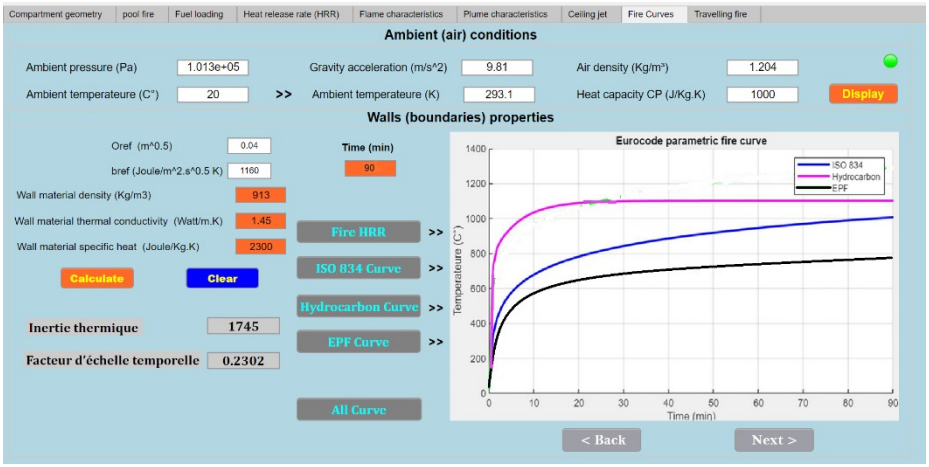


Fig. 2. Fire Curves tab: the resulting ISO 834, hydrocarbon and EPFC curves (EPFC: heating phase only).

4.2 **Traveling Fire Results**

Figures 3 and 4 show the two complementary outputs of the Travelling Fire tab for the case study compartment, computed with a heat release rate per unit area of 500 kW/m², a fuel load density of 570 MJ/m², and a reference flame flapping angle of 6.5°. Figure 3 plots the predicted temperature as a function of location along the fire path at three fixed times (30, 60, and 120 min). Figure 4 plots the same simulation as a function of time at three fixed locations (0, 13, and 20 m from the ignition point). Both views show a near-field plateau of about 1145°C, constant in amplitude since it is governed by the local plume/flame temperature rather than by position but mobile in space. In Fig. 3 the plateau is centered near x =2.5–3 m at t = 30 min, x =6.5 m at t = 60 min, and x = 14 m at t = 120 min, giving an average fire spread rate of about 2×10^{-3} m/s, consistent with the time lag between the near-field peaks observed at the fixed locations of Fig. 4 (immediately at t = 0 for the 0 m location, around t = 110 min for 13 m, and around t =170 min for 20 m). Each fixed location is exposed to the near-field plateau for only about 20 minutes as the fire front passes, before decaying towards a far-field level that itself rises gradually over the course of the fire (from about 200°C at early times to about 270–280°C in the tail of the simulation), reflecting the progressive accumulation of hot gases in the compartment as the fire advances. All three locations return to ambient temperature simultaneously at around 210 min, marking the exhaustion of the compartment's total fuel load rather than a local burnout. Table 2 summarizes the resulting TFM parameters.

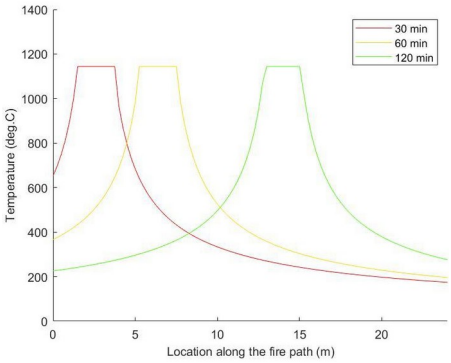


Fig. 3. Predicted temperature as a function of location along the fire path, at t = 30, 60 and 120 min.

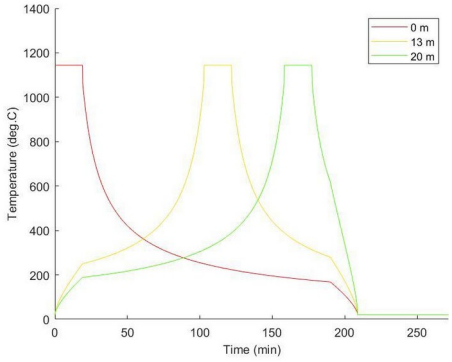


Fig. 4. Predicted temperature as a function of time, at fixed locations 0, 13 and 20 m from the ignition point.

Table 2. Travelling fire (TFM) results for the case-study compartment.

Parameter	Value
Heat release rate per unit area (kW/m²)	500
Fuel load density (MJ/m²)	570
Reference flame flapping angle (°)	6.5
Near-field peak temperature (°C)	≈1145

Near-field exposure duration at a fixed location (min)	≈20
Average fire spread rate (m/s)	$\approx 2\times10^{-3}$
Total fire duration (min)	≈210

5 Conclusion and Future Work

An interactive MATLAB® App Designer tool has been presented to make thermal-action selection, the first step in any structural fire design, faster and more representative. The key findings are:

- Integrated tool: plume/ceiling-jet correlations, regulatory fire curves and a traveling fire curve are coupled within a single, engineer-oriented interface, so that thermal-action selection becomes an immediate, comparative step rather than a disjointed, standard-by-standard calculation.
- Regulatory and parametric curves: in the carpentry-workshop case study (32 m × 21 m floor area, 5 m ceiling height), the tool reproduces the expected ordering between the hydrocarbon, ISO 834 and Eurocode parametric (EPFC) curves.
- Travelling fire: the TFM module captures the expected near-field/far-field temperature pattern as the fire moves across the compartment.
- Thermal exposure: the near-field peak is approximately 1145°C, sustained for approximately 20 minutes at any given location, for a total fire duration of approximately 210 minutes.
- Added value: these results differ significantly from, and are more spatially informative than, the uniform regulatory curves.

Ongoing work focuses on coupling the tool with CFD (FDS®) computations to benchmark the correlative and TFM predictions against field-resolved simulations, with the broader aim of feeding automated design loops for structural fire engineering practice.

References

1. Buchanan, A.H., Abu, A.K.: Structural Design for Fire Safety, 2nd edn. John Wiley & Sons, Chichester (2017).
2. Franssen, J.M., Real, P.V.: Fire Design of Steel Structures, 2nd edn. ECCS – European Convention for Constructional Steelwork, Berlin (2015).
3. Drysdale, D.: An Introduction to Fire Dynamics, 3rd edn. John Wiley & Sons, Chichester (2011).
4. Torero, J.L., Majdalani, A.H., Abecassis-Empis, C., Cowlard, A.: Revisiting the compartment fire. Fire Safety Science 11, 28–45 (2014).

5. Oli, T., Ha, D., Jang, T., Park, C., Kim, G., Kim, S.: Temperature distribution curve analysis on concrete through LS-DYNA. Fire 7(1), 15 (2011).
6. Stern-Gottfried, J., Rein, G.: Travelling fires for structural design – Part I: Literature review. Fire Safety Journal 54, 74–85 (2012).
7. Rackauskaite, E., Hamel, C., Law, A., & Rein, G.: Improved formulation of travelling fires and application to concrete and steel structures. In Structures, Vol. 3, 250-260 (2015).
8. Heskestad, G.: Fire plumes, flame height, and air entrainment. In: SFPE Handbook of Fire Protection Engineering, 3rd edn., pp. 2.1–2.17. NFPA-SFPE (2002).
9. Alpert, R.L.: Calculation of response time of ceiling-mounted fire detectors. Fire Technology 8, 181–195 (1972).
10. Hachana, R., Benarous, A., Roudane, M., Piloto, P.: An attempt to predict temperature-time histories for large scale compartment fires: A parametric analysis. Annals of the University of Craiova, Physics, 34, 128-136 (2024).
11. Lucherini, A., Jovanovic, B., Van Coile, R., Merci, B.: Background and limitations of the Eurocode parametric fire curves, including the fire decay phase. In: Proc. ASFE'21, pp. 330–335. University of Ljubljana (2021).
12. Santarpia, L., Bologna, S., Ciancio, V., Golasi, I., & Salata, F.: Fire temperature based on the time and resistance of buildings—predicting the adoption of fire safety measures. Fire, 2(2) (2019).

Extremely Randomized Trees Approach for Predicting the Axial Capacity of FRP-Reinforced Concrete Compressive Members

Sarra Sendjasni[1a], Klaus Holschemacher[2], Mohammed Berradia[1], Riad Benzaid[3], Asmaa Djilali Bouzina[1]

[1] Department of Civil Engineering, Laboratory of Structures, Geotechnics and Risks (LSGR), Hassiba Benbouali University of Chlef, B.P 78C, Ouled Fares, Chlef 02180, Algeria
[2] Structural Concrete Institute, HTWK Leipzig, Leipzig, Germany
[3] S.T.U. Department, L.G.G research laboratory, University of Jijel - Mohamed Seddik Benyahia, B.P 98, Cité Ouled Aissa, 18000 Jijel, Algeria

[a] s.sendjasni@univ-chlef.dz

Abstract. Reinforced concrete compressive members internally reinforced with fibre-reinforced polymer (FRP) bars are increasingly specified where long-term corrosion resistance is a priority. However, their axial capacity is still commonly assessed using code provisions and empirical regressions developed from limited experimental datasets and, in most cases, neglecting load eccentricity. This study develops a data-driven alternative based on Extremely Randomized Trees (Extra-Trees), an ensemble learning method that constructs highly randomized decision trees, to predict the ultimate axial capacity of FRP-reinforced concrete members under concentric and eccentric compression. A database of 308 tested specimens collected from the literature is employed, with a dimensionless eccentricity ratio derived directly from specimen geometry to explicitly represent the loading condition. Following hyperparameter optimization using cross-validated random search, the proposed model achieves coefficients of determination (R^2) of 0.99 for the training set, 0.92 for the independent test set, and 0.98 for the complete database, with a mean absolute percentage error below 9%. Compared with four representative code-based and empirical formulations, which yield R^2 values ranging from 0.34 to 0.88 on the same database, the proposed model reduces the root-mean-square error by more than half, reaching a reduction of up to 80% relative to the least accurate formulation. SHapley Additive exPlanations (SHAP) indicate that predictions are governed primarily by cross-sectional area, eccentricity ratio, and concrete strength, consistent with established structural mechanics. Overall, the results demonstrate that the validated Extra-Trees model provides accurate and interpretable predictions of axial capacity under both concentric and eccentric compression for engineering applications.

Keywords: *FRP-reinforced concrete · axial capacity · Extremely Randomized Trees · machine learning · SHAP · eccentric compression*

1 Introduction

Fibre-reinforced polymer (FRP) bars are used as an alternative to steel reinforcement where corrosion is a major concern. Typical applications include marine substructures, bridge decks, and industrial platforms. However, FRP bars have lower compressive strength than steel bars of equal diameter. Consequently, FRP-reinforced concrete (FRP-RC) members generally exhibit lower axial capacities. This reduction depends on geometric, material, and loading variables that current design provisions do not fully account for.

The axial capacity of FRP-RC members is estimated using design codes, such as ACI 440.1R-15, CSA S806-12, and AS 3600, or empirical formulas. These approaches mainly differ in their treatment of longitudinal FRP bars. Most do not explicitly

consider eccentric compression. A recent study evaluated fifteen formulas using 308 experimental specimens. Their R^2 values ranged from 0.34 to 0.71 [1]. An empirical model incorporating eccentricity achieved an R^2 of 0.88 [1]. These results highlight the need for more accurate prediction methods.

Machine learning has been applied to FRP members using neural networks [7,8], group method of data handling (GMDH) [9], and gradient-boosted trees with SHAP interpretation [6]. However, bagging-based methods remain less explored for predicting FRP-RC axial capacity. Unlike boosting, bagging builds multiple trees independently. This offers a potentially simpler approach to prediction.

This study applies Extremely Randomized Trees (Extra-Trees) to predict the axial capacity of FRP-RC members. The method randomly selects splitting features and thresholds during tree construction. It uses the 308-specimen database compiled by Sendjasni et al. [1]. The main contributions are: (i) introducing a dimensionless eccentricity ratio derived from specimen geometry; (ii) developing and evaluating an Extra-Trees model against existing design codes and empirical formulas; and (iii) applying SHAP analysis to interpret the model predictions.

2 Materials and Methods

2.1 Experimental Database

The specimen-level database analysed in this study originates from Sendjasni et al. [1], who assembled results from forty-two independent experimental campaigns published between 2010 and 2023 into a single, consistently formatted record of 308 FRP-RC compressive members. The database spans both circular (187 specimens) and square (121 specimens) cross-sections, three transverse-reinforcement configurations - spirals (172), ties (108) and hoops (28) and three longitudinal FRP materials, namely glass (239 specimens), carbon (53) and basalt (16) FRP. Of the 308 members, 151 were loaded concentrically and 157 eccentrically, and the recorded ultimate axial capacity ranges from 90 kN to 15,234 kN, a span of almost 170:1 that reflects the wide variety of specimen scales represented.

Table 1 groups the twelve continuous predictors retained in this study by geometric, unconfined-concrete and FRP-material category, together with the target variable. Four categorical descriptors complete the input set: cross-section shape, longitudinal FRP type, transverse FRP type and tie configuration.

Table 1. Range of the continuous variables retained as model inputs, grouped by category.

Category	Variable	Min	Max	Mean	SD
Geometric	H (mm)	500	3730	1409.7	583.9
Geometric	A_g (mm²)	12,272	372,100	63,158.9	43,746.3
Geometric	$\frac{e}{h}$	0.00	0.82	0.10	0.16
Concrete	f'_c (MPa)	20.8	90.0	42.3	13.8
FRP material	ρ_{frp} (%)	0.55	4.80	2.02	0.85
FRP material	A_{frp} (mm²)	199.7	3721.0	1232.9	723.1
FRP material	^{n}	3	16	6.8	2.3
FRP material	d_m (mm)	8.0	25.4	14.5	3.0
FRP material	E_{frp} (GPa)	39.0	151.0	65.9	31.9
FRP material	f_{frp} (MPa)	574	2000	1158.4	393.2
Transverse steel/FRP	d_s (mm)	4.0	12.7	8.8	1.9
Transverse steel/FRP	S_v (mm)	30	305	92.5	53.0
Target	P_u (kN)	90.0	15,234.4	1980.8	1830.6

2.2 Overview of Existing Models

Five existing equations for calculating the axial capacity of FRP-reinforced concrete columns are presented and evaluated for their accuracy and effectiveness using the developed database.

Table 2. Existing Axial capacity formulas

Code/Research	Developed formulas, P_u
ACI 440.1R-15 [10]	$\mathbf{P_u = 0.85\, f'_c\, (A_g - A_{frp})}$
CSA S806-12 [11]	$P_u = \alpha_1 f'_c\, (A_g - A_{frp})$ $\alpha_1 = \mathbf{0.85 - 0.0015 f'_c \geq 0.67}$
Tobbi et al. (2012) [12]	$P_u = \alpha_1\, f'_c\, (A_g - A_{frp}) + \alpha_{frp} f_{frp} A_{frp}$ $\alpha_1 = 0.85\ ;\ \alpha_{frp} = 0.35$
Hadhood et al. (2017) [13]	$P_u = \alpha_1\, f'_c\, (A_g - A_{frp}) + 0.0035 E_{frp} A_{frp}$ $\alpha_1 = 0.85 - 0.0015 f'_c \geq 0.67$
Unified empirical model [1]	$P_u = 0.85 f'_c (A_g - A_{frp}) + 0.0037 E_{frp} A_{frp} - 5{,}022 (e/h) \times 100 \times n$

2.3 Eccentricity-Ratio Feature Engineering

The source records report the applied eccentricity together with a binary concentric/eccentric label, but not a size-normalized eccentricity measure. Since the influence of eccentricity on axial capacity depends on its magnitude relative to the section depth rather than on *e* alone, a characteristic section dimension *h* was first derived from the gross cross-sectional area and the section shape (the diameter for circular sections, the side length for square sections), and the ratio $(\frac{e}{h})$ was computed for every specimen. Replacing the binary loading flag with this continuous ratio lets the tree ensemble select any split threshold along the eccentricity axis rather than only two branches, and it is the only engineered feature used in this study; all other predictors are taken directly from the source database.

2.4 Extremely Randomized Trees

Extremely Randomized Trees (Extra-Trees) build a forest of decision trees in which the randomization that distinguishes a Random Forest [6] is pushed one step further [7]. A Random Forest bootstraps the training set for each tree and then searches, at every node, for the locally optimal split threshold among a random subset of features; Extra-Trees instead draw the split threshold itself at random for each candidate feature and keep whichever random split maximizes the variance-reduction criterion, using the full training set for every tree rather than a bootstrap resample. This extra randomization decouples the individual trees more strongly, trading a small increase in bias for a larger reduction in ensemble variance, and it removes the search for an optimal cut-point as a source of overfitting - useful on a database of only a few hundred specimens, where an exhaustive threshold search can lock onto sampling noise. Every categorical predictor was one-hot encoded before training, since Extra-Trees does not natively partition categorical levels the way some boosting implementations do.

2.5 Model Training and Validation Protocol

The 308 specimens were split once into a training partition of 246 specimens (80%) and a testing partition of 62 specimens (20%), using a fixed random seed for reproducibility. Five hyperparameters number of trees, maximum tree depth, minimum samples to split or to form a leaf, and the feature fraction considered per split together with the choice of bootstrapping, were tuned with the scikit-learn implementation of Extra-Trees [8] via a randomized search of 40 candidate configurations, each scored by five-fold cross-validation on RMSE. The configuration with the lowest mean cross-validated error was retained and re-fitted on the full training partition (Table 2). To check that the single train/test split was not an unrepresentative draw, the tuned configuration was reevaluated over fifteen further five-fold cross-validation runs on independent partitioning of the full database; the mean cross-validated R^2 across these fifteen runs was 0.91 with a standard deviation of 0.03, consistent with the held-out test result reported in Sect. 3.

Table 2. Extra-Trees hyperparameters selected by cross-validated random search.

Hyperparameter	Selected value
Number of trees	200
Maximum tree depth	12
Minimum samples to split a node	3
Minimum samples per leaf	1
Feature fraction considered per split	0.8
Bootstrap resampling	Disabled

2.6 Performance Indicators

Model accuracy is reported using four standard regression indicators: the coefficient of determination (R^2), the root-mean-square error (RMSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE). To keep the comparison with the code and empirical formulas of Sect. 3.5 on a common footing regardless of the wide range of specimen capacities in the database, RMSE is also expressed as a percentage of the mean experimental capacity (normalized RMSE, NRMSE). Lower RMSE, MAE, MAPE and NRMSE, together with R^2 closer to unity, indicate closer agreement between predicted and experimental capacities.

3 Results and Discussion

3.1 Correlation Structure of the Database

Figure 1 shows the pairwise Pearson correlation between the twelve continuous variables and the experimental capacity P_u. Gross cross-sectional area A_g correlates most strongly with capacity (r = 0.9), as expected from its role in the concrete-contribution term, while the eccentricity ratio $\frac{e}{h}$ is the only variable to correlate negatively with capacity (r = −0.4), confirming its physical role as a capacity-reducing factor rather than a contributing one. Several FRP-material descriptors are themselves strongly intercorrelated notably A_{frp} with bar diameter and bar count ($r \geq 0.8$) a pattern typical of reinforcement design tables in which these quantities are not chosen independently. Such correlated predictors pose little difficulty for a tree ensemble, which selects one variable at a time rather than requiring the linear independence that a regression model would.

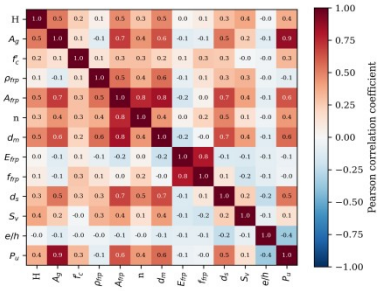


Fig. 1. Pairwise correlation between the continuous input variables and the experimental axial capacity.

3.2 Predictive Accuracy

Table 3 reports the accuracy of the tuned Extra-Trees model on the training partition, the testing partition and the full database. The model reaches R^2 = 0.992 on the training data, falling to 0.923 on the independent test set and averaging 0.978 over all 308 specimens; MAPE follows the same pattern, from 4.7% on training data to 22.4% on the test set. This gap between partitions is expected given that the test set contains only 62 specimens drawn from a target variable spanning almost three orders of magnitude, and it is corroborated by the repeated cross-validation result reported in Sect. 2.4 rather than being an artefact of one particular split.

Table 3. Accuracy of the tuned Extra-Trees model.

Indicator	Training	Testing	Full database
R^2	0.992	0.923	0.978
RMSE (kN)	166.1	501.9	269.7
MAE (kN)	80.1	292.1	122.8
MAPE (%)	4.74	22.36	8.29
NRMSE (%)	8.42	24.93	13.62
N	246	62	308

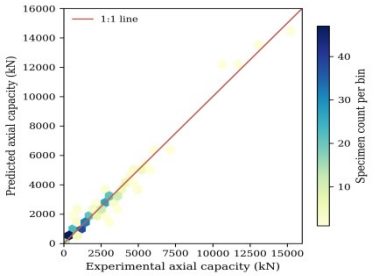


Fig. 2. Density of predicted versus experimental axial capacity, pooling the training and testing partitions; darker hexagons indicate more specimens.

The hexagonal binning in Fig. 2 makes clear that the bulk of the database, which is concentrated below about 4000 kN, sits tightly along the 1:1 line, while the handful of very large specimens above 10,000 kN too few to form a dense cluster are still tracked closely by the model. This indicates that the model's accuracy is not confined to the densely populated part of the capacity range.

3.3 Residual Analysis

Figure 3 plots the distribution of residuals (experimental minus predicted capacity) separately for the training and testing partitions. Both distributions are centred close to

zero and are approximately symmetric, indicating that the model neither systematically over- nor under-predicts capacity. A small number of testing-set residuals exceed 1000 kN in magnitude; inspection of these cases shows they correspond to the largest specimens in the database, where even a small relative error translates into a large absolute residual simply because the capacities themselves are an order of magnitude above the database mean.

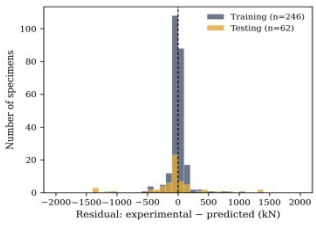


Fig. 3. Distribution of prediction residuals for the training and testing partitions.

3.4 Feature Importance and Physical Interpretation

To verify that the Extra-Trees model bases its predictions on physically defensible relationships rather than incidental patterns in the training data, SHAP (SHapley Additive exPlanations) values [9] were computed for every specimen and averaged, with the three dummy variables belonging to each one-hot-encoded categorical predictor summed into a single grouped score. Figure 4 shows the resulting ranking: gross cross-sectional area dominates, contributing more than 1.5 times the next most influential variable, the engineered eccentricity ratio, ahead of concrete strength and FRP reinforcement area. Section shape is the most influential categorical descriptor, consistent with its role in determining the characteristic dimension used to compute $\frac{(e}{h)}$. This ordering mirrors the structure of the concrete-contribution term common to every code and empirical formula reviewed in Sect. 1, and the prominence of $\frac{e}{h}$ a feature that was engineered rather than supplied directly indicates that recasting the loading condition as a continuous ratio was informative to the model, not merely a convenient input format.

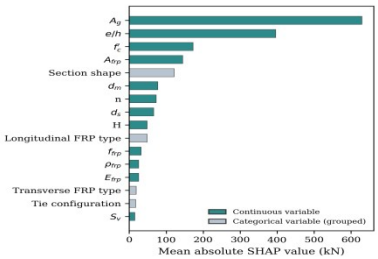


Fig. 4. Mean absolute SHAP value by input variable, with one-hot-encoded categorical predictors aggregated by category.

3.5 Benchmarking Against Design Equations

Table 4 compares the Extra-Trees model, evaluated on the full 308-specimen database, against four representative formulas assessed by Sendjasni et al. [1] on the same data: the ACI 440.1R-15[10] and CSA S806-12 [11]code equations, the empirical model of Tobbi et al. [12], the empirical model of Hadhood et al. [13], and the unified empirical expression that Sendjasni et al. [1] fitted with an explicit eccentricity term. The complete set of fifteen formulas reviewed in that source spans R^2 = 0.34-0.71 for the code and single-mechanism empirical models, rising to 0.88 once eccentricity is fitted explicitly; the four rows shown span this range without repeating every formula.

Table 4. Extra-Trees model versus representative code and empirical formulas (formula results from [1]).

Model	R^2	RMSE (kN)	MAE (kN)
ACI 440.1R-15 [10]	0.70	909.1	662.2
CSA S806-12 [11]	0.71	846.6	651.9
Tobbi et al. (2012) [12]	0.61	1245.8	861.7
Hadhood et al. (2017) [13]	0.34	1380.9	960.0
Unified empirical model [1]	0.88	635.8	405.1
Proposed Extra-Trees	0.98	269.7	122.8

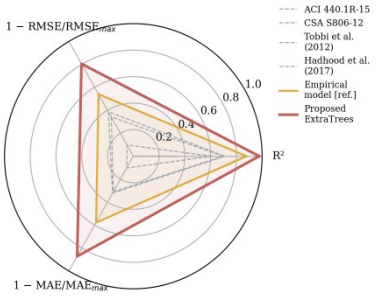


Fig. 5. Multi-indicator comparison of the proposed model against the code and empirical formulas of Table 4; the outer vertex on each axis represents the best observed value.

The three axes in Fig. 5 - R^2, and RMSE and MAE each expressed as one minus their ratio to the worst-performing formula, so that a larger enclosed area is always better show the proposed model enclosing every code and empirical formula on every axis simultaneously, including the unified empirical model that already incorporates an eccentricity correction. Relative to the four code/empirical rows in Table 4, the Extra-Trees model reduces RMSE by 58-80% and MAE by 68-87%, while raising R^2 from a maximum of 0.88 to 0.98. Because the same database and the same statistical indicators were used throughout, this comparison isolates the benefit of a flexible, nonparametric predictor from any difference in the underlying test population.

4 Conclusions

An Extremely Randomized Trees model was developed to estimate the axial capacity of FRP-reinforced concrete compressive members under concentric and eccentric compression, using a 308-specimen database and a dimensionless eccentricity ratio engineered from the raw specimen geometry. The findings support the following conclusions:

1. Cross-validated random search tuning produced a model with $R^2 = 0.992$ (training), 0.923 (testing) and 0.978 (full database), stable across fifteen repeated cross-validation runs on independent partitionings (mean $R^2 = 0.91$, SD = 0.03).
2. Benchmarked against four code and empirical formulas spanning $R^2 = 0.34$-0.88 on the same database, the model reduced RMSE by 58-80% and MAE by 68-87%, a substantial and consistent accuracy gain over the closed-form expressions currently used in design.
3. SHAP analysis showed predictions governed primarily by cross-sectional area, the engineered eccentricity ratio and concrete strength, in that order - consistent with the concrete-contribution term shared by every formula reviewed, and evidence that the engineered eccentricity feature carries genuine predictive information.
4. Residual and density analysis showed no systematic bias across the capacity range, with the largest absolute errors confined to the few very-high-capacity specimens where relative error, not model bias, dominates.

5. Requiring only a modest, easily interpreted hyperparameter set and no sequential residual-fitting, Extra-Trees offers a practical route to an accurate, auditable design aid for FRP-RC compressive members, complementary to the code and empirical formulas it was benchmarked against.
6. Compared with the code and empirical formulas, the proposed model predicts axial capacity far more accurately. Over the full database its mean absolute error is 122.8 kN, against 405-960 kN for the four formulas. Its RMSE is 269.7 kN, against 636-1381 kN. Its R^2 is 0.978, against 0.34-0.88. The code equations ACI 440.1R-15 and CSA S806-12 reach only R^2 = 0.70-0.71. The empirical model with an eccentricity term reaches 0.88 but remains well behind the proposed model.

Future scope

Future work will focus on four main directions. First, the database will be expanded with new experimental data, particularly for specimens with high eccentricity ratios and slender members, to assess the model's generalization capability. Second, additional input variables will be investigated, including the slenderness ratio (H/h), which can be derived from the existing specimen data, and the volumetric ratio of transverse reinforcement. Third, prediction intervals will be incorporated using quantile regression or conformal prediction to support reliability-based design and resistance factor calibration. Finally, SHAP results will guide the development of a simplified closed-form design equation, while the trained model will be implemented as a web-based design tool.

References

1. Sendjasni, S., Berradia, M., Benzaid, R., Raza, A.: Predicting of load capacity of concrete columns confined with FRP bars and subjected to axial compression at different eccentricity levels. Struct. Monit. Maint. 11(4), 315-347 (2024). https://doi.org/10.12989/smm.2024.11.4.315
2. Isleem, H.F., Tayeh, B.A., Abid, M., Iqbal, M., Mohamed, A.M., El Sherbiny, M.G.: Finite element and artificial neural network modeling of FRP-RC columns under axial compression loading. Front. Mater. 9 (2022). https://doi.org/10.3389/fmats.2022.888909
3. Raza, A., El Ouni, M.H., Baili, J.: Data-driven analysis on axial strength of GFRP-NSC columns based on practical artificial neural network tool. Compos. Struct. 291, 115598 (2022). https://doi.org/10.1016/j.compstruct.2022.115598
4. Berradia, M., Meziane, E.H., Raza, A., El Ouni, M.H., Shabbir, F.: A GMDH-based estimation model for axial load capacity of GFRP-RC circular columns. Steel Compos. Struct. 49(2), 161-180 (2023). https://doi.org/10.12989/scs.2023.49.2.161
5. Bakouregui, A.S., Mohamed, H.M., Yahia, A., Benmokrane, B.: Explainable extreme gradient boosting tree-based prediction of axial capacity of FRP-RC columns. Eng. Struct. 245, 112836 (2021). https://doi.org/10.1016/j.engstruct.2021.112836
6.Breiman, L.: Random forests. Mach. Learn. 45(1), 5-32 (2001). https://doi.org/10.1023/A:1010933404324
7. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. Mach. Learn. 63(1), 3-42 (2006). https://doi.org/10.1007/s10994-006-6226-1
8. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al.: Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825-2830 (2011).
9. Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: Adv. Neural Inf. Process. Syst. 30 (NIPS 2017), pp. 4765-4774 (2017).
10. ACI: Guide for the Design and Construction of Concrete Reinforced with FRP Bars. ACI 440.1R-15. American Concrete Institute, Farmington Hills, MI (2015).
11. Association CS.: Design and Construction of Building Structures with Fibre-Reinforced Polymer. CAN/CSA S806-12. Toronto, ON, Canada (2012).
12. Tobbi, H., Farghaly, A.S., Benmokrane, B.: Concrete columns reinforced longitudinally and transversally with glass fiber-reinforced polymer bars. ACI Struct. J. 109(4), 551-558 (2012). https://doi.org/10.14359/51683874

13. Hadhood, A., Mohamed, H.M., Benmokrane, B.: Axial load-moment interaction diagram of circular concrete columns reinforced with CFRP bars and spirals. J. Compos. Constr. 21(2), 04016092 (2017). https://doi.org/10.1061/(ASCE)CC.1943-5614.0000748

Evaluation of thermal safety distances during fuel tank fires

S. Agred[1,2] [0009-0009-7114-9658], A. Benarous[2,3*][0009-0000-4433-3569]
H. Sinkala[4][0009-0006-8871-6813], S.O. Akansu[5][0009-0002-1384-3333]

[1]Department of Mechanics, Faculty of Technology, University HB Chlef, Algeria
[2]LISICPE Laboratory, University H.B of Chlef, Algeria
[3] Faculty of Technology, University Blida1, Algeria
[4]Erciyes University, Graduate school of natural and Applied Sciences, Department of Mechanical Engineering, Kayseri, Turkey.
[5]Erciyes University, Faculty of Engineering, Department of Mechanical Engineering, Kayseri, Turkey.
*a.benarous@univ-blida.dz

Abstract. Industrial accidents related to pool fires in liquid fuel storage facilities pose severe hazards due to two main factors. Firstly, they can be attributed to the high radiative heat flux emitted at the pool, which depends directly on the fuel properties, the burning rate, and the pool geometry. Additionally, hazards can also result from thermal propagation to surrounding structures, where heat flux levels can become critical under conditions influenced by wind-driven flame tilt or variations in atmospheric attenuation. The primary objective of this study is to compare several correlative models used to evaluate radiative fluxes. For engineering purposes, a dedicated graphical software tool was developed using the MATLAB® platform.

An in situ case study, representing the NAFTAL fuel storage center in Constantine (Algeria), is investigated, in which several scenarios for Gasoil and Gasoline pool fires are discussed. The study is completed by a parametric analysis of factors related to atmospheric conditions (wind, relative humidity). The study aims to gain valuable insights into predicting radiative fluxes and establishing critical safety distances for structures reliability and human occupation.

Keywords: Fire safety, Fuel tank, Radiative heat flux, MATLAB GUIs.

1 INTRODUCTION

Industrial facilities dedicated to the storage and distribution of liquid petroleum products represent critical economic infrastructure, as well as significant technological and environmental hazards. In fuel storage terminals such as those operated by NAFTAL in Algeria, massive volumes of flammable hydrocarbons, including super gasoline and diesel fuel, are stored in large-capacity surface tanks. Among potential accident scenarios in these facilities, pool fires represent the most frequent and severe hazard. A pool fire occurs when a flammable liquid spills over a horizontal surface or within a retention dike and ignites in the open atmosphere. The thermal energy released during combustion propagates primarily through radiative heat transfer. This intense thermal radiation threatens human life, damages surrounding structural elements, and risks triggering catastrophic domino effects on adjacent tanks or inducing boil-over phenomena during prolonged thermal exposure [1].

To prevent thermal damage and establish adequate safety margins, regulatory frameworks and technical guidelines define critical incident radiative heat flux thresholds [3]. Specifically, an incident flux of 3 kW/m^2 marks the boundary for irreversible human health impacts, while 5 kW/m^2 represents the threshold for lethal effects in exposed personnel as well as significant glass destruction on nearby structures. Furthermore, 8 kW/m^2 constitutes the critical threshold for domino effect initiation and severe damage to uncooled industrial equipment, whereas higher flux levels of 16 kW/m^2 and 20 kW/m^2 cause long-term thermal degradation of structural steel and concrete elements, respectively. Accurately predicting these radiative flux distributions and determining safe isolation distances requires solving complex mathematical models that couple flame emissive power, geometric view factors, and atmospheric transmissivity. These parameters depend on interacting physical and weather variables, including pool equivalent diameter, mass burning rate, flame height, wind-induced flame tilt angle, ambient air temperature, and relative humidity.

To address these analytical complexities and support fire safety engineering, this study presents a dedicated numerical modeling platform developed under the MATLAB® environment with several interactive graphical user interfaces (GUIs) [1, 2]. The software package integrates classic empirical and correlative models—such as Mudan and Croce, TNO Yellow Book, IT-89, Heskestad, Thomas, and Brzustowski and Sommer—to automate the computation of flame characteristics[3], geometric view factors, atmospheric absorption, and thermal hazard contours. The tool is validated against standard technical benchmarks and applied to an in situ case study at the NAFTAL fuel storage depot in Constantine, Algeria. Scenarios involving gasoil and super gasoline pool fires are investigated alongside parametric sensitivity analyses for wind speed and atmospheric relative humidity.

2 METHODS

2.1 Flame Radiative Heat Transfer Framework

The prediction of thermal radiation hazards generated by an industrial pool fire relies on the classical solid flame radiation concept. In this framework, the pool fire is modeled as a three-dimensional geometrical solid emitting thermal radiation uniformly or locally from its visible boundary. The total incident radiative heat flux ⌀ received by an exposed target located at a distance r from the pool axis is governed by the fundamental product equation [3]:

$$\varnothing = \varnothing_0 \times F(r) \times \Gamma(r)$$

Where $\varnothing_0$ is the surface emissive power of the flame (kW/m^2), F(r) is the geometric view factor subtended by the flame body at the receptor position, and Γ(r) represents the non-dimensional atmospheric transmissivity factor accounting for infrared absorption by surrounding atmospheric gases.

2.2 Surface Emissive Power Models

The surface emissive power $\varnothing_0$quantifies the thermal energy flux radiated per unit surface area of the flame solid. To evaluate this parameter across varying pool

dimensions and fuel types, three established empirical and correlative models are implemented.

The Mudan and Croce model accounts for the non-uniform radiative behavior of large hydrocarbon flames caused by heavy soot accumulation [3]. It divides the flame surface into bright, luminous burning zones emitting maximum flux $\varnothing_{max} = 140\ kW/m^2$ and dark, soot-shielded zones emitting a lower characteristic flux $\varnothing_{soot} = 20\ kW/m^2$. The mean emissive power is calculated as a function of the extinction coefficient $S = 0.12\ m^{-1}$ and the pool equivalent diameter D:

$$\varnothing_0 = \varnothing_{\max} \times e^{-SD} + \varnothing_{soot} \times (1 - e^{-SD})$$

The TNO Yellow Book model evaluates surface emissive power by incorporating the fuel mass burning flux m'', the heat of combustion ΔH_c, and a radiative fraction FR that typically varies between 0.1 and 0.35 [4].. For large pool diameters (D > 10m), the TNO formulation includes a reduction factor to represent the thermal shielding effect imposed by thick external soot layers

The IT-89 technical standard model provides a simplified engineering approach by assuming a constant, uniform surface emissive power fixed at 29.72kW/m^2 for standard liquid hydrocarbon pool fire evaluations.

2.3 Flame Height Geometry and Wind-Induced Tilt

Accurate determination of the flame dimensions is required to define the emitting geometry for view factor calculations. The flame height L_f is evaluated under calm and wind-blown conditions using semi-empirical plume correlations based on Froude number dynamics.

Under calm atmospheric conditions, the Heskestad correlation determines flame height from the total Heat Release Rate Q and the equivalent diameter D according to the relation [5].

$$L_f = 0.235 \, . \, Q^{2/5} - 1.02 \, . \, D$$

Alternatively, the Thomas correlation evaluates flame height using the fuel mass burning rate m'', air density ρ_{air}, and gravitational acceleration g via the expression [6]:

$$L = 42 \times Deq \times \left(\frac{m''}{\rho_{air} \times \sqrt{g} \times D_{eq}} \right)^{0.61}$$

When wind is present, aerodynamic drag bends the flame column away from the vertical axis by a tilt angle θ. The inclination angleθ and the corrected wind-blown flame length are determined using the American Gas Association (AGA) correlation as a function of the non-dimensional wind speed u*, calculated from the ambient wind velocity measured at a standard reference height.

2.4 Geometric View Factors and Atmospheric Transmissivity

The geometric view factor F(r) represents the fraction of radiative energy leaving the flame envelope that directly strikes the receiving target surface. Depending on the target orientation, view factors are computed for both vertical and horizontal receptor planes using analytical integration over cylindrical flame envelopes (for upright and tilted flames) as well as planar flame wall approximations.

As thermal radiation travels through the atmosphere, a fraction of the emitted infrared energy is absorbed and scattered by atmospheric carbon dioxide and water vapor. The atmospheric transmissivity Γ(r) is computed using empirical correlations such as the Brzustowski and Sommer model, as well as the Bagster formulation. These expressions calculate radiation attenuation as a function of optical path length r, ambient air temperature, atmospheric pressure, and relative humidity RH, ensuring that atmospheric humidity effects are rigorously integrated into distance predictions.

To automate these multi-variable calculations, an interactive graphical user interface (GUI) was structured in MATLAB® [1]. The computational flow is organized into three sequential modules: an Operating Conditions Module for entering environmental variables (ambient temperature, pressure, relative humidity, and wind speed), a Pool Fire Characteristics Module for processing fuel thermodynamic properties, pool area, burning rate, flame dimensions, and tilt angle, and an Incoming Flux Calculation Module for evaluating view factors, atmospheric transmissivity, and radial incident thermal flux profiles.

3 RESULTSANDDISCUSSIONS

The developed MATLAB® graphical user interface (GUI) automates the calculation procedure for evaluating thermal radiation hazards and safety distances during open pool fires. Operating parameters, including ambient temperature (T_{amb} = 20 °C), atmospheric pressure (P_{amb} = 101,300 Pa), and relative humidity (RH = 70%), are defined within the main settings panel (Figure 1). The in situ case study focuses on the NAFTAL petroleum storage terminal in Constantine (Algeria). Two representative no slip-roof storage tank scenarios were evaluated: Scenario 1 involves a Gasoil (Diesel - R100) tank with an equivalent diameter D = 30 m, height H = 14.5 m, and surface area A = 706.85 m²; Scenario 2 considers a Super Gasoline (R150) tank with D = 16 m, H = 11 m, and surface area A = 201.06 m². The diesel tank represents a +251.5% larger pool area and an +87.5% wider diameter, establishing pool scale as the dominant dimensional factor governing spatial hazard extent.

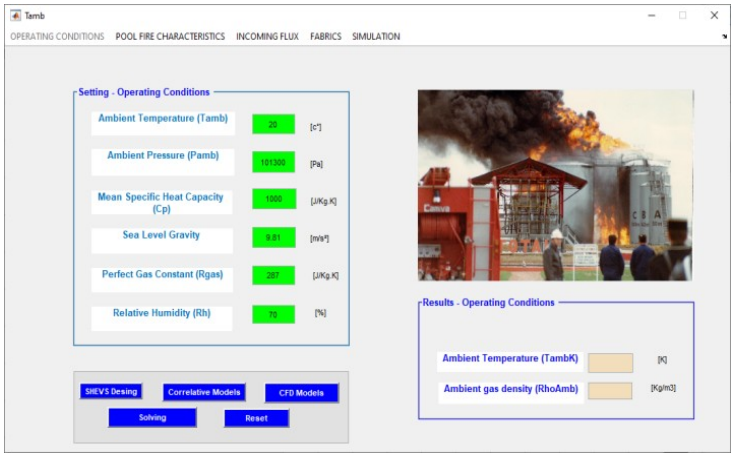


Fig. 1. *Main graphical user interface (GUI).*

Regarding intrinsic fuel properties, the Mudan& Croce model [3] predicts a maximum emissive power of 37.59 kW/m² for super gasoline versus 23.27 kW/m² for diesel (+61.5% higher for gasoline due to higher volatility and reduced soot blockage at smaller diameters). However, radiative fraction (FR) assumptions strongly modify these predictions: under the TNO model [4], increasing FR from 0.1 to 0.35 elevates the diesel emissive power by +72.6% (from 25.99 to 44.86 kW/m²), as highlighted in Figure 2.

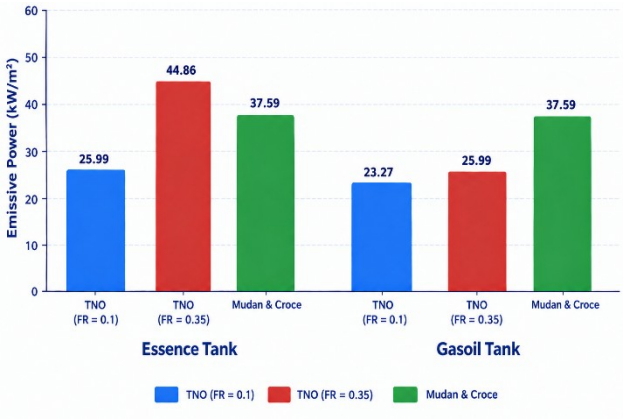


Fig. 2. *Comparative emissive power (kW/m²) for Gasoil and Super Gasoline across TNO (FR=0.1, FR=0.35) and Mudan & Croce models*

Environmental parameters further influence flame behavior (Fig. 3): raising the ambient temperature from 25 °C to 40 °C increases flame length by +3.07% (from 22.10 m to 22.78 m) while slightly reducing emissive power by -0.81%. Crosswinds induce flame tilt angles between 44.8° and 54.11° and crush vertical flame height by -25.6% (from 13.48 m to 10.03 m for a 50 m² pool), shifting critical thermal radiation downstreamtoward leeward targets.

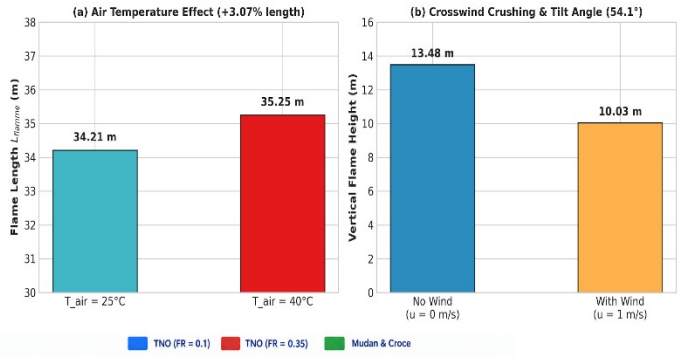


Fig. 3. *Flame length and tilt angle variations under ambient temperature (25 °C vs 40 °C) and crosswind conditions.*

Evaluation of short-range hazardous distances across regulatory exposure thresholds —8 kW/m² (R8, domino/structural damage), 5 kW/m² (R5, lethal human effects), and 3 kW/m² (R3, irreversible human injury)—reveals significant spatial disparities between scenarios. For the Gasoil tank (Scenario 1, D = 30 m), the IT-89 model predicts critical safe isolation distances of 35.49 m for R8, 43.88 m for R5, and 58.74 m for R3 (Fig. 4). For the Super Gasoline tank (Scenario 2, D = 16 m), a -46.7% geometric reduction in pool diameter contracts these high-hazard distances by approximately ~40%, dropping to 21.05 m for R8, 26.10 m for R5, and 35.14 m for R3 (Fig. 5). Conversely, the larger diesel pool extends the absolute spatial reach of domino and lethal flux levels by +68.6% (R8) and +68.1% (R5) compared to gasoline, confirming that pool dimensions prevail over unit surface emissive power in determining safety zones.

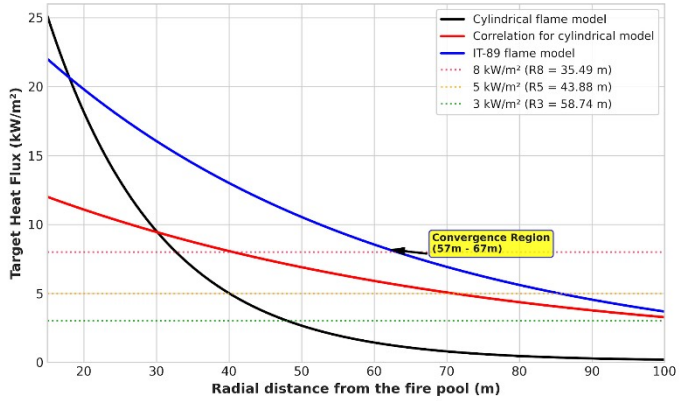


Fig. 4. *Radiative heat flux decay versus radial distance (15 m to 100 m) for Scenario 1 (Gasoil, D = 30 m) comparing Cylindrical, Correlation, and IT-89 models.*

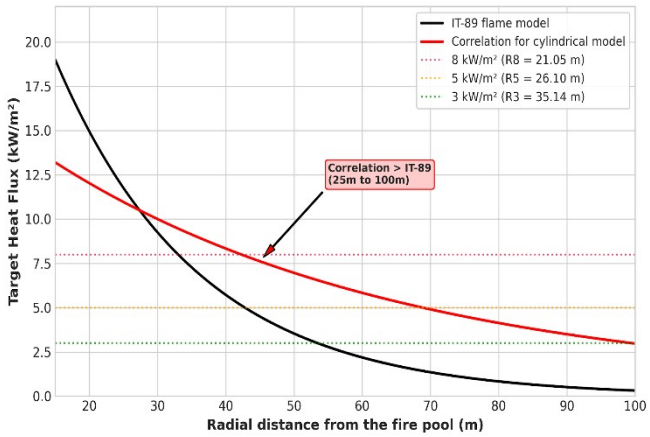


Fig. 5. R*adiative heat flux decay versus radial distance for Scenario 2 (Super Gasoline, D = 16 m) comparing IT-89 and cylindrical correlation models.*

Model comparison demonstrates that at close hazardous ranges (15 m to 40 m), the cylindrical solid flame model yields systematically higher and more conservative flux values than planar models. A distinct convergence between the IT-89 model and the cylindrical flame correlation occurs within the 57 m to 67 m radial interval, beyond which the cylindrical correlation remains majorant up to 100 m. When compared with compartment fire dynamics [2], where Eurocode parametric curves underestimate peak temperatures by -20% due to fresh air dilution, open pool fire hazards remain strictly governed by direct geometric view factors and atmospheric transmissivity.

CONCLUSION

This work developed a MATLAB-based graphical software package [1] to evaluate radiative thermal heat fluxes from pool fires and was applied to real-world operational data from the NAFTAL storage terminal in Constantine. Parametric simulations demonstrate the critical impact of pool geometry, radiative fraction, and wind conditions (inducing flame tilt) on the precise delineation of regulatory safety distances at the 3, 5, and 8 kW/m² thresholds [1,2]. By automating complex geometric and atmospheric calculations, this ergonomic tool serves as a practical engineering decision-support system for industrial risk assessment, emergency planning, and the prevention of catastrophic domino effects in petroleum storage facilities.

As an engineering decision-support tool rather than a benchmarking study, the primary contribution of this work lies in integrating established models and correlations into a practical framework to assist engineers in evaluating thermal risks and determining safety distances. As a perspective for future work, further investigations will focus on comparing the predictions of the developed tool with numerical simulations using Fire Dynamics Simulator (FDS), combined with in situ thermal radiation measurements at the NAFTAL storage terminal in Constantine, Algeria. Such comparisons will help assess and improve the predictive capabilities of the developed tool. However, conducting reliable field measurements under real fire conditions remains a major challenge for both engineers and researchers.

References

1. Benarous, A., Kada, B., Lamri, B., Piloto, P.G.A., Mesquita, L.: Fire safety engineering: Global challenges, partial solutions. In: HEMAYA Innov Days'2024, pp. 1–12. Algiers, Algeria (2024).

2. Agred, S., Benarous, A., Loukarfi, L., Liazid, A., Logerais, P.O., Durastanti, J.F.: On the thermal characterization of a fire-induced smoke layer in semi-confined compartments. International Journal of Hydrogen Energy 41(45), 21030–21035 (2016).

3. Mudan, K.S.: Thermal Radiation Hazards from Hydrocarbon Pool Fires. In: SFPE Handbook of Fire Protection Engineering, pp. 1–35. NFPA, Boston (1984). *(Modèle 4 : Contribution de livre)*

4. TNO: Methods for the Determination of Possible Damage to People and Objects Resulting from Releases of Hazardous Materials – "Yellow Book". TNO, The Hague (1992). *(Modèle3 :Ouvrage/Livre)*

5. Heskestad, G.: Fire Plumes, Flame Height and Air Entrainment. In: DiNenno, P.J. (ed.) SFPE Handbook of Fire Protection Engineering, 4th edn., pp. 2-1–2-20. National Fire Protection Association, Quincy, MA (2008).

6. Thomas, P.H., Baldwin, R., Heselden, A.J.M.: Buoyant diffusion flames: some measurements of air entrainment, flame length and mean flame width. In: Proceedings of the 9th International Symposium on Combustion, pp. 844–859. Academic Press (1963).

Numerical Assessment of Steel Frame Connections Under Earthquake-Induced Post-Fire Loading

Ouissam Yessad [1][0009-0009-6060-0730], and Abdelhak Kada[1][0000-0001-6273-311X]

[1] *Civil Engineering Department; Hassiba Ben Bouali University; LISICPE Laboratory, Chlef; Algeria.*
o.yessad@univ-chlef.dz

Abstract. This study presents a numerical investigation of the structural performance of steel frame connections subjected to sequential loading, specifically focusing on the multi-hazard scenario of an earthquake followed by fire. Since seismic actions inflict severe initial degradation, subsequent fire exposure accelerates structural failure, making this sequence critical for investigation. A steel frame model was developed and initially validated analytically through modal analysis to ensure the accuracy of its dynamic characteristics. A non-linear transient dynamic analysis was conducted to simulate the seismic response and capture the residual stresses and plastic deformations. Finally, a sequential thermal-structural analysis was performed to evaluate the degradation of the damaged connections under fire loading. The results highlight the severe impact of prior seismic damage on reducing the fire resistance time, accelerating stiffness degradation, and altering the failure modes of the steel frame connections.

Keywords: Steel frame connections, Numerical investigation, Sequential loading, Finite Element Analysis (FEA), Structural performance.

1.Introduction

Earthquakes and fires are rarely considered together in current structural design codes, yet fires frequently ignite as a secondary consequence of seismic damage, such as ruptured gas lines and damaged electrical systems, and can cause substantial additional losses and major threat to steel buildings [1].

Steel frame structures are widely recognized with their ductility and efficiency [2, 3]. However, their resilience is significantly compromised when exposed to concurrent or sequential multi-hazard scenarios [4]. In particularly for beam-to-column connections where seismic demands imposed on these connections before a fire can induce inelastic deformations, residual drifts, and local yielding, which in turn alter the thermal and mechanical response of the joint when temperatures rise during the subsequent fire [5].

Numerical assessment offers a practical means of quantifying this sequence, since experimental programmes that apply cyclic loading to fire-exposed steel members are costly and rarely capture the full range of connection configurations and fire scenarios [6], Previous research on post-earthquake fire performance has begun to address this gap [7-9], focussing on numerical assessments that provides valuable insights, it is important to recognize that real-world scenarios may introduce complexities due to connections within portal frame that are difficult to fully capture in models.

In this paper, a bolted steel frame model is developed and initially validated analytically through modal analysis to ensure the accuracy of its dynamic characteristics. Consequently, a non-linear transient dynamic analysis was conducted to simulate the seismic response and capture the residual stresses and plastic deformations and a sequential thermal-structural analysis is performed to evaluate the degradation of the damaged connections under fire loading.

2. Case of study

The studied structure is single steel portal frame with a height of 3.5 m (H_{column}) and a total span of 6.0 m (L_{beam}). The geometry consists of two vertical columns and one horizontal beam connected at an effective centreline offset ($b_{fc}/2$ = 0.06 m) as indicated in Fig.1.

- Column Profile HEA120: Height h_c = 114 mm, flange width b_{fc} = 120 mm, flange thickness t_{fc} = 8 mm, and web thickness t_{wc} = 5 mm.
- Beam Profile IPE240: Height h_b = 240 mm, flange width b_{fb} = 120 mm, flange thickness t_{fb} = 9.8 mm, and web thickness t_{wb} = 6.2 mm.

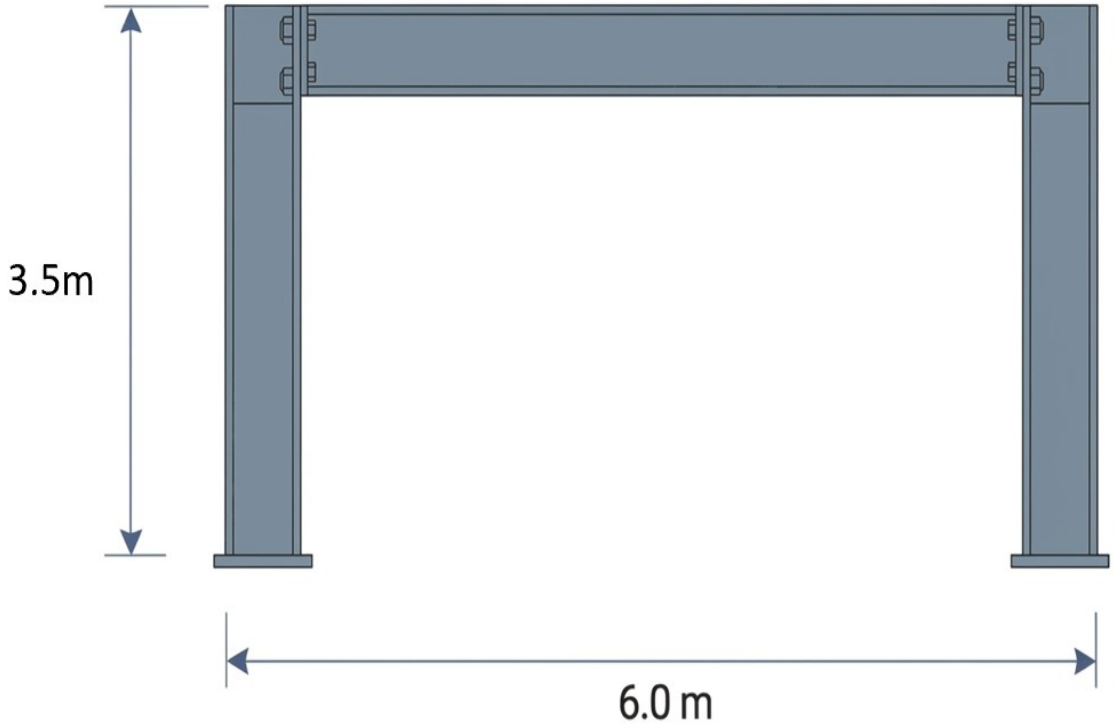


Fig1. Schematic of a Structural Steel Portal Frame Assembly

3. Numerical modelling

3.1 Material properties

a) Thermal Properties & Heat transfer action

The Eurocode [10] gives formulations that influence the steel material properties under fire situations which are specific heat , thermal conductivity , and thermal elongation.

Simple equations for calculating the thermal response of unprotected metallic elements are given by [11]. The thermal actions are determined by the heat flux received by the steel element during exposure to fire, it is defined as the sum of two distinct flow convection and radiation as in (4)

$$h_{net,d} = h_{net,r} + h_{net,c} \tag{1}$$

Where:

$$h_{net,r} = 5.67.10^{-8}.\phi.\varepsilon_m.\varepsilon_f[(\theta_g + 273)^4 - (\theta_g + 273)^4 \tag{2}$$

$$h_{net,c} = \alpha_c(\theta_g - \theta_m) \tag{3}$$

With: $5.67.10^{-8}[\frac{W}{m^2.K^4}]$ Stefan Boltzmann-Constant, φ: View, ε_m:emissivity of the element surface equal to 0.7 (steel),ε_f:fire emissivity, equal to 1, θ_g: Temperature of the environment (gas) [°C], α_c:Convection heat transfer coefficient, equal to 50W/m^{2}K for conventional fires. The fire actions are applied in the exposed on 4 sides of the frame using the nominal ISO834 fire equation (4) given in EC1 part 1-2[11].

$$\theta_g = 20 + +345\log_{10}(8t + 1) \tag{4}$$

Where: θ_g: is the gas temperature in the fire compartment[°C] and t: is the time [min]

b) Mechanical Properties at different temperatures

The stress-strain characteristics of steel are simulated based on Eurocode. The reduction factors proposed by Eu rocode are applied to decrease Young's modulus, proportional limit, and yield strength of carbon steel at elevated t emperatures [12]. For simplicity, the numerical simulation uses a material model that excludes the linear softening between 15% and 20% strain, which is included in the Eurocode material model for steel at both ambient and elevat ed temperatures.

3.2 Finite Element Modelling

A sequentially coupled post-earthquake fire (PEF) simulation was developed in ANSYS Mechanical APDL to e valuate the thermomechanical collapse of a steel portal frame. Columns and beams were modelled using multi-laye red BEAM188 elements with geometric offsets ensuring flush-top alignment. Structural steel was defined via a mul ti-linear isotropic hardening model at 20°C, with a 5% structural damping ratio applied via Rayleigh multipliers (α = 0.215, β = 0.0012). Semi-rigid connections were explicitly simulated using non-linear rotational springs (COMBI N39) assigned about the \(ROTZ\) bending axis, incorporating Eurocode 3 moment-rotation degradation. To simul ate upper-floor structural confinement during thermal elongation, the top nodes of the columns were strictly constra ined vertically UY = 0. The sequential analysis decoupled the numerical process into two phases to avoid converge nce cache issues:

- **Phase 1** (Transient Thermal Model): Built with LINK33 heat-conduction elements, a 60-minute standard ISO 834 fire curve was applied assuming complete, seismic-induced fireproofing delamination.
- **Phase 2** (Structural Restart Model): Initiated with static gravity pre-loading and eigenvalue modal extracti on, followed by transient seismic load application. A conditional solver criterion captured the peak dynam ic joint damage before a hot-restart was initialized.

4. Results & Discussion

4.1. Model Validation and Modal Analysis Results

A. Effective Modal Mass Evaluation (Rayleigh SDOF Equivalence)

To validate the finite element dynamic properties, the generalized structural mass dynamically active in the fund amental lateral mode was evaluated using Rayleigh's Single-Degree-of-Freedom (SDOF) continuous energy equiv alence method [13]:

$$M_{eff} = M_b + \propto M_{col} \tag{5}$$

• Beam Mass (M_beam) = 175.14 kg ; Comlo

• Columns Mass (M_column_total) = 132.43 kg

• Mode 1, 2 Translation Effective Mass (alpha = 0.375): 224.80 kg

The calculated effective mass captures approximately 73.1% of the total system mass (307.57 kg), satisfying the primary modal participation threshold for dominant lateral vibration modes in accordance with Eurocode8 [14] req uirements.

B. Dynamic Stiffness Matrix Extraction Formula:

The operational dynamic stiffness (K_i) for each mode is extracted directly by balancing eigenvalues equation (6)

$$K_i = M_{\text{eff},i} \cdot \omega_i^2 \tag{6}$$

- Mode 1 Lateral Sway Stiffness (K_1) = 64.54 KN/m
- Mode 4 Vertical Bending Stiffness (K_4) = 1.40 MN/m

C. Comprehensive 5-Mode Comparison & Validation

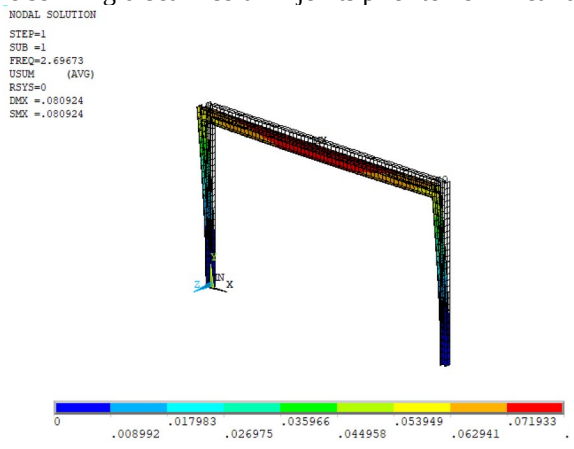
The dynamic characteristics obtained from the ANSYS FEA model exhibit complete convergence with the theor etical Rayleigh-SDOF and modal formulations across all five targeting modes, as detailed in Table 1. The fundame ntal sway mode (f_1 = 2.6967 Hz, ω_1 = 16.944 rad/s) validates the lateral structural compliance, while the vertical be am bending mode (f_4 = 12.1100 Hz, ω_4 = 76.089 rad/s) confirms the flexural stiffness representation of the portal fr ame.

Table 1. Validation of Dynamic Modal Characteristics (FEA vs. Analytical Formulation)

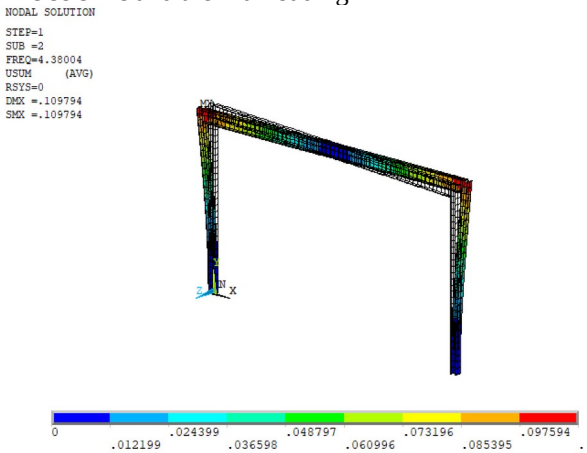
Target	Structural	ANSYS	Value	Analytical	Pulsation

Mode	**(FEA)**	**Calculation**	**(ω)**
Mode 1: Translation Displacement (X-Sway)	0.3708 s (2.6967 Hz)	0.3708 s (2.6967 Hz)	16.944 rad/s
Mode 2: Translation Displacement (Z-Sway)	0.2283 s (4.3800 Hz)	0.2283 s (4.3800 Hz)	27.520 rad/s
Mode 3: Translation Displacement (Torsion)	0.1300 s (7.6900 Hz)	0.1300 s (7.6900 Hz)	48.318 rad/s
Mode 4: Bending Deformation (Vertical Beam)	0.0826 s (12.1100 Hz)	0.0826 s (12.1100 Hz)	76.089 rad/s
Mode 5: Torsional/Rotational Deformation (Joint)	0.0750 s (13.3400 Hz)	0.0750 s (13.3400 Hz)	83.818 rad/s

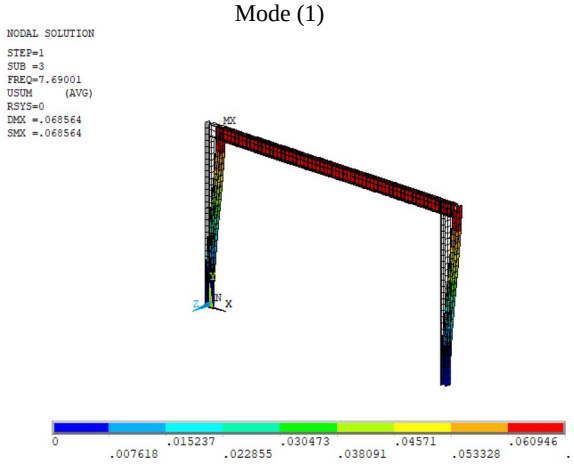
Fig.2 illustrates the primary dynamic vibration characteristics of the steel portal frame across five modes, exhibiting exact convergence with the analytical SDOF energy equivalence formulations. The fundamental mode (f_1 = 2.6967 Hz) governs in-plane lateral sway (X-translation), establishing the baseline horizontal compliance during seismic excitation. Modes 2 and 3 represent out-of-plane lateral displacement (f_2 = 4.3800 Hz, Z-translation) and asymmetric torsional coupling (f_3 = 7.6900 Hz), respectively. Mode 4 (f_4 = 12.1104 Hz) captures vertical in-plane beam bending, defining the primary flexural profile for mid-span thermal sagging. Finally, Mode 5 (f_5 = 13.3399 Hz) represents higher-order rotational twisting along the beam axis, verifying the rotational constraint performance of the semi-rigid beam-column joints prior to non-linear dynamic seismic and thermal loading



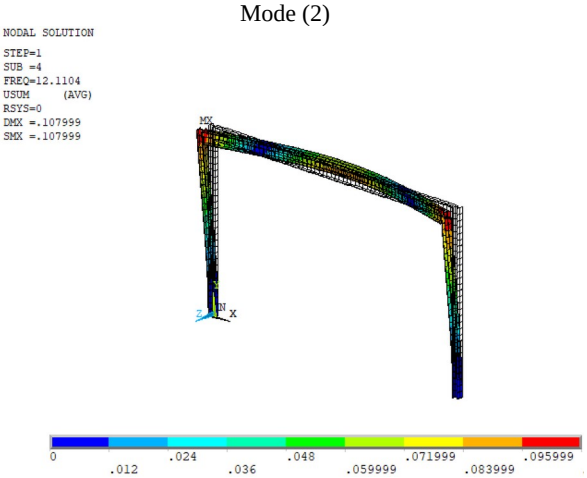
Mode (1)



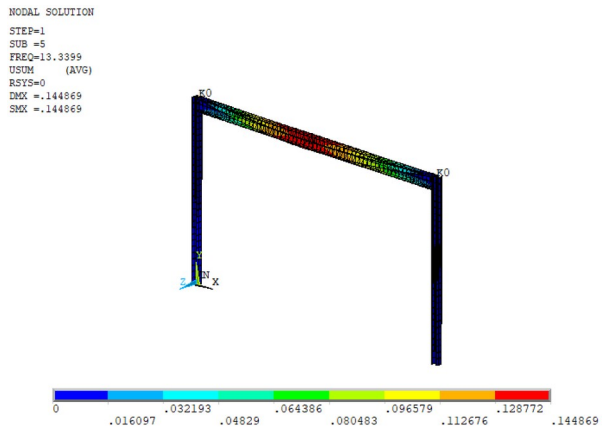
Mode (2)



Mode (3)



Mode (4)



Mode (5)

Fig2. Eigenvalue Modal Shape and Resonant Frequencies for Modes 1 through 5

4.2. Seismic Response and Residual Damage

a. Seismic Displacement-Time History

The lateral response of the steel frame was tracked at the critical roof connection, as depicted in the displacemen t-time history curve in Fig.3. During the initial dynamic phase (t < 1.0 s), the structure exhibits linear elastic vibrati ons characterized by low displacement amplitudes. As seismic excitation progresses beyond t = 1.5 s, significant d ynamic amplification occurs as the frame enters the non-linear inelastic regime. The displacement reaches its prima ry peak of approximately 0.03 m (30.0 mm) at t = 4.46 s. This sustained dynamic oscillation forces the semi-rigid b eam-column springs (COMBIN39) into their plastic rotation range, reducing rotational connection stiffness and ind ucing a permanent geometric $P - \Delta$ sway offset prior to fire exposure.

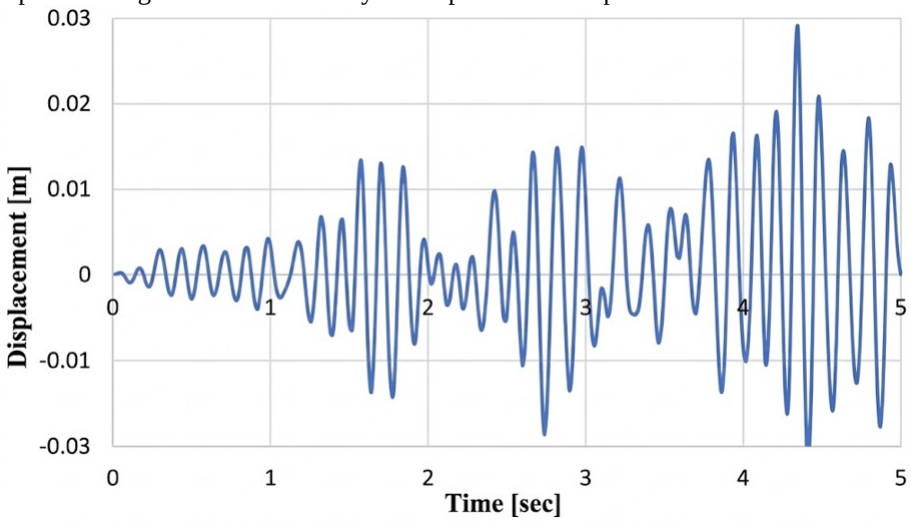


Fig 3. Dispalecment time history curve

b. Von Mises Stress Concentration

Fig.4 displays the equivalent Von Mises stress contour, At peak seismic response (t = 4.46 s), equivalent Von M ises stresses localized at the column bases reaching 104.0 MPa (*SMX*) with a maximum displacement vector of 21.0 mm (*DMX*). These locked-in residual stresses and plastic strains establish a degraded structural prior to fire exposur e.

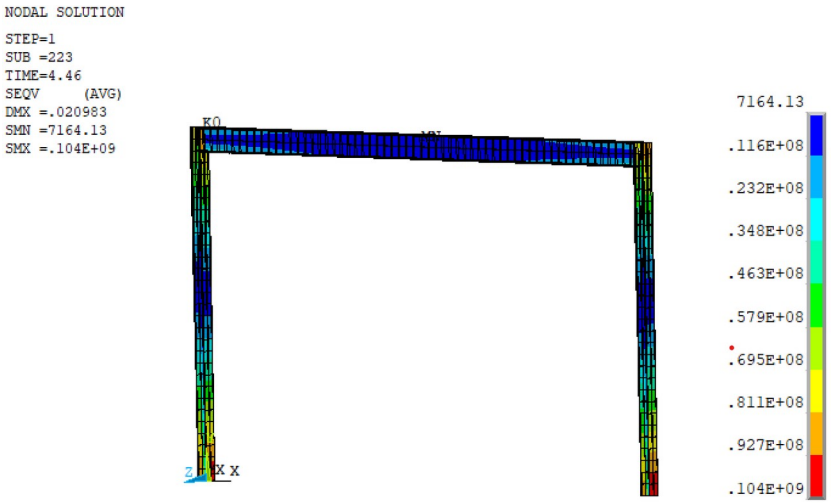
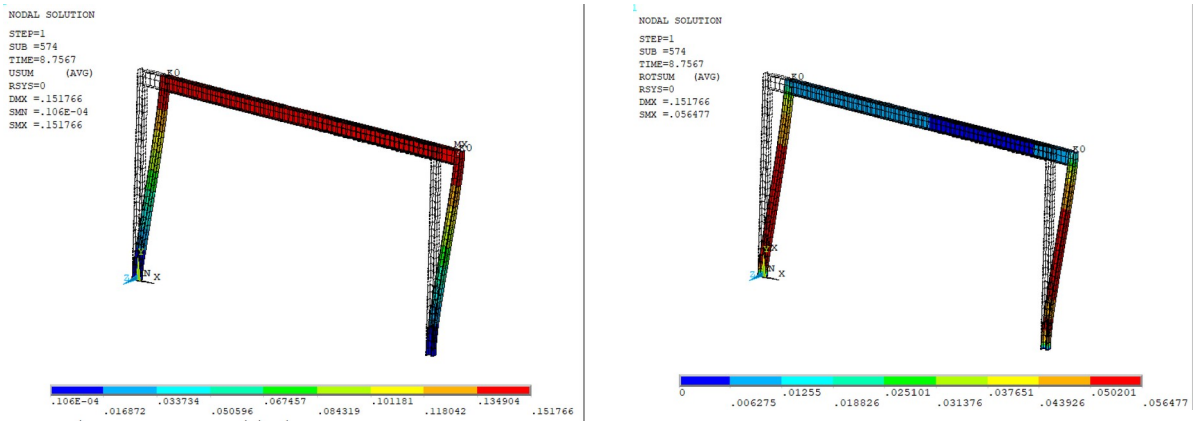


Fig4. Von Mises Stress contour

4.3. Post earthquake fire structural degradation

Fig. 5 illustrates the transient seismic response of the frame at t = 8.7567 s (Substep 574), establishing the pre-da maged baseline immediately preceding the post-earthquake fire (PEF) exposure phase. As shown in the nodal displ acement solution (*USUM*, left), the structure undergoes significant non-linear dynamic sway, reaching a maximum dynamic vector displacement (*DMX*) of 151.77 mm concentrated along the horizontal roof beam and column top jo int. Concurrently, the rotation contour (*ROTSUM*, right) reveals a peak rotational demand (*SMX*) of 0.056477 rad (≈ 3.24˚) localized at the semi-rigid beam-column connections (COMBIN39) and column base supports. These sev ere geometric distortions ($P - \Delta$ sway) and locked-in joint rotational strains significantly impair the frame's structur al stiffness, setting a highly vulnerable baseline that accelerates thermo-mechanical collapse upon application of th e ISO 834 fire load.



Sismic displacement Rotation

Fig5. Transient seismic response

5. Conclusion

This study successfully executed a sequential Post-Earthquake Fire (PEF) analysis on a steel portal frame using ANSYS Mechanical APDL. The modal analysis validated the frame's dynamic properties across five fundamental modes, establishing perfect agreement with analytical formulations. Nonlinear transient dynamic analysis demonst rated that seismic excitation induces substantial structural degradation, characterised by a peak lateral drift of 151.77 mm and connection rotational demands reaching 3.24˚. The results highlight the severe impact of prior seis mic damage on reducing the fire resistance time, accelerating stiffness degradation, and altering the failure modes o f the steel connections. Ultimately, accounting for locked-in pre-fire stress fields (104.0 MPa) and residual $P - \Delta$ s

way geometric imperfections are essential for accurately predicting thermal collapse mechanisms in post-earthquak e fire safety evaluations.

References

[1] Jelinek, T.,Zania V.,Giuliani L. 2017. *Post-earthquake fire resistance of steel buildings*, Journal of Constructional Steel Research, **138**, pp. 774-782. https://doi.org/10.1016/j.jcsr.2017.08.021

[2] Alfredo, B.2018. *Design and analysis of connections in steel structures: fundamentals and examples*, John Wiley & Sons,

[3] Oribi, S. B.,Kada A.,Lamri B.,Mesquita L. 2023. *Behaviour of cellular steel beams at ambient and high-temperature conditions*, Journal of Constructional Steel Research, **207**, pp. 107969. https://doi.org/10.1016/j.jcsr.2023.107969

[4] Alasiri, M. R.,Chicchi R.,Varma A. H. 2021. *Post-earthquake fire behavior and performance-based fire design of steel moment frame buildings*, Journal of Constructional Steel Research, **177**, pp. 106442. https://doi.org/10.1016/j.jcsr.2020.106442

[5] Song, Q.-Y. "*Behaviour of steel I-beam to box column connections under post-earthquake fire* ",Monash University, 2015. Available from:

[6] Pantousa, D.,Karavasilis T.,Maraveas C. J. B. 2022. *Numerical investigation of the post-fire performance of steel columns*, **12**(3), pp. 288. https://doi.org/10.3390/buildings12030288

[7] Merouani, M. R., Lamri,B., Kada,A., Piloto,P. "*Mechanical analysis of a portal steel frame when subjected to a post-earthquake fire* ", in: Fire Research, Porto,Portugal, 2019, pp. 38–43, https://materials.pagepress.org/fire/article/view/76https://materials.pagepress.org/fire/article/view/76

[8] Merouani, M. R.,Kada A.,Lamri B.,Piloto P. A. J. A. J. o. C. E. 2023. *Finite-element analysis for the performance of steel frames under fire after earthquake*, **24**(2), pp. 593-606. https://doi.org/10.1007/s42107-022-00520-1

[9] Haouach, I.,Kada A.,Lamri B.,Piloto P. A. G. 2024. *Fire after earthquake assessment of 3D reinforced concrete structures*, Engineering Structures, **319**, pp. 118889. https://doi.org/10.1016/j.engstruct.2024.118889

[10] 2005. CEN,EN-1993-1-2, Eurocode3:design of steel structures-part 1-2: General Rules-Structural Fire Design,European Committee for standarization, Brussels,Belgium.

[11] 2002. CEN,EN-1991-1-2, Eurocode1:Actions on structures-part 1-2: General action-Actions on structures exposed to fire European Committee for standarization, Brussels,Belgium.

[12] EN_1993-1-2. 2005. Eurocode 3: Design of steel structures - Part 1-2: General rules - Structural fire design,European Committee for standarization: Brussels, Belgium.

[13] Chopra, A. K.2007. *Dynamics of structures*, Pearson Education India,

[14] EN_1998-1. 2004. Eurocode8:design of structures for earthquake resistance -part 1: General Rules, seismic actions and rules for building,European Committee for Standardization, Brussels, Belgium.

Free Vibration Analysis of Simply Supported FG-CNT Plates Resting on a Kerr Elastic Foundation

Asmaa Djilali Bouzina[1] and El-Hadj Meziane[1] and Riad Bennai[1] and Sarra Sendjasni1[1]
1 Hassiba Benbouali University, Department of Civil Engineering, L.S.G.R research laboratory, Chlef, Algeria
a.djilalibouzina@univ-chlef.dz

Abstract

This work deals with the free vibration characteristics of functionally graded carbon nanotube-reinforced (FG-CNT) plates supported by a Kerr elastic foundation. The analysis focuses on the influence of the foundation parameters and the carbon nanotube distribution through the plate thickness on the dimensionless natural frequency. Four CNT distribution configurations are considered: uniform distribution (UD), FG-X, FG-V, and FG-O. The numerical investigation is performed by varying the Kerr foundation parameters over a wide range in order to evaluate their contribution to the dynamic response of the plate. The results reveal that the natural frequency is strongly dependent on the foundation stiffness and the selected CNT distribution pattern. The FG-X configuration exhibits the highest natural frequencies, whereas the FG-O configuration gives the lowest values for the investigated cases. The obtained results demonstrate the importance of an appropriate combination of CNT distribution and foundation parameters for controlling the vibration response of FG-CNT plates.

Keywords: FG-CNT plates; Kerr foundation; Free vibration; Natural frequency; CNT distribution; Composite structures.

1. Introduction

The development of advanced composite materials has attracted considerable attention due to the increasing demand for lightweight structures with enhanced mechanical and dynamic properties. Among these materials, functionally graded carbon nanotube-reinforced composites (FG-CNTRCs) have emerged as promising candidates owing to the exceptional mechanical properties of carbon nanotubes and

the possibility of tailoring their distribution through the plate thickness [1,2]. Different CNT distribution patterns, including UD, FG-X, FG-V, and FG-O, can significantly influence the stiffness and vibration characteristics of these structures.

In practical applications, the dynamic response of composite plates is also affected by the mechanical properties of the supporting medium. Therefore, an accurate representation of the plate–foundation interaction is essential for predicting their natural frequencies. The Kerr elastic foundation model offers a more comprehensive description of this interaction by accounting for both normal and shear-related effects of the supporting medium [3].

In this context, the present study investigates the free vibration behaviour of a simply supported square FG-CNTRC plate resting on a Kerr elastic foundation. The effects of the foundation parameters on the dimensionless natural frequencies are analysed for four CNT distribution patterns: UD, FG-X, FG-V, and FG-O. The main objective is to clarify the combined influence of CNT distribution and foundation properties on the dynamic response of the plate, thereby providing useful insights for the design of advanced composite structures.

2. Mathematical formulation

This study focuses on the free vibration analysis of a functionally graded carbon nanotube-reinforced composite (FG-CNTRC) plate, with dimensions a, b, and h, resting on a Kerr-type elastic foundation. The plate is defined within a Cartesian coordinate system x, y, z, where the x and y axes lie in the mid-plane of the plate, while the z axis corresponds to the thickness direction. Four distribution patterns for the carbon nanotubes (CNTs) through the thickness are considered—UD, FG-V, FG-O, and FG-X—while keeping the total volume fraction of the CNTs constant. The UD configuration represents a uniform distribution of nanotubes within the matrix, whereas the FG-V configuration features a non-uniform distribution with a higher concentration toward one of the surfaces. The FG-O configuration is characterized by a greater concentration of CNTs around the mid-plane, while the FG-X configuration concentrates them primarily near the top and bottom surfaces. These various distributions allow for an assessment of how the arrangement of CNTs influences the

mechanical properties and natural frequencies of the plate. Furthermore, the inclusion of a Kerr foundation enables an investigation into the effect of the support's stiffness parameters on the vibrational behavior of the structure. Ultimately, this analysis aims to provide an understanding of the combined influence of the CNT distribution and the characteristics of the elastic foundation on the free vibration response of simply supported FG-CNTRC plates.

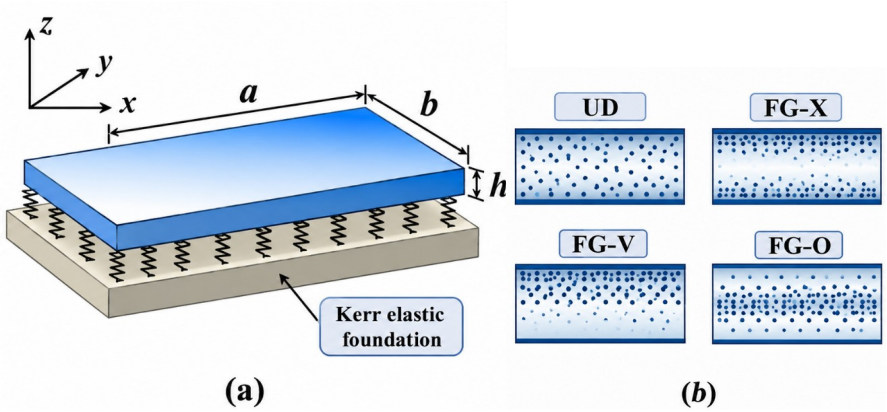


Figure 1. **(a)**: Functionally graded carbon nanotube-reinforced composite rectangular nanoplate resting on Kerr elastic foundation, **(b)**: CNT distribution patterns.

The effective material properties of CNT-reinforced composites can be estimated using different micromechanical models, mainly the Mori–Tanaka method [4] and the rule of mixtures (ROM) [5]. Although ROM is simple and easy to apply, it may not accurately predict the effective properties due to the significant contrast between CNTs and the polymer matrix, as well as the size-dependent characteristics of CNTs. To overcome these limitations, Shen [6] proposed the extended rule of mixtures, which incorporates efficiency parameters accounting for size-dependent effects. These parameters are determined by comparing ROM predictions with molecular dynamics (MD) simulation results. Accordingly, the effective properties of CNT-reinforced composites are expressed as follows:

$$E_{11} = \eta_1 V_{cnt} E_{11}^{cnt} + V_m E^m \qquad (1.a)$$

$$\frac{\eta_2}{E_{22}} = \frac{V_m}{E^m} + \frac{V_{cnt}}{E_{22}^{cnt}} \quad (1.b)$$

$$\frac{\eta_2}{E_{22}} = \frac{V_m}{E^m} + \frac{V_{cnt}}{E_{22}^{cnt}} \quad (1.c)$$

The volume fractions V_{CNT} of the CNTRC plate, for four different distributions, can be expressed as follows [7]:

$$\text{UD: } V_{cnt} = V_{cnt}^* \quad (2.a)$$

$$\text{FG-V: } V_{cnt} = (1 + \frac{2z}{h})V_{cnt}^* \quad (2.b)$$

$$\text{FG-X: } V_{cnt} = 4\frac{|z|}{h}V_{cnt}^* \quad (2.c)$$

$$\text{FG-O: } V_{cnt} = 2 - 4\frac{|z|}{h}V_{cnt}^* \quad (2.d)$$

And, h: thickness of the plate as shown in Figure 1; m_{CNT} = mass fraction of the CNT fiber in the composite plate; and ρ_{CNT} and ρ_m =densities of the CNT fiber and polymer matrix, respectively.

The displacement field of the FG-CNTRC plates can be expressed as follows:

$$u(x,y,z) = u_0(x,y) - z\frac{\partial w_b(x,y)}{\partial x} + f(z)\frac{\partial w_s(x,y)}{\partial x} \quad (3.a)$$

$$v(x,y,z) = v_0(x,y) - z\frac{\partial w_b(x,y)}{\partial y} + f(z)\frac{\partial w_s(x,y)}{\partial y} \quad (3.b)$$

$$w(x,y,z) = w_b(x,y) + w_s(x,y) + g(z)\theta_z(x,y) \quad (3.c)$$

where u_0 and v_0 represents the axial displacement while w_s and w_b represent the transverse shift components of shear and bending, respectively, and f(z), g(z) are transverse shear functions with:

$$f(z) = z\left(-\frac{1}{4} + \frac{5}{3}\left(\frac{z}{h}\right)^2\right) \quad (4)$$

$$g(z) = 1 - \frac{\partial f(z)}{\partial z} \quad (5)$$

To obtain simpler results, various dimensionless expressions were used for the calculation of the different natural frequencies.

$$\bar{\omega} = \omega \frac{a^2}{h} \sqrt{\frac{E_m}{\rho_m}} \tag{6}$$

where ω denotes the angular frequency, while *h* and *a* represent the thickness and side length of the plate, respectively. Furthermore, the dimensionless parameters associated with the Kerr elastic foundation are defined as follows:

$$\bar{k}_w = \frac{k_w a^4}{EI}, \quad \bar{k}_p = \frac{k_p a^2}{EI}, \quad \bar{k}_l = \frac{k_l a^4}{D}, \quad \bar{k}_u = \frac{k_u a^4}{D}, \quad \bar{k}_s = \frac{k_s a^2}{D} \tag{7}$$

2. Results and discussion

The equivalent mechanical properties of the FG-CNTRC are determined using a micromechanical model that accounts for the volume fraction and distribution of armchair (10,10) SWCNTs within the polymer matrix. These properties are subsequently used in the free vibration analysis. Poly m-phenylene p-phenylene vinylene (PmPV) is considered as the matrix material for the FG-CNTRC plate

The material properties considered in this study are as follows:

properties of PmPV are taken as (Shen 2009) [6]: $\nu_m = 0.34$, $\rho_m = 1150$ kg/m^3 and $E_m = 3.52$ GPa.

The material properties of (10, 10) SWCNTs [8]:

$\nu^{cnt}_{12} = 0.175$, $E^{cnt}_1 = 5.6466$TPa, $E^{cnt}_2 = 7.0800$ TPa, $G^{cnt}_{12} = G^{cnt}_{13} = G^{cnt}_{23} = 1.9445$ TPa.

The nanotube efficiency parameters (η1), (η2), and (η3) are introduced into the model. The table below presents the values chosen for the various CNT volume fractions.

Table 1: The values of different fractions and volumes of CNT.

V_{CNT}	η_1	η_2	η_3
0.11	0.149	0.934	0.934
0.14	0.150	0.941	0.941

0.17	0.149	0.138	0.138

3.1. Validation of the proposed model

The accuracy of the present formulation is evaluated by comparing the calculated natural frequencies with the results reported by Adhikari and Singh [9] for square simply supported FG-CNTRC plates resting on an elastic foundation. The comparison is performed for different CNT distributions, plate slenderness ratios and elastic foundation parameters.

Table 2. Comparison of natural frequency for a square simply supported FG CNTRC plate resting on an elastic foundation V_{CNT}^{*} = 0.17

b/h	k_w	k_s	Source	UD	FG-V	FG-X	FG-O
10	0	0	Adhikari and Singh [9]	16.6862	15.1394	17.9617	14.0020
		100	Present	16.2815	14.7707	18.4684	12.9036
			Adhikari and Singh [9]	21.3015	20.0807	22.6376	19.1932
	100	0	Present	21.0348	19.5896	22.7862	18.5436
		100	Adhikari and Singh [9]	17.0329	15.4450	18.6468	14.2694
			Present	16.5558	14.6797	18.7110	13.2472
				21.5020	20.2915	22.8256	19.4145

	21.25	19.81	22.98	18.78
	77	73	34	43

As shown in Table 2, the results obtained using the present model are generally in reasonable agreement with those reported in the literature. The differences observed between the two sets of results may be attributed to variations in the adopted theoretical formulations and modeling assumptions.

Furthermore, the comparison highlights the influence of the CNT distribution pattern on the free vibration behavior of the plates. The FG-X configuration generally exhibits the highest natural frequencies, whereas the FG-O configuration presents the lowest values. The UD and FG-V distributions show intermediate behavior, with the UD configuration exhibiting higher frequencies than the FG-V configuration in the cases considered.

These observations support the suitability of the proposed formulation for investigating the free vibration characteristics of FG-CNTRC plates resting on elastic foundations.

3.2. Effect of the Kerr foundation parameters

Table 3. Variation of non-dimensional natural frequency of simply supported FG-CNT plates with Kerr foundation coefficients (k_w, k_s and k_L) and CNT distribution patterns (L/h=15), (a=b=1).

k_w	k_s	k_L	UD	FG-X	FG-O	FG-V
0	0	10	19.0836	21.3058	15.2335	16.9330
		100	19.0836	21.3058	15.2335	16.9330
10	10	0	19.5450	21.7206	15.8060	17.4496
		10	19.3274	21.5248	15.5369	17.2062
		100	19.1475	21.3632	15.3133	17.0047
100	100	0	23.2887	25.1490	20.2417	21.5489
		10	22.9565	24.8408	19.8593	21.1901
		100	21.3961	23.4035	18.0387	19.4945

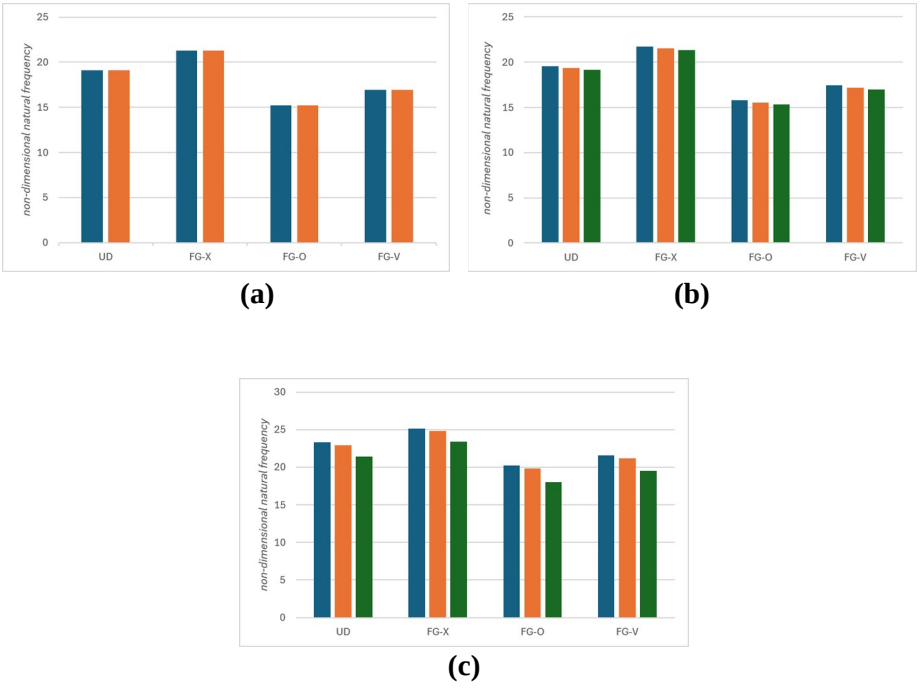


Figure 1. Variation of the non-dimensional natural frequency of simply supported FG-CNT plates as a function of the Kerr foundation parameter k_L for different CNT distribution patterns: (a): $k_w = k_s = 0$, (b): $k_w = k_s = 10$, and (c): $k_w = k_s = 100$.

The CNT distribution pattern has a significant influence on the natural frequencies. For the configuration with $k_w = k_s = k_L = 0$, the natural frequencies are 21.4206, 18.5166, 24.7339, and 16.4439 for the UD, FG-V, FG-X, and FG-O distributions, respectively. The FG-X configuration yields the highest natural frequency, whereas FG-O exhibits the lowest value. This trend is maintained across the investigated foundation configurations, demonstrating the importance of CNT distribution through the plate thickness in determining the effective mechanical properties and vibration response.

The influence of the Kerr elastic foundation is examined by varying k_w, k_s, and k_L. When $k_w = k_s = 10$ and $k_L = 0$, the natural frequency of the FG-X plate is 21.7206. Increasing k_w and k_s simultaneously to 100, while maintaining $k_L = 0$, increases the frequency to 25.1490. This increase indicates that the modified foundation

configuration enhances the dynamic stiffness of the plate-foundation system in this case.

However, the influence of k_L differs from that observed when k_w and k_s are varied. For $k_w = k_s$ =100\), increasing k_L from 0 to 10 and then to 100 reduces the FG-X natural frequency from 25.1490 to 24.8408 and 23.4035, respectively. Similar decreasing trends are observed for the UD, FG-V, and FG-O distributions. For example, the UD frequency decreases from 23.2887 to 21.3961, while the FG-V frequency decreases from 21.5489 to 19.4945 over the same range of k_L.

For $k_w = k_s$ =10, increasing k_L from 0 to 100 also produces a gradual reduction in natural frequencies. The FG-X frequency decreases from 21.7206 to 21.3632, whereas the UD frequency decreases from 19.5450 to 19.1475. These results confirm that the effect of the Kerr foundation depends on the combination of its stiffness parameters and cannot be characterized by a single trend for all parameters.

Overall, the results demonstrate that both CNT distribution and Kerr foundation characteristics govern the free vibration response of FG-CNT plates. The FG-X distribution consistently provides the highest natural frequencies, while FG-O produces the lowest values. The combined influence of the CNT distribution and foundation parameters should therefore be considered when predicting and optimizing the dynamic performance of composite plates.

5. Conclusion

This study investigated the free vibration behavior of simply supported functionally graded carbon nanotube-reinforced composite (FG-CNT) plates resting on a Kerr elastic foundation. Particular attention was paid to the effects of CNT distribution patterns and foundation stiffness parameters on the natural frequencies. The results lead to the following conclusions:

- The distribution pattern significantly affects the natural frequencies. Among the configurations considered, FG-X consistently yields the highest frequencies, followed by UD and FG-V, whereas FG-O produces the lowest values.

- The Kerr foundation modifies the natural frequencies according to the selected combination of k_w, k_s, and k_L. The observed trends indicate that the influence of an individual parameter depends on the values of the other parameters.
- For k_L =0, increasing k_w and k_s simultaneously from 10 to 100 increases the natural frequency of the FG-X plate.
- For k_w = k_s =100), increasing k_L from 0 to 100 reduces the FG-X natural frequency. This trend illustrates the sensitivity of the vibration response to the foundation configuration.

The CNT distribution and Kerr foundation parameters jointly determine the dynamic response of the plate. Their combined consideration is therefore essential for accurately predicting natural frequencies and selecting suitable configurations for vibration-sensitive applications.

The findings of this study provide useful guidance for the design and optimization of composite plate structures subjected to vibration. Understanding the combined influence of CNT distribution and elastic foundation properties can help engineers select suitable material configurations and support conditions to achieve the desired dynamic performance. These results may be relevant to structural and mechanical engineering applications in which lightweight composite components interact with elastic supports.

Future research could extend the present investigation to different boundary conditions, plate geometries, and loading environments, as well as to the effects of temperature, damping, and material uncertainties. Such developments could provide a more comprehensive understanding of the dynamic behavior of FG-CNT structures and support their potential application in advanced engineering systems.

References

1. Ansari, R., Torabi, J., & Hassani, R. (2019). A comprehensive study on the free vibration of arbitrary shaped thick functionally graded CNT-reinforced composite plates. Engineering Structures, 181, 653–669.

2. Amir, M., Tayeb, T. S. I., Kim, S.-W., & Choi, D. (2024). An analytical modelling of free vibration in porous FG-CNTRC plate resting on elastic foundations. Materials Today Communications, 40, 109925.
3. Mindlin's strain gradient theory for vibration analysis of FG-CNT-reinforced composite nanoplates resting on Kerr foundation in thermal environment. Journal of Vibration and Control.
4. Liu, L., and Z. Huang. 2014. "A Note on mori-tanaka's method." Acta Mech. Solida Sin. 27 (3): 234–244. https://doi.org/10.1016/S0894 -9166(14)60033-1.
5. Kiani, Y. 2016. "Free vibration of FG-CNT reinforced composite skew plates." Aerosp. Sci. Technol. 58: 178–188. https://doi.org/10.1016/j.ast.2016.08.018.
6. Shen, H. S. 2009. "Nonlinear bending of functionally graded carbon nanotube-reinforced composite plates in thermal environments." Compos. Struct. 91 (1): 9–19. https://doi.org/10.1016/j.compstruct.2009.04026.
7. Z. X. Lei, K. M. Liew, and J. L. Yu, "Free vibration analysis of functionally graded carbon nanotube-reinforced composite plates using the element-free kp-Ritz method in thermal environment," Compos. Struct., vol. 106, pp. 128–138, 2013.
8. H.-S. Shen and C.-L. Zhang, "Thermal buckling and postbuckling behavior of functionally graded carbon nanotube-reinforced composite plates," Mater. Des., vol. 31, no. 7, pp. 3403–3411, 2010.
9. Adhikari, B., & Singh, B. N. (2019). Dynamic response of FG-CNT composite plate resting on an elastic foundation based on higher-order shear deformation theory. Journal of Aerospace Engineering, 32(5), 04019061. DOI: 10.1061/(ASCE)AS.1943-5525.0001052

Interpretable Machine Learning for Predicting the Compressive Strength of FRP-Confined Concrete: A Rigorous Data-Curation and Bayesian Optimization Framework with TabPFN

Ali Benzaamia[1][0000-0002-2128-1396] and Mohamed Ghrici[2][0000-0001-9541-1862]

[1] National Earthquake Engineering Research Center, C.G.S., Rue Kaddour RAHIM, Hussein Dey, BP 252, Algiers, Algeria
[1] Geomaterials Laboratory, Hassiba Benbouali University of Chlef, P.O. Box 151, Chlef 02000, Algeria
benzaamia.ce@gmail.com

Abstract. Accurate prediction of the compressive strength of concrete confined with fiber-reinforced polymers (FRP) remains challenging because of nonlinear interactions among geometric, concrete, and FRP properties that analytical models capture only partially. This study presents a rigorous machine learning pipeline built on an experimental database of 1,157 raw observations of FRP-confined cylindrical specimens (CFRP, GFRP, AFRP, and HM/UHM-CFRP). After physical plausibility checks, slenderness filtering ($H/D \leq 3$), removal of exact duplicates, and averaging of replicate tests, 693 unique specimens were retained and split into training (554) and test (139) sets. Seven algorithms were compared: decision tree, random forest, AdaBoost, XGBoost, CatBoost, LightGBM, and TabPFN. All used a log-transformed target and seven inputs, including the physically derived confining pressure and ultimate FRP strain. Hyperparameters were tuned with Optuna Bayesian optimization and repeated 5-fold cross-validation. TabPFN, a pretrained transformer that needs no tuning, gave the best test performance ($R^2 = 0.958$, RMSE = 7.81 MPa, MAE = 4.83 MPa, A20 = 96.4%) with low overfitting ($\Delta R^2 = 0.039$), ahead of CatBoost ($R^2 = 0.936$). SHAP analysis identified lateral confining pressure as the dominant variable, followed by unconfined concrete strength, with trends consistent with confinement mechanics. The results are comparable to published models evaluated under less stringent protocols. A desktop application, PrédFRP, was developed so practicing engineers can use the model directly.

Keywords: FRP-confined concrete, Machine learning, Bayesian optimization, TabPFN, SHAP.

1 Introduction

Externally bonded fiber-reinforced polymer (FRP) jackets are widely used to strengthen reinforced concrete columns, because the passive lateral pressure they develop

increases both strength and ductility of the confined core [1]. Predicting the compressive strength of FRP-confined concrete, f'_{cc}, is a basic ingredient of design and assessment, but it is difficult: f'_{cc} depends on nonlinear and interacting effects of geometry, unconfined concrete strength, and jacket stiffness, strength and thickness, which design-oriented analytical models capture only partially [1,2].

Machine learning (ML) offers a data-driven alternative that imposes no functional form a priori, and has been applied to FRP-confined concrete [3–5]. Reported results can, however, be optimistic: databases compiled from many sources contain duplicates, replicate tests, slender specimens that fail by buckling and implausible records, and when such data are split randomly, replicates fall on both sides of the training/test boundary. Interpretability is also rarely checked against confinement mechanics, and few studies deliver a tool for practicing engineers.

This paper proposes a rigorous and reproducible pipeline with four contributions: (i) a curated database of 693 unique specimens obtained from 1,157 raw records; (ii) a fair comparison of seven algorithms on a stratified, strictly independent test set, with Bayesian hyperparameter optimization and repeated cross-validation; (iii) a SHAP interpretation of the best model checked against confinement mechanics; and (iv) PrédFRP, a desktop application that gives engineers direct access to the model.

2 Database and Data Curation

2.1 Source, Cleaning and Composition

The database compiles axial compression tests on FRP-confined circular concrete cylinders under monotonic loading. Its primary source is the dataset of Valasaki and Papakonstantinou [6], complemented with data from [4,5]. The raw file contains 1,157 observations on carbon (CFRP), glass (GFRP), aramid (AFRP) and high/ultra-high-modulus carbon (HM/UHM-CFRP) jackets. Each specimen is described by its diameter D and height H, the unconfined concrete strength f'_{c0}, the total FRP thickness t_f, the number of layers n, and the FRP modulus E_f and tensile strength f_f; the output is the confined strength f'_{cc}.

Physical plausibility. All dimensions and material properties had to be positive, the FRP volume ratio within (0, 1) and $n \geq 1$. Passive confinement cannot reduce the strength [1,2], so the nine specimens with $f'_{cc} < f'_{c0}$ were removed as probable transcription errors or premature debonding failures. **Slenderness.** Following ACI 318 [7], the 28 specimens with $H/D > 3$, which may fail by buckling instead of crushing, were discarded. **Duplicates.** Thirteen exact duplicates were deleted, and 263 groups of replicate tests (same configuration, different f'_{cc}) were replaced by their mean f'_{cc}, a common practice [3]. The final database contains 693 unique specimens (1,157 − 9 − 28 − 13 = 1,107 valid records, reduced by 414 after averaging).

Justification of the exclusions. Although the database shrinks by 40% (1,157 → 693), only 50 records (4.3%) were excluded; the other 414 (35.8%) were merged into the mean of their replicate group, so no configuration was lost and replicates cannot straddle the training/test split. The rules are objective and physical: $f'_{cc} < f'_{c0}$ contradicts passive confinement [1,2]; $H/D > 3$ may fail by buckling [7]; exact duplicates only

repeat an entry. They were applied before modeling and independently of model errors. The model therefore applies to cylinders with $H/D \leq 3$.

Fig. 1 shows the composition of the final database: CFRP dominates (n = 420, 60.6%), followed by GFRP (162), AFRP (74) and HM/UHM-CFRP (37), and normal-strength concrete (15–55 MPa) accounts for 78.5% of the specimens, an imbalance that mirrors the published experimental record.

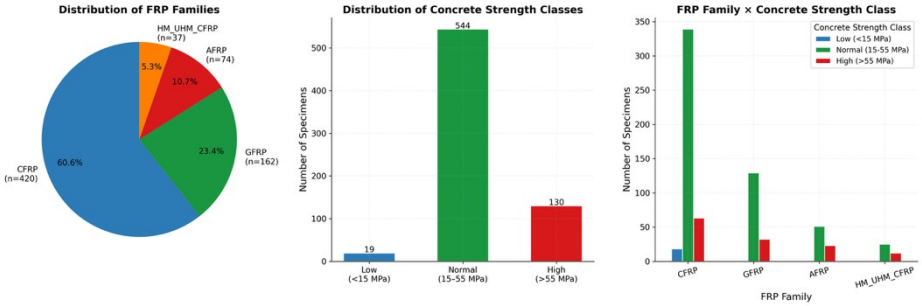


Fig. 1. Composition of the final database (N = 693): (left) FRP families; (center) concrete strength classes; (right) FRP family × concrete class.

2.2 Features and Target Variable

Two physically derived inputs complement the measured variables: the ultimate FRP strain and the lateral confining pressure, with f_f in MPa and E_f in GPa,

$$\varepsilon_{fu} = \frac{f_f}{E_f \cdot 10^3} \tag{1}$$

$$f_l = \frac{2t_f f_f}{D} \tag{2}$$

Eq. (2) is the radial pressure exerted by the jacket at hoop rupture, the basic quantity of FRP confinement models [1]. The input vector is [D, f'_{c0}, t_f, n, E_f, ε_{fu}, f_l], selected from correlation, mutual information (MI) and preliminary multi-model trials; all inputs are known before the test. Table 1 gives their statistics. The confining pressure is by far the best individual predictor of f'_{cc} (Pearson r = 0.811, Spearman ρ = 0.802, MI = 0.65), followed by f'_{c0} (r = 0.676). The target is $\ln(f'_{cc})$: the raw f'_{cc} is strongly right-skewed (skewness 1.59), and the logarithm reduces the skewness to 0.03; predictions are exponentiated back to MPa for evaluation.

Table 1. Descriptive statistics of the final database (N = 693 for every variable). SD: standard deviation; CV: coefficient of variation.

Variable	Unit	Min	Max	Mean	SD	CV (%)	Median
D	mm	47.0	600.0	151.5	55.8	36.8	150.0
f'_{c0}	MPa	6.2	169.7	43.4	23.5	54.2	38.2
t_f	mm	0.060	15.0	0.90	1.25	139.7	0.44
n	–	1	15	2.5	1.9	76.6	2.0
E_f	GPa	3.0	640.0	163.6	111.9	68.4	226.0
ε_{fu}	–	0.002	0.051	0.018	0.007	40.1	0.016
f_l	MPa	0.17	108.7	17.7	14.4	81.2	13.0
f'_{cc}	MPa	12.8	303.6	82.0	40.7	49.7	73.1
$\ln(f'_{cc})$	–	2.549	5.716	4.299	0.464	10.8	4.292

2.3 Stratified Train/Test Split

The data were split into a training set (80%, 554 specimens) and a test set (20%, 139 specimens) with a stratification key combining the FRP family, the concrete strength class and the quantile of f'_{cc} (eight bins), so that every sub-region of the target space appears in both sets (seed 47). Kolmogorov–Smirnov tests found no significant difference between the two sets ($p > 0.05$) for all variables except D ($p = 0.040$), which is concentrated around 150 mm in both. The test set was never used for training, tuning or model selection.

3 Modeling Methodology

3.1 Algorithms and Hyperparameter Optimization

Seven regressors were compared: decision tree (DT), random forest (RF), AdaBoost, XGBoost [8], LightGBM [9], CatBoost [10] and TabPFN [11]. TabPFN is a transformer pretrained by meta-learning on synthetic tabular data; it predicts by in-context learning, needs no tuning, and was used in its default configuration. All models use the seven inputs above and the log-transformed target.

The six other algorithms were tuned with Optuna [12]. Unlike grid search and unguided random search [13], Bayesian optimization builds a surrogate of the objective from previous trials and selects the next one with an acquisition function, reaching good configurations with far fewer evaluations [14]. We used the multivariate Tree-structured Parzen Estimator [15] after 20 random start-up trials, with the RMSE of repeated 5-fold cross-validation (three repeats) on the training set as objective. A median pruner stops unpromising trials, and each study stops after 200 consecutive trials without improvement (budget: 500 trials). The optimal settings are not reproduced for lack of space.

3.2 Evaluation Metrics

Seven metrics are used: the coefficient of determination R^{2}, the Kling–Gupta efficiency (KGE), the confidence index (CI, Pearson correlation times Willmott's index of agreement), the A20 index (fraction of predictions within ±20% of the measured value), and the MAE, RMSE and MAPE. Results are computed in MPa after exponentiation.

4 Results and Discussion

4.1 Predictive Performance

Table 2 reports the test metrics of the seven models, with the training R^{2} and the generalization gap $\Delta R^{2} = R^{2}_{\text{train}} - R^{2}_{\text{test}}$ as an indicator of overfitting. DT and RF give the weakest results (R^{2} = 0.796 and 0.836, A20 below 80%). AdaBoost reaches the same test R^{2} as RF but overfits strongly (training R^{2} = 0.975, ΔR^{2} = 0.139). LightGBM, XGBoost and CatBoost improve steadily (R^{2} = 0.903, 0.914 and 0.936). TabPFN is the best model on every metric except KGE, where CatBoost is marginally higher (0.966 versus 0.959): R^{2} = 0.958, RMSE = 7.81 MPa, MAE = 4.83 MPa, MAPE = 5.9% and A20 = 96.4%. Its gap (ΔR^{2} = 0.039) is the smallest among the accurate models; only RF, which is far less accurate, has a smaller one (0.024). This is consistent with the regularizing effect of the meta-training of TabPFN [11]. The ranking is TabPFN > CatBoost > XGBoost > LightGBM > AdaBoost ≈ RF > DT.

Table 2. Performance of the seven models on the test set (N = 139). Training R² and the gap ΔR² are given for overfitting assessment. Best values in bold.

Model	R^{2} train	R^{2} test	ΔR^{2}	KGE	CI	MAE (MPa)	RMSE (MPa)	MAPE (%)	A20 (%)
DT	0.838	0.796	0.042	0.898	0.852	11.62	17.25	15.2	76.3
RF	0.860	0.836	0.024	0.864	0.871	10.58	15.49	13.8	79.1
AdaBoost	0.975	0.836	0.139	0.893	0.873	9.50	15.46	11.9	84.2
XGBoost	0.996	0.914	0.082	0.956	0.938	6.28	11.17	7.7	95.0
CatBoost	0.996	0.936	0.060	0.966	0.952	5.97	9.69	7.3	95.0
LightGBM	0.979	0.903	0.076	0.949	0.932	7.21	11.87	8.8	92.1
TabPFN	**0.997**	**0.958**	**0.039**	0.959	**0.970**	**4.83**	**7.81**	**5.9**	**96.4**

Fig. 2 confirms these results for TabPFN: the points follow the 1:1 line on both sets, including strengths above 200 MPa that usually produce large errors, very few test points leave the ±20% band, and the training and test residual distributions are almost identical. All models have larger errors for strong confinement (f_l > 30 MPa) and high-performance concrete (f'_{c0} > 50 MPa), two regimes under-represented in the database.

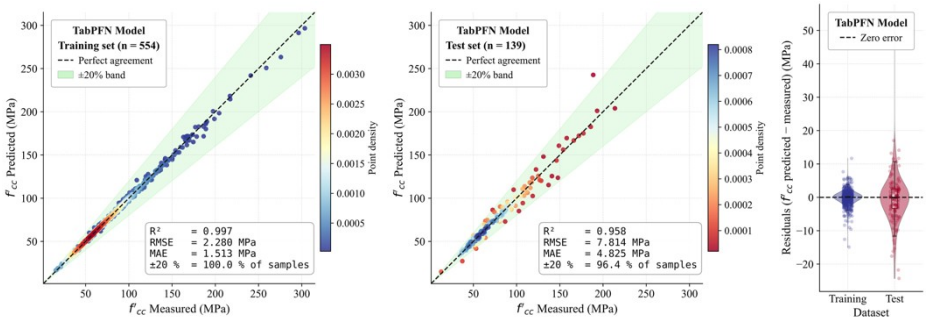


Fig. 2. Measured versus predicted f'_{cc} for TabPFN: (a) training and (b) test scatter plots, and (c) distribution of the centered residuals.

4.2 Interpretation with SHAP

SHAP [16] was applied to TabPFN. Fig. 3(a) ranks the mean absolute SHAP values: f_l (0.27), f'_{c0} (0.18), E_f (0.09), ε_{fu} (0.07), t_f and n (0.05 each) and D (0.03). The dominance of the confining pressure, followed by the unconfined strength, agrees with confinement mechanics [1,2]. The beeswarm plot (Fig. 3(b)) shows that high f_l values strongly increase the predicted strength, while f'_{c0} acts positively but with more dispersion, revealing interactions.

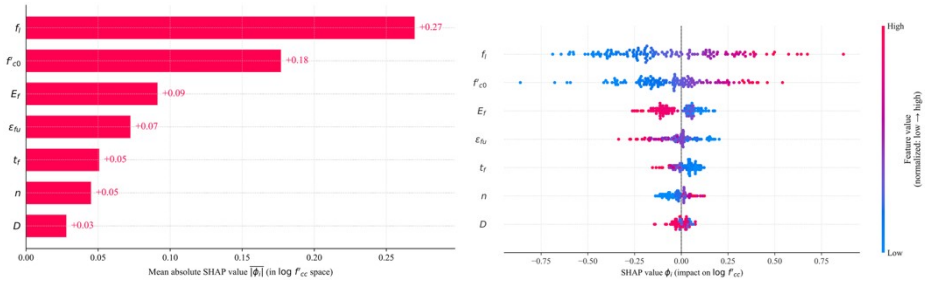


Fig. 3. Global SHAP analysis of TabPFN: (a) mean absolute SHAP values; (b) beeswarm plot of the SHAP values of each input.

SHAP dependence plots show monotone increasing relations for f_l (Spearman correlation between feature and SHAP value ρ_s = 0.994, saturating shape), f'_{c0} (ρ_s = 0.984, reflecting the lower relative efficiency of confinement in strong concrete) and n (0.933), and a slightly decreasing effect of D (−0.636), as expected since a larger diameter lowers f_l for the same jacket (Eq. (2)). TabPFN has thus learned relations consistent with the physics of confinement.

4.3 Comparison with the Literature

Zeng et al. [3] used the same source database [6] under two protocols: TRTR (train and test on real data, 70/30 split of the raw database) and TSTR (train on CTGAN-generated synthetic data, test on real data). Their best models reach test R^2 = 0.963 (XGBoost,

TRTR; RMSE = 7.50 MPa) and 0.972 (GBDTR, TSTR; RMSE = 6.30 MPa), against R^{2}_{2} = 0.958 and RMSE = 7.81 MPa for TabPFN. These values are not directly comparable: according to the protocol described in [3], no de-duplication was applied, so replicates may lie on both sides of a random split and inflate test scores, and in the TSTR protocol the CTGAN was trained on all real data, test set included. Here, the 139 test specimens are independent of every training and data-generation step, and TabPFN still reaches comparable accuracy with a smaller training set (554 versus 778 specimens), without tuning or synthetic augmentation.

4.4 The PrédFRP Application

TabPFN was embedded in PrédFRP, a desktop application written in Python with PyQt5. The engineer enters the seven inputs and obtains f'_{cc} instantly, together with a SHAP waterfall plot that explains the prediction (Fig. 4). To discourage extrapolation, each field shows the range of the training database and a green or red indicator tells whether the entered value lies inside it. For the specimen of Fig. 4 (D = 150 mm, f'_{c0} = 35 MPa, t_f = 0.26 mm, n = 2, E_f = 211 GPa, ε_{fu} = 0.015, f_l = 10.97 MPa), the model returns f'_{cc} = 70.77 MPa. The model should be applied within the validated ranges (D = 75–600 mm, f'_{c0} = 16–120 MPa, f_l = 0.5–110 MPa).

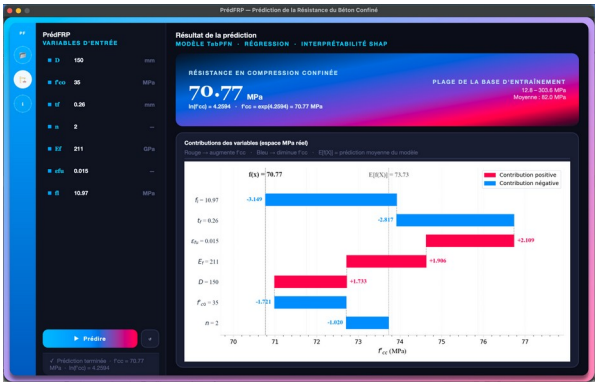


Fig. 4. Main window of PrédFRP showing a complete prediction: predicted f'_{cc} = 70.77 MPa and SHAP waterfall plot of the entered specimen.

5 Conclusions

Seven ML algorithms were compared on a curated database of 693 FRP-confined concrete cylinders with a stratified independent test set. TabPFN gave the best test performance (R^{2}_{2} = 0.958, RMSE = 7.81 MPa, MAE = 4.83 MPa, A20 = 96.4%) with a small generalization gap (ΔR^{2}_{2} = 0.039) and no tuning, ahead of CatBoost (R^{2}_{2} = 0.936) and XGBoost (R^{2}_{2} = 0.914). SHAP analysis identified the confining pressure as the dominant variable, followed by the unconfined concrete strength, with trends consistent with confinement mechanics. The main limitation is the larger error for strong confinement (f_l > 30 MPa) and high-performance concrete (f'_{c0} > 50 MPa), which calls

for more experimental data in these regimes. PrédFRP makes the model directly available to practicing engineers.

References

1. Lam, L., Teng, J.G.: Design-oriented stress–strain model for FRP-confined concrete. Constr. Build. Mater. 17(6–7), 471–489 (2003).
2. Teng, J.G., Huang, Y.L., Lam, L., Ye, L.P.: Theoretical model for fiber-reinforced polymer-confined concrete. J. Compos. Constr. 11(2), 201–210 (2007).
3. Zeng, S., Wang, X., Hua, L., Altayeb, M., Wu, Z., Niu, F.: Prediction of compressive strength of FRP-confined concrete using machine learning: a novel synthetic data driven framework. J. Build. Eng. 94, 109918 (2024).
4. Jamali, F., Mousavi, S.R., Peyma, A.B., et al.: Prediction of compressive strength of fiber-reinforced polymers-confined cylindrical concrete using artificial intelligence methods. J. Reinf. Plast. Compos. 41(17–18), 679–704 (2022).
5. Kumar, P., Arora, H.C., Bahrami, A., et al.: Development of a reliable machine learning model to predict compressive strength of FRP-confined concrete cylinders. Buildings 13(4), 931 (2023).
6. Valasaki, M.K., Papakonstantinou, C.G.: Fiber reinforced polymer (FRP) confined circular concrete columns: an experimental overview. Buildings 13(5), 1248 (2023).
7. ACI Committee 318: Building Code Requirements for Structural Concrete (ACI 318-19) and Commentary. American Concrete Institute, Farmington Hills (2019).
8. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. ACM, New York (2016).
9. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y.: LightGBM: a highly efficient gradient boosting decision tree. In: Advances in Neural Information Processing Systems 30, pp. 3146–3154 (2017).
10. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A.: CatBoost: unbiased boosting with categorical features. In: Advances in Neural Information Processing Systems 31, pp. 6638–6648 (2018).
11. Hollmann, N., Müller, S., Purucker, L., Krishnakumar, A., Körfer, M., Hoo, S.B., Schirrmeister, R.T., Hutter, F.: Accurate predictions on small data with a tabular foundation model. Nature 637, 319–326 (2025).
12. Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.: Optuna: a next-generation hyperparameter optimization framework. In: 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 2623–2631. ACM, New York (2019).
13. Bergstra, J., Bengio, Y.: Random search for hyper-parameter optimization. J. Mach. Learn. Res. 13, 281–305 (2012).
14. Shahriari, B., Swersky, K., Wang, Z., Adams, R.P., de Freitas, N.: Taking the human out of the loop: a review of Bayesian optimization. Proc. IEEE 104(1), 148–175 (2016).
15. Bergstra, J., Yamins, D., Cox, D.D.: Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures. In: 30th International Conference on Machine Learning, pp. 115–123 (2013).
16. Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30, pp. 4765–4774 (2017).

Numerical Investigation of the Behavior of Cold-Formed Steel Profiles with Web Openings under Fire Exposure

Taibi Sarra[1], Kada Abdelhak[1], Hassoune Mohammed[1], Mus'ab Salameh[1]

[1]Laboratory Fire Safety Engineering of Constructions and Protection of their Environment (LISICPE), Faculty of Civil Engineering and Architecture, Hassiba Benbouali University of Chlef, Ouled Fares, N19 Chlef, Algeria
s.taibi@univ-chlef.dz

Abstract. Cold-formed steel (CFS) profiles are widely used in lightweight structural systems because of their high strength-to-weight ratio, ease of fabrication and versatility. However, their thin-walled nature and high thermal conductivity make them vulnerable to elevated temperatures and fire exposure. Web openings, which are commonly introduced to accommodate services and reduce weight, may affect the thermal and structural behaviour of these members. This study numerically investigates the thermal response of single and back-to-back cold-formed steel lipped-channel profiles with web openings under the ISO 834 standard fire. Thermal finite element models are developed using ANSYS Mechanical APDL, with temperature-dependent thermal properties of steel defined in accordance with Eurocode 3, Part 1-2. The investigated configurations include profiles with different opening arrangements and profile lengths. The results show that the presence of web openings has a negligible influence on the overall temperature–time response under the investigated conditions. After 60 min of fire exposure, the temperature of the single CFS profile reaches approximately 944 °C. For the back-to-back configuration, the maximum temperatures are approximately 944 °C at the web and 844 °C at the flanges. The temperature–time curves for profiles with openings remain close to those of the corresponding reference profiles. These findings indicate that the investigated web openings have a limited effect on the thermal response of the studied CFS members.

Keywords: Cold Formed steel sections with web openings, Fire resistance, ANSYS, ISO834 Fire.

1 Introduction

In In recent years, cold-formed steel (CFS) sections have attracted considerable attention in modern construction because of their low self-weight, high strength-to-weight ratio and ease of fabrication. However, their thin-walled nature makes them vulnerable to strength reduction at elevated temperatures, and their fire performance is therefore an important consideration in building safety. Several studies have investigated the thermal behaviour of cold-formed steel members exposed to fire, including individual beams and columns [1, 2] and stud walls [3]. Rajeendra Godakandage et al. [4] studied the effects of cavity fires in modular buildings. Other

researchers have examined the structural response and buckling behaviour of CFS members at elevated temperatures. Varun Sabu Sam et al. [5] conducted an experimental and numerical investigation into the influence of elevated temperature on the buckling capacity of mild-steel-based CFS column sections. Kazuya Mitsui and Kikuo Ikarashi [6] analysed the elastic critical local buckling of cold-formed lipped-channel sections using the energy method. Mohammed Hassoune et al. [7] investigated the performance of CFS stud walls under compression loading and one-sided and double-sided fire exposure. Yecheng Dai et al. [8] studied the structural performance of face-to-face built-up CFS channel sections subjected to axial compression at high temperatures. Abdelhak Kada and Belkacem Lamri [9] investigated the mechanical behaviour of unrestrained solid steel I-beams exposed to fire. Carla de Amorim Lana Dib et al. [10] analysed the distortional failure of CFS hat-section beams under non-uniform bending. Other studies have focused on fire-protection methods, including fire-resistant coatings and insulation. Mohammed Hassoune et al. [11] numerically analysed the thermal performance of light-steel-framed walls protected with plasterboard or magnesium oxide board and cavity insulation. Weihua Wang et al. [12] experimentally examined the fire resistance of insulated steel beams protected with fire-protection coatings and fibre-cement boards. Varun Sabu Sam et al. [13] investigated the buckling performance of back-to-back CFS channel columns with protective coatings. Pegah Aghabozorgi et al. [14] conducted a detailed investigation of protected steel columns under standard fire exposure. Xinxin Zhang et al. [15] examined the fire resistance of prefabricated steel tubular columns with membrane protection. Although these studies address the fire performance of steel members and protection systems, further investigation is needed into the thermal response of CFS profiles containing web openings, particularly for single and back-to-back configurations. Therefore, the aim of this study is to numerically investigate, using ANSYS APDL, the thermal behaviour of single and back-to-back lipped-channel CFS profiles with different web-opening arrangements under the ISO 834 standard fire. The study compares their temperature–time responses and evaluates the influence of the openings on temperature development.

2 Material properties of CFS

The thermal properties of steel—specific heat, thermal conductivity and density—vary with temperature, as shown in Figure 1. These properties are evaluated in accordance with Eurocode 3, Part 1-2, and are used as input data for the transient thermal analyses.

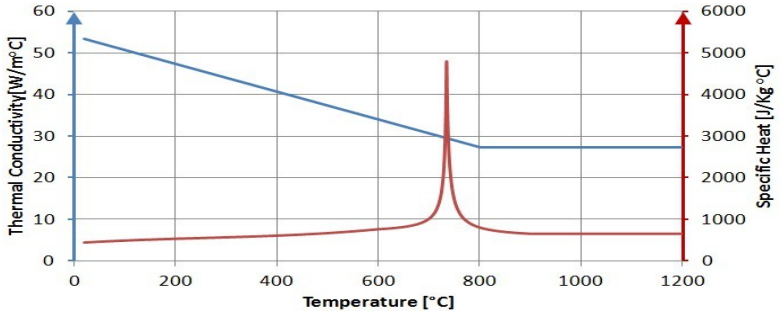


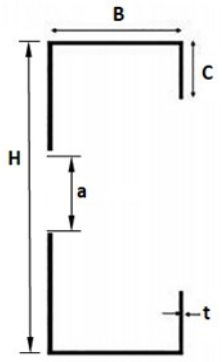
Figure 1. *Thermal properties of CFS steel.*

3 Finite element modeling and validation

3.1 Case of stud of CFS profiles

The present study investigates cold-formed steel (CFS) beams with lipped C-sections, considering variations in profile length and in the size and configuration of the web openings, as illustrated in Figure 2. The geometric details of the profiles considered in this study are presented in Table1.

Table 1. Dimensions of profiles

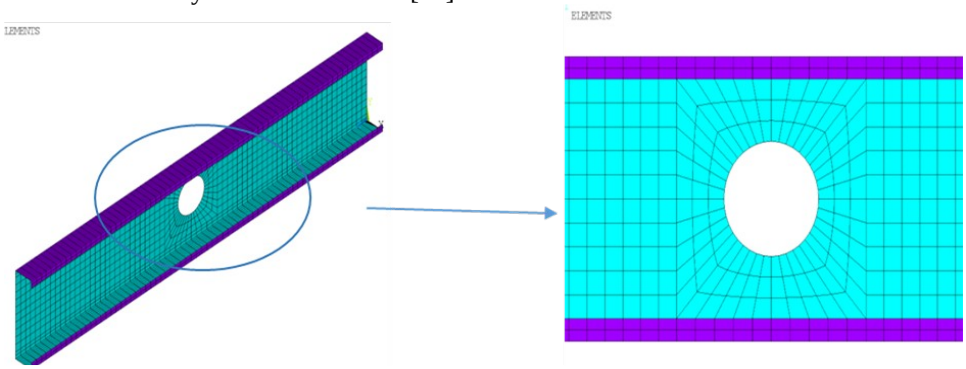


Model	Type of profile	H (mm)	B (mm)	t (mm)	c (mm)	a (mm)	L (mm)
1	Simple	300	75	3	25	60* / 120* / 2×120**	3300
2	Back-to-back	300	150	3	25	120 / 2×120	3300
3	Simple	300	75	3	25	120 / 2×120	1600

Note: (*) indicates one web opening, while (**) indicates two web openings

3.2 Element type, mesh size and boundary coditions

Thermal The thermal simulations are conducted using ANSYS APDL. The lipped-channel sections are exposed to the ISO 834 standard fire. Thermal finite element models are developed using the SHELL131 element, which is suitable for the thermal analysis of thin-walled CFS profiles. A mesh size of 20 × 20 mm is adopted for models without web openings, while a finer mesh of 12.5 × 12.5 mm is applied at the beam ends to improve the accuracy of the results in these critical regions. For models incorporating web openings, a specific meshing strategy is used to ensure a structured and regular mesh around the openings (Figure 2). The thermal boundary conditions represent exposure to the ISO 834 fire. Convection coefficients of 25 W/(m²·K) for the exposed face and 9 W/(m²·K) for the unexposed face are used, including the radiation effect, as recommended by EN 1991-1-2 [16]. An



emissivity coefficient of 1 is applied to the exposed face.

Figure 2. *Mesh size used models. CFS steel.*

4 Results and discussions

4.1 Thermal response

The temperature–time curves show that the presence of web openings has a negligible influence on the thermal response of both single and back-to-back cold-formed steel (CFS) profiles exposed to the ISO 834 standard fire. For single profiles, the maximum temperature reaches approximately 944 °C after 60 min of exposure, with only a slight temperature difference between the web and the flanges (Figure 3). Similarly, for back-to-back profiles, the maximum temperatures reach approximately 944 °C at the web and 844 °C at the flanges (Figure 4). In both configurations, the temperature–time curves remain close to those of the reference profiles, regardless of whether a central opening or two lateral openings are present. These results indicate that, under the investigated conditions, web openings have a limited effect on the overall thermal behaviour of single and back-to-back CFS profiles.

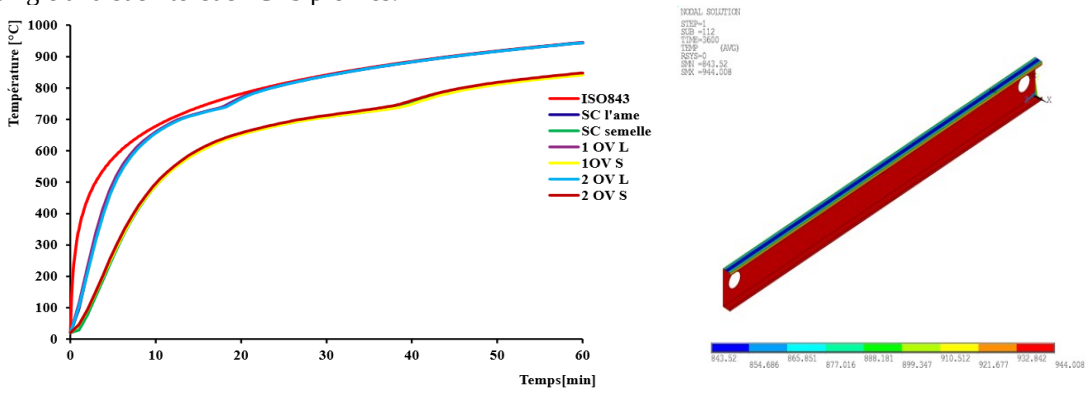


Figure 3. *Temperature time curve of single CFS.*

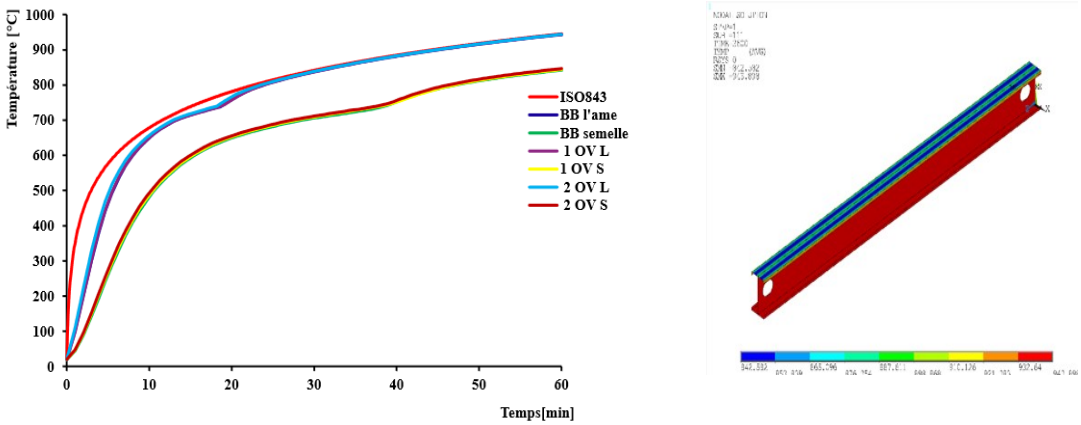


Figure 4. *Temperature time curve of back-to-back CFS.*

5 Conclusion

The thermal analysis indicates that web openings have a negligible influence on the overall thermal response of single and back-to-back cold-formed steel (CFS) profiles exposed to the ISO 834 standard fire. The temperature–time curves remain similar across the investigated configurations, with only minor variations in temperature evolution. After 60 min of exposure, the temperature of the single profile reaches approximately 944 °C. For the back-to-back profiles, the maximum temperatures are approximately 944 °C at the web and 844 °C at the flanges. These findings suggest that the investigated openings have a limited effect on the thermal response of the studied members. However, thermal behaviour alone is insufficient to assess the structural performance of perforated profiles. Further work should therefore investigate their mechanical behaviour, particularly the influence of web openings on structural stability, load-bearing capacity and buckling behaviour at elevated temperatures.

References

1. Al-Kaseasbeh, Q., *Analysis of hydrocarbon fire-exposed cold-formed steel columns.* Results in Engineering, 2023. **20**: p. 101400.
2. Prabowo, A., Y. Huang, and B. Young, *Cold-formed stainless steel beams with single web hole at elevated temperatures.* Thin-Walled Structures, 2024. **195**: p. 111321.
3. Vy, S.T., A. Ariyanayagam, and M. Mahendran, *Behaviour and design of CFS stud walls under both sides fire exposure.* Thin-Walled Structures, 2024. **197**: p. 111619.
4. Godakandage, R., et al., *Structural and Fire Performance of Cold-Formed Steel Columns Subjected to Cavity Fire in Modular Buildings.* Fire, 2025. **8**(5): p. 190.
5. Sabu Sam, V., et al., *Influence of elevated temperature on buckling capacity of mild steel-based cold-formed steel column sections– experimental investigation and finite element modelling.* Journal of Structural Fire Engineering, 2023. **15**(3): p. 314-337.
6. Mitsui, K. and K. Ikarashi, *Elastic critical local buckling load and behavior of local buckling for cold-formed lipped channel steel member under compression.* Japan Architectural Review, 2024. **7**(1): p. e12432.
7. Hassoune, M., et al., *Performance of single and built-up I-shaped cold formed steel stud under double sided walls fire exposure.* Engineering Structures, 2025. **335**: p. 120392.
8. Dai, Y., et al., *Structural performance of cold-formed steel face-to-face built-up channel sections under axial compression at high temperatures through finite element modelling.* Buildings, 2023. **13**(2): p. 305.
9. Kada, A. and B. Lamri, *Numerical analysis of non-restrained long-span steel beams at high temperatures due to fire.* Asian Journal of Civil Engineering, 2019. **20**(2): p. 261-267.
10. de Amorim Lana Dib, C., G.H. dos Santos Ramos, and G.S. Vieira, *Numerical analysis on distortional failure of cold-formed steel hat-section beams under non-uniform bending.* International Journal of Steel Structures, 2023. **23**(5): p. 1191-1201.
11. Hassoune, M., et al., *Numerical thermal performance analysis of light steel insulated walls under fire.* Građevinski materijali i konstrukcije= Building Materials and Structures, 2023. **66**(1): p. 43-54.
12. Wang, W., et al., *Fire Test on Insulated Steel Beams with Fire-Protection Coating and Fiber Cement Board.* Buildings, 2025. **15**(12): p. 2121.

13. Sam, V.S., et al., *Effect of Protective Coatings on Post-Fire Performance and Behavior of Mild Steel-Based Cold-Formed Steel Back-to-Back Channel Columns with Bolted Connections.* Fire, 2025. **8**(3): p. 107.
14. Aghabozorgi, P., L. Laím, and A. Santiago, *Numerical and Design Analysis of Protected Steel Columns in Standard Fire.* Fire Technology, 2025: p. 1-43.
15. Zhang, X., X.Y. Zheng, and W. Li, *Fire Resistance of Prefabricated Steel Tubular Columns with Membrane Protections.* Buildings, 2025. **15**(10): p. 1730.
16. European Committee for Standardization (CEN), Eurocode 3: Design of Steel Structures – Part 1-2: General Rules – Structural Fire Design, EN 1993-1-2:2005, Brussels, Belgium, 2005.

Prediction of the Chloride Permeability of Fly Ash-Containing Self-Compacting Concrete Using an Optimized XGBoost Model

Kechroud Fatima[1][0000-0001-7427-693X], Benzaamia Ali[2][0000-0002-2128-1396], and Ghrici Mohamed[1][0000-0001-9541-1862]
[1] Geomaterials Laboratory, Hassiba Benbouali University of Chlef, 1, Chlef 02000, Algeria
[2] National Earthquake Engineering Research Center, C.G.S., Rue Kaddour RAHIM, Hussein Dey, BP 252, Algiers, Algeria
Kechroude1992@gmail.com

Abstract. Self-compacting concrete (SCC) is increasingly used in construction, but its resistance to chloride ingress remains a major durability concern. This study develops an Extreme Gradient Boosting (XGBoost) model to predict the chloride permeability of fly ash-containing SCC, as measured by the Rapid Chloride Permeability Test (RCPT). The model was trained and tested on a database of 126 observations compiled from the published literature (94 for training and 32 for testing), with six input variables: curing age, binder content (B), fly ash content (FA), water-to-binder ratio (W/B), superplasticizer dosage (SP) and compressive strength (f_{cj}). Hyperparameters were tuned with a Bayesian (TPE) search over 1000 Optuna trials and five-fold cross-validation. The optimized model achieved coefficients of determination (R^2) of 0.995 for the training set and 0.915 for the test set, RMSE values of 54.9 C and 308.0 C, and a_{30}-index values of 1.000 and 0.875, respectively. SHAP, gain-based importance and partial dependence analyses showed that curing age is by far the most influential variable (about 40% of the total mean |SHAP| value), followed by FA, B and W/B. Chloride permeability decreased sharply with age up to about 90 days and with increasing fly ash content, whereas the influence of W/B was weaker and non-monotonic within the compiled database. These findings highlight the potential of XGBoost as an efficient tool for supporting the durability assessment of fly ash-based SCC.

Keywords: Self-compacting concrete, Chloride permeability, RCPT, Machine learning, XGBoost.

1 Introduction

Self-compacting concrete (SCC) flows under its own weight and fills the formwork without vibration [1]. It generally contains a higher binder content and a lower water-

to-binder ratio than conventional concrete and is often combined with fly ash. Chloride-induced corrosion of the reinforcement is a leading cause of deterioration of reinforced concrete, so the chloride resistance of such mixtures is a primary durability concern.

The Rapid Chloride Permeability Test (RCPT) measures the electrical charge (in coulombs) passed through a specimen over six hours and is widely used to indicate resistance to chloride penetration [2]. The charge depends on interacting parameters such as age, binder content, fly ash replacement [3], water-to-binder ratio and admixture dosage, so simple empirical equations have limited generality.

Machine learning (ML) offers a data-driven alternative. Neural networks have been used to predict the RCPT results of SCC [4], and XGBoost combined with metaheuristic optimizers has recently been applied to SCC with fly ash and silica fume [5]. Gradient-boosted trees [6], and in particular Extreme Gradient Boosting (XGBoost) [7], are well suited to small tabular datasets because they represent nonlinear responses and interactions without assuming a functional form. This study (i) characterizes a literature database on the RCPT of fly ash-containing SCC, (ii) develops an XGBoost model tuned by Bayesian optimization, (iii) evaluates it on unseen data and (iv) interprets it with SHAP, gain importance and partial dependence, with attention to W/B and FA.

2 Experimental Database

2.1 Description of the Database

The database was compiled from the published literature and contains 126 RCPT observations for fly ash-containing SCC. The inputs are the curing age (Age, days), binder content (B, kg/m^3), fly ash content (FA, %), water-to-binder ratio (W/B), superplasticizer dosage (SP, kg/m^3) and compressive strength (f_{cj}, MPa); the output is the RCPT charge (C). Descriptive statistics are given in Table 1.

Table 1. Descriptive statistics of the database (n = 126).

Variable	Min.	Max.	Mean	Median	Std.	CV (%)
Age (days)	28	365	98.4	28	111.4	113.1
B (kg/m^3)	350	610	497.0	500	59.0	11.9
FA (%)	0	70	42.1	44	16.3	38.7
W/B	0.24	0.45	0.361	0.35	0.045	12.5
SP (kg/m^3)	0.40	11.40	5.45	6.75	3.26	59.8
f_{cj} (MPa)	31	84	53.1	52.1	11.3	21.2
RCPT (C)	230	4600	1256	975	882	70.2

Std.: standard deviation; CV: coefficient of variation.

The database covers wide ranges of mixture proportions (Table 1; distributions in Fig. 1). RCPT values span 230–4600 C, with a mean of 1256 C and a median of 975 C; according to ASTM C1202 [2] the mean lies in the “low” class and the upper tail

reaches the “moderate” and “high” classes. Curing age is discrete and right-skewed: data are concentrated at six ages (28, 56, 90, 91, 180 and 365 days) and 52 % of the observations (65 of 126) are at 28 days. RCPT is positively skewed (skewness = 1.35); since tree-based models are invariant to monotonic transformations, no variable was transformed. Observations flagged by the 1.5 × IQR criterion (16 for age, all at 365 days; one for W/B; six for RCPT above about 3100 C) correspond to genuine mixtures and were all retained.

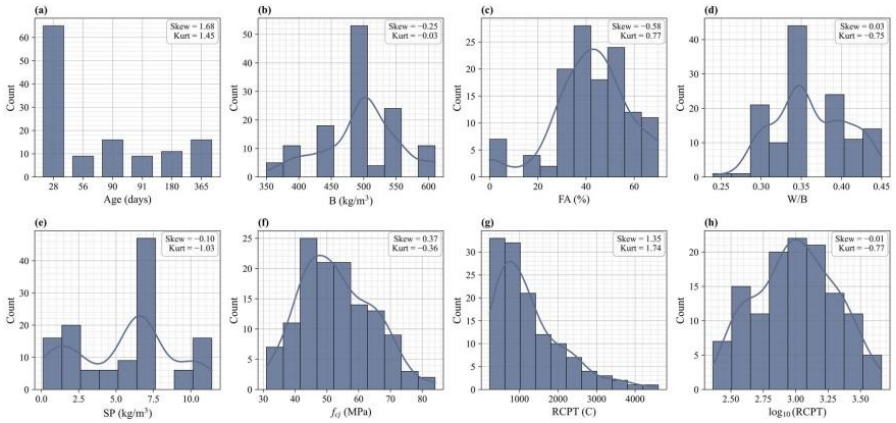


Fig. 1. Histograms with kernel density estimates of (a) age, (b) B, (c) FA, (d) W/B, (e) SP, (f) f_{cj}, (g) RCPT and (h) $\log_{10}$(RCPT).

2.2 Data Partition and Correlation Analysis

The database was divided into a training set of 94 observations (74.6 %) and an independent test set of 32 observations (25.4 %), which remained unseen until the final evaluation. Means and standard deviations are similar in both subsets and Kolmogorov–Smirnov tests do not reject equality of distributions for any variable (all $p > 0.05$; smallest $p = 0.12$, for f_{cj}). However, the test set contains the highest (4600 C versus 4000 C) and lowest RCPT values (230 C versus 300 C) and a lower minimum W/B (0.24 versus 0.27), so part of it lies slightly outside the training domain, which is a demanding test for a tree-based model.

Linear (Pearson) and rank (Spearman) correlations are shown in Fig. 2. Curing age is the variable most strongly correlated with RCPT ($r = -0.50$; $\rho = -0.72$), and the larger rank correlation indicates a monotonic but markedly nonlinear relationship. SP ($r = -0.36$) and FA ($r = -0.29$) show weak-to-moderate negative correlations, whereas B, W/B and f_{cj} are practically uncorrelated with RCPT ($|r| \leq 0.06$). The strongest correlations among inputs are B–SP ($r = 0.66$), B–W/B ($r = -0.51$) and FA–f_{cj} ($r = -0.51$); the variance inflation factors (1.90–3.87) are below 5, so multicollinearity is not problematic, although the effects extracted in Section 4 must be interpreted with caution.

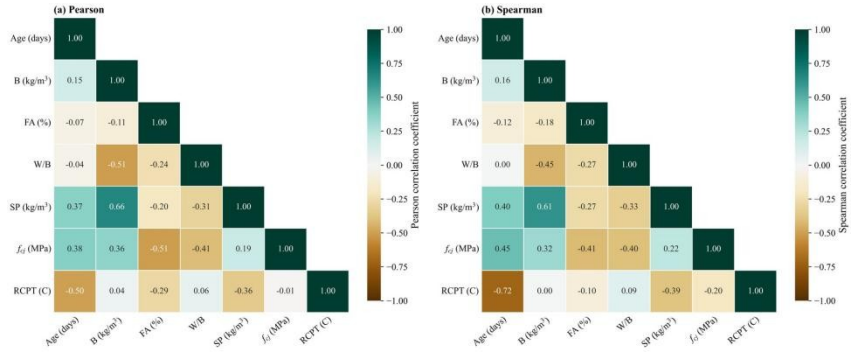


Fig. 2. Correlation matrices of the variables: (a) Pearson and (b) Spearman coefficients.

3 Methodology

3.1 Extreme Gradient Boosting

In XGBoost [7] the prediction for a sample with input vector $\mathbf{x}_i$ is the sum of the outputs of K regression trees:

$$\hat{y}_i = \Sigma_{k=1}^{K} f_k(\mathbf{x}_i), \;\; f_k \in F \tag{1}$$

where F is the space of regression trees. The trees are added sequentially, each one reducing the residual error of the current ensemble by minimizing the regularized objective

$$L = \Sigma_i \, l(y_i, \hat{y}_i) + \Sigma_k \, \Omega(f_k), \;\; \Omega(f) = \gamma T + \tfrac{1}{2}\lambda\|\mathbf{w}\|^2 + \alpha\|\mathbf{w}\|_1 \tag{2}$$

in which l is the squared-error loss, T the number of leaves, $\mathbf{w}$ the vector of leaf weights, γ the minimum loss reduction required to split a node, and λ and α the L2 and L1 regularization coefficients.

3.2 Hyperparameter Optimization and Metrics

The hyperparameters were tuned with Optuna [8] using a Tree-structured Parzen Estimator (TPE) sampler [9] and a median pruner. The objective was the RMSE averaged over five-fold cross-validation of the training set only, so that the test set played no role in the tuning; the number of trees was set by early stopping (maximum 2000, patience 50).

Of 1000 trials, 904 were completed and 96 pruned; the lowest cross-validated RMSE was 355.2 C. All best-ranked trials used the loss-guided policy with comparable settings. The configuration of trial 981 (cross-validated RMSE = 371.2 C, 217 trees) was adopted (learning rate 0.209, 31 leaves, minimum child weight 3.59, row subsampling 0.744), trained on the full training set and applied unchanged to the test set.

The model was assessed with the coefficient of determination (R^2), mean absolute error (MAE), RMSE, mean absolute percentage error (MAPE), Kling–Gupta efficiency (KGE), confidence index (CI) and the a_n-index, defined as $a_n = m_n/N$, where m_n is the number of samples for which the ratio of experimental to predicted value lies between $1 - n/100$ and $1 + n/100$ (a_{10}, a_{20} and a_{30} were computed). Interpretation relied on the mean absolute SHapley Additive exPlanations (SHAP) value [10] computed on all 126 observations, the gain-based importance,within a ±30% error margin(a_{30}) and individual conditional expectation (ICE) and partial dependence (PD) curves.

4 Results and Discussion

4.1 Predictive Performance

The metrics are reported in Table 2 and the predictions are plotted in Fig. 3a. The model reproduces the training data almost perfectly (R^2 = 0.995, RMSE = 54.9 C, MAPE = 4.0 %). On the unseen test set it retains a high accuracy: R^2 = 0.915, CI = 0.943 and KGE = 0.850, with RMSE = 308.0 C, MAE = 224.5 C and MAPE = 19.3 %. The a_{30}-index of the test set is 0.875, i.e. 28 of 32 specimens are predicted within ±30 % of the measured value, while a_{20} and a_{10} are 0.625 and 0.344.

Table 2. Performance of the XGBoost model on the training and test sets.

Set	*N*	R^2	CI	KGE	a_{20}	a_{30}	MAE (C)	RMSE (C)	MAPE (%)
Training	94	0.995	0.997	0.988	1.000	1.000	37.2	54.9	3.98
Test	32	0.915	0.943	0.850	0.625	0.875	224.5	308.0	19.31

The drop from training to test performance (R^2 decreases by 0.080, RMSE increases about 5.6 times) is expected for a flexible ensemble trained on 94 observations: the test RMSE (308.0 C) is even lower than the cross-validated RMSE obtained during tuning (371.2 C), which confirms a realistic error estimate. The test RMSE is about 22 % of the mean test RCPT (1412 C), acceptable given the scatter of RCPT and the combination of tests from different laboratories.

The test points scatter around the line of agreement without strong global bias (Fig. 3a): the median ratio of experimental to predicted RCPT is about 1.02 (interquartile range 0.82–1.15). Two systematic patterns are visible. First, the four highest test values (≈ 3100–4600 C) are underpredicted; the largest error (970.7 C) concerns the highest RCPT of the database (4600 C, predicted at about 3630 C), because predictions of tree-based models are bounded by the target range seen in training (maximum 4000 C). Second, the largest relative errors (Fig. 3b) occur at low RCPT values (below about 1000 C), up to +79 % for a specimen of about 740 C and +48 %

for the 230 C specimen, which lies below the training minimum; these are absolute errors of a few hundred coulombs at most.

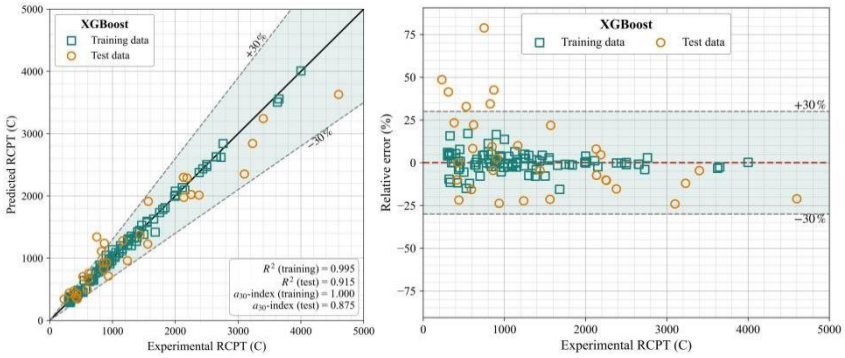


Fig. 3. (a) Predicted versus experimental RCPT (solid line: perfect agreement; dashed lines: ±30 % band) and (b) relative error of the predictions versus experimental RCPT.

4.2 Relative Importance of the Input Variables

The importance of the inputs was assessed with the mean absolute SHAP . value and the relative gain, which give the same ranking (Fig. 4): curing age ≫ FA > B > W/B > SP > f_{cj}. Curing age dominates, with a mean |SHAP| value of 436.0 C (about 40 % of the total) and a gain share of 31.2 %, about 2.5 times the |SHAP| value of FA (174.7 C; gain 21.4 %). B (162.3 C; 17.3 %) and W/B (122.1 C; 11.5 %) follow, while SP (103.3 C; 9.6 %) and f_{cj} (89.1 C; 9.0 %) contribute least.

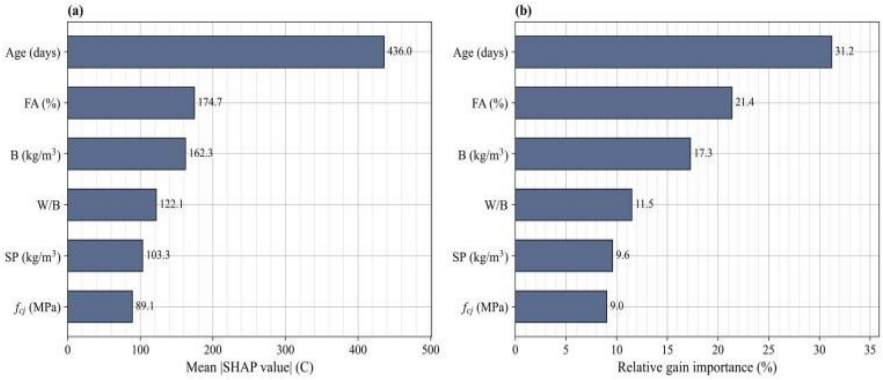


Fig. 4. Importance of the input variables: (a) mean absolute SHAP value and (b) relative gain.

4.3 Effect of the Input Variables on the Predicted RCPT

The marginal effects are shown in Fig. 5 by the ICE curves of individual samples (grey) and their average, the PD curve (blue); the step-like curves are less reliable in sparsely populated regions (rug marks).

Curing age (Fig. 5a). The average predicted RCPT decreases monotonically with age, from about 1720 C at 28 days to 1250 C at 56 days and 800–870 C at 90–91 days, and levels off at about 650 C at 180 days and 610 C at 365 days, reflecting continuing hydration and the delayed pozzolanic reaction of fly ash.

Fly ash content (Fig. 5c). The predicted RCPT is highest (about 2570 C) with no or very little fly ash, drops sharply at about 15 % and to about 1400 C at 20 %, and approaches 1100 C above about 45 %. The benefit of fly ash is thus largest for the first 20–30 % of replacement, in line with the literature [3]; since few observations exist between 0 % and 30 % FA, the abruptness of the drop partly reflects data sparsity and is not a material threshold.

Water-to-binder ratio (Fig. 5d). The influence of W/B is weaker and non-monotonic: about 1100 C below W/B = 0.32, a local maximum of about 1600 C at 0.40 and about 1200 C above 0.42. The weak tendency of higher RCPT at higher W/B agrees with a more porous matrix, but it is modest and partly confounded with B and f_{cj}.

B, SP and f_{cj} (Figs. 5b, 5e, 5f). The predicted RCPT increases in steps with binder content, from about 930 C below 450 kg/m^3 to about 1550 C above 515 kg/m^3, a conditional effect given the correlation of B with SP and W/B. SP and f_{cj} have little consistent influence.

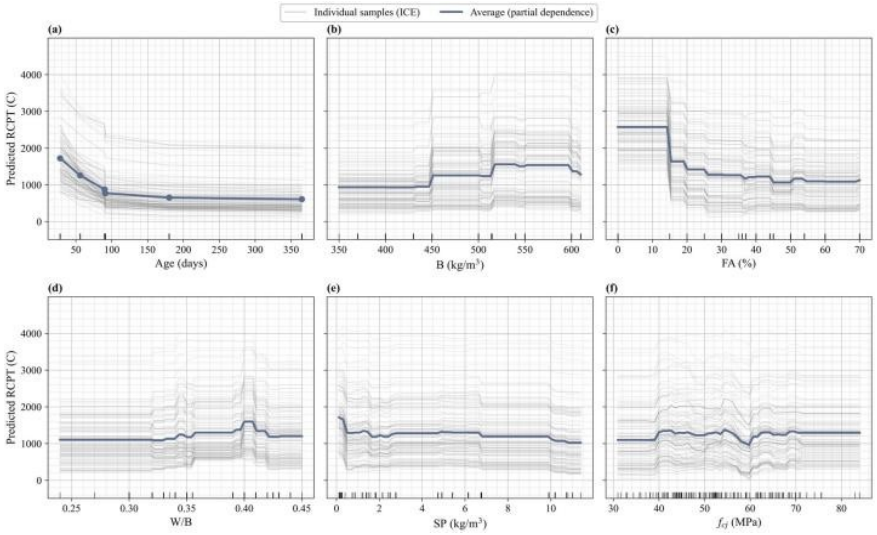


Fig. 5. ICE curves (grey) and partial dependence (blue) of the predicted RCPT for (a) age, (b) B, (c) FA, (d) W/B, (e) SP and (f) f_{cj}. Rug marks show the distribution of the observations.

Limitations. The database is small and compiled from different sources, curing age takes only six values, the model cannot extrapolate beyond the training ranges (Table 1), the PD curves describe the fitted model and not necessarily causal effects, and a single random partition was used.

5 Conclusions

An XGBoost model tuned by Bayesian optimization was developed to predict the RCPT of fly ash-containing SCC from six variables, using 126 literature results. The main conclusions are:

1. The training and test sets are statistically consistent (Kolmogorov–Smirnov $p > 0.05$) and the multicollinearity among inputs is acceptable (VIF ≤ 3.87).
2. The model predicted the unseen test data with high accuracy ($R^2 = 0.915$; CI = 0.943; KGE = 0.850; RMSE = 308.0 C; MAPE = 19.3 %), with 87.5 % of the test specimens within ±30 % of the experimental value.
3. The largest errors occur at the extremes of the RCPT range (underprediction of the highest values, overprediction of some low values); no systematic bias was detected.
4. SHAP and gain analyses both ranked the variables as age ≫ FA > B > W/B > SP > f_{cj}; curing age accounts for about 40 % of the total mean |SHAP| value.
5. The predicted RCPT decreases strongly with age up to about 90 days and with fly ash content (largest benefit in the first 20–30 % of replacement), whereas the effect of W/B is weaker, non-monotonic and partly confounded with that of binder content.

The model is an efficient tool for the preliminary durability assessment of fly ash-based SCC within the ranges of the database.

References

1. Okamura, H., Ouchi, M.: Self-compacting concrete. J. Adv. Concr. Technol. 1(1), 5–15 (2003).
2. ASTM C1202: Standard test method for electrical indication of concrete's ability to resist chloride ion penetration. ASTM International (2019).
3. Boğa, A.R., Topçu, İ.B.: Influence of fly ash on corrosion resistance and chloride ion permeability of concrete. Constr. Build. Mater. 31, 258–264 (2012).
4. Mohamed, O.A., Ati, M., Al Hawat, W.: Using artificial neural networks to predict chloride penetration of sustainable self-consolidating concrete. In: Civil-Comp Proceedings 109, Paper 37. Civil-Comp Press (2015).
5. Kumar, D.R., Kumar, S., Senjuntichai, T., Kaewunruen, S.: Hybrid explainable machine learning models for predicting rapid chloride penetration test and sorptivity of self-compacting concrete with fly ash and silica fume under thermal exposure. Front. Struct. Civ. Eng. (2026).
6. Friedman, J.H.: Greedy function approximation: a gradient boosting machine. Ann. Stat. 29(5), 1189–1232 (2001).
7. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proc. 22nd ACM SIGKDD, pp. 785–794. ACM (2016).
8. Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.: Optuna: a next-generation hyperparameter optimization framework. In: Proc. 25th ACM SIGKDD, pp. 2623–2631. ACM (2019).
9. Bergstra, J., Bardenet, R., Bengio, Y., Kégl, B.: Algorithms for hyper-parameter optimization. In: NIPS 24, pp. 2546–2554 (2011).
10. Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: NIPS 30, pp. 4765–4774 (2017).

Machine Learning-Based Prediction of Seismic Response and Fragility Curves for Reinforced Concrete Structures

Nesrine Atallah[1[0009-0009-1396-2261]], Kawther Arbia Boudjelthia[1[0009-0002-2418-4425]], Zamila Harichane[1[0000-0003-2581-0344]] and Halima Belbahi[2[0009-0001-8307-6224]]

[1] Department of Civil Engineering, Hassiba Benbouali University of Chlef, Algeria
[2] Laboratory of Acoustics and Civil Engineering, Djilali Bounaama University of Khemis Miliana, Algeria
z.harichane@univ-chlef.dz

Abstract. Seismic vulnerability assessment of reinforced concrete (RC) buildings is a central task in performance-based earthquake engineering, yet nonlinear time-history analysis (NLTHA) becomes computationally prohibitive for large building inventories. This paper investigates a machine learning workflow for predicting the peak and residual interstorey drift ratios (IDR) of RC moment-resisting frames and for deriving fragility curves from the predicted response. A numerical database was generated from 34 building models (336 storey-level analysis points) subjected to 148 ground-motion records, giving 49,728 simulations, of which 40,017 valid observations remained after removing diverged or collapse-type runs. Decision Tree, Random Forest and XGBoost regressors were trained on 12 seismic and structural features, tuned with Optuna and evaluated with R^2, MAE and RMSE. XGBoost performed best, with test R^2 values of 0.847 for peak IDR and 0.749 for residual IDR. SHAP analysis showed that moment magnitude and epicentral distance dominated both targets, whereas lateral stiffness and yield strength gained importance for residual drift. The predicted peak IDR was then converted into magnitude-based fragility curves with a logistic generalised linear model. Because the database derives from a single reference building and excludes collapse-type runs, the results demonstrate the feasibility of the workflow, and the fragility curves are conditional on convergent response.

Keywords: reinforced concrete frames, seismic fragility curves, machine learning, XGBoost, interstorey drift ratio.

1 Introduction

Performance-based earthquake engineering (PBEE) links ground-motion intensity to structural response, damage and loss [1]. For frame buildings, the peak interstorey drift ratio (IDR), defined as the relative displacement between adjacent floors divided by the storey height, is the most widely adopted demand measure. Demand is conventionally estimated by NLTHA or incremental dynamic analysis (IDA) [2], and fragility functions convert it into exceedance probabilities of damage states. These analyses are

dependable, but they must be repeated for every building and record, which becomes prohibitive for large inventories.

Machine learning (ML) has been proposed as a way to reduce this burden [3]. Among the building-level studies considered here, the learned quantity is usually a fragility parameter or a damage class rather than the structural response itself. Decision-tree models estimated the median and dispersion of fragility curves from 214 collected cases [4], and several algorithms, XGBoost among them, were trained on IDA results for 165 RC frame models to predict fragility curves [5].

Two aspects receive less attention. First, the studies cited above do not report residual drift as a prediction target, although it governs post-earthquake repairability and demolition decisions [6]. Second, a surrogate inherits the scope of its training simulations, so the treatment of numerically diverged runs determines what can legitimately be concluded.

This study therefore trains and compares three tree-based regressors to predict the peak and residual IDR of RC frame buildings, interprets the best model with SHapley Additive exPlanations (SHAP), and builds magnitude-based fragility curves from its predictions with a logistic generalised linear model (GLM). The scope is deliberately limited: all models derive from one reference building, and the aim is to test the feasibility of the workflow, not to provide a general-purpose predictor.

2 Database Generation and Preparation

2.1 Reference Building and Structural Models

All models derive from one reference building: a 15-storey RC moment-resisting frame built in 1950 in Mexico City, 53.5 m high, with a rectangular plan of about 17.0 m × 53.0 m and three transverse and thirteen longitudinal bays [7]. Controlled variations of lateral stiffness and floor mass produced 34 building models in five height classes (2, 4, 6, 10 and 15 storeys, with 2, 4, 6, 10 and 12 models respectively). Each storey is a separate data point, giving 336 storey-level analysis points. The workflow is summarised in Fig. 1.

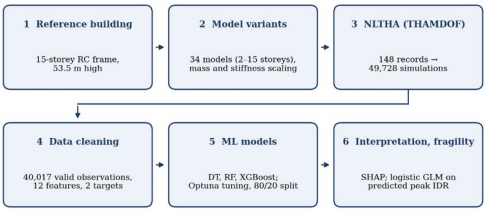


Fig. 1. Workflow from the reference building to the interpretation and fragility stages.

NLTHA was carried out with THAMDOF [8], a MATLAB-based tool that idealises each building as a shear-type lumped-mass system with nonlinear interstorey springs. Storey hysteresis follows a tri-linear model with stiffness degradation, yielding and

post-yield softening; second-order (P–Δ) effects are included, and integration uses Newmark's constant average acceleration scheme with Newton–Raphson iterations. THAMDOF was used as published [8], without calibration against test data, so results are model-based and not validated against real behaviour.

2.2 Ground Motions, Response Quantities and Input Features

The excitation comprises 148 records. Moment magnitude ranges from 3.4 to 9.0, epicentral distance from 14.9 to 277.0 km, focal depth from 5 to 50 km, and the average shear-wave velocity of the top 30 m (V_{s30}) from 237 to 585 m/s. Each row of the database holds the response of one storey of one model under one record, so 336 × 148 gives 49,728 rows. Two response quantities were extracted. The peak IDR is the largest relative displacement between adjacent floors divided by the storey height. The residual IDR is the permanent drift left after shaking; its sign depends on the direction of the last large inelastic excursion [6], so its magnitude is used here. The 12 input features comprise four seismic and eight structural parameters (Table 1).

Table 1. Input features of the ML dataset.

Group	Features
Seismic (4)	M_w moment magnitude; D_{epi} epicentral distance; H_{foc} focal depth; V_{s30} shear-wave velocity of the top 30 m
Structural (8)	T fundamental period; N_{st} number of storeys; n_{st} storey index; h storey height; W_i floor seismic weight; P gravity load on the storey; K_i storey lateral stiffness; F_y storey yield strength

2.3 Data Cleaning

When demand approaches the collapse capacity, time-domain integration may fail to converge and return unbounded displacements (P–Δ instability, Newton–Raphson divergence or negative tangent stiffness), so IDR values above 1.0 were treated as non-converged or collapse-type outcomes and excluded, since a shear-type model cannot separate numerical divergence from true collapse once the response is unbounded. A four-step sequential filter removed (i) 9,602 runs with IDR > 1.0 (19.31 % of the database), (ii) 80 runs with IDR ≤ 0, (iii) 27 runs in which the peak response converged but no residual drift was returned, and (iv) two runs with a sentinel residual value above 1.0 or an exact zero residual value, which cannot be log-transformed. In total 9,711 records were removed, leaving 40,017 valid observations (80.47 %).

Invalid runs concentrated where demand was highest (Fig. 2): the invalid share rose from 6.4 % for the 2-storey models to 21.8 % for the 15-storey models, and was 81.2 % for records with M_w between 8 and 9, against 0.3 % below M_w 5. Divergence is therefore demand-driven, and collapse-type cases are absent from the retained data, which affects the fragility analysis (Section 5).

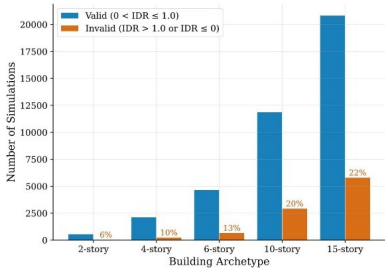


Fig. 2. Database: valid and invalid simulations by building archetype, with the invalid share in percentage.

2.4 Targets and Data Split

Both targets span several orders of magnitude and were $\log_{10}$-transformed. This gave near-symmetric distributions (skewness −0.298 for $\log_{10}$ IDR and 0.153 for $\log_{10}$ | residual IDR|), in line with the lognormal description commonly adopted for demand parameters [1]. The data were split 80/20 by stratified random sampling over 25 strata (five magnitude bins × five height classes), giving 32,013 training and 8,004 test observations. Kolmogorov–Smirnov tests found no difference between the target distributions of the two subsets (p = 0.978 and 0.687). A random split was used because all 34 models derive from one building, so no independent building exists to hold out (Section 5).

3 Methodology

3.1 Models, Optimisation and Interpretation

Three tree-based regressors were compared: a Decision Tree (DT), a Random Forest (RF) and XGBoost [9]. Hyperparameters were tuned with Optuna using the Tree-structured Parzen Estimator sampler. The objective was the mean root-mean-square error (RMSE) over five cross-validation folds of the training set. Each study ran for up to 500 trials, with a median pruner and early stopping after 200 trials without improvement. Because the objective is shared, one hyperparameter set per model serves both targets. Each model was then refitted to the full training set and evaluated on the test set, which was not used during tuning. Performance was measured by the coefficient of determination (R^2), the mean absolute error (MAE) and the RMSE, all computed on the $\log_{10}$-transformed targets, so that errors are in $\log_{10}$ units. SHAP [10] assigns each feature an additive contribution to every prediction. For the best model, the TreeSHAP algorithm was applied to the training set; mean absolute SHAP values give global importance, and beeswarm plots show the direction of each effect.

3.2 Fragility Curves from Predicted Response

For each IDR limit s in {0.01 %, 0.05 %, 0.10 %, 1.00 %}, every response is converted into a binary exceedance indicator. The probability of exceedance is modelled with a binomial GLM with logit link as a function of the standardised moment magnitude z, which is the only intensity measure used:

$$z = (M_w - \mu) / \sigma \quad (1)$$

$$P_f(M_w) = 1 / (1 + \exp[-(\beta_0 + \beta_1 z)]) \quad (2)$$

where μ and σ are the sample mean and standard deviation of the magnitudes, and the coefficients β_0 and β_1 are estimated by maximum likelihood. Applying the procedure to the predicted IDR gives the ML-based curves. The four limits serve as exceedance levels of drift; their relation to code-based damage states is discussed in Section 5. Magnitude was the only scenario variable common to all records; PGA and spectral ordinates are left for future work.

4 Results

4.1 Predictive Performance

XGBoost gave the best results for both targets (Table 2). For peak IDR it reached a test R^2 of 0.847, with MAE 0.279 and RMSE 0.326 ($\log_{10}$ units), against 0.827 for DT and 0.828 for RF. For residual IDR the test R^2 was 0.749 for XGBoost, against 0.726 and 0.732 for DT and RF. From training to test, R^2 fell by 0.009–0.011 for peak IDR and by 0.012–0.021 for residual IDR, so overfitting was limited. DT and RF performed almost identically. Observed and predicted values for XGBoost are compared in Fig. 3.

Table 2. Performance of the three models ($\log_{10}$-transformed targets).

Target	Model	Train R^2	Test R^2	Test MAE	Test RMSE
$\log_{10}$ IDR	DT	0.838	0.827	0.290	0.348
	RF	0.837	0.828	0.289	0.346
	XGBoost	**0.858**	**0.847**	**0.279**	**0.326**
$\log_{10}$ |residual IDR|	DT	0.747	0.726	0.460	0.617
	RF	0.744	0.732	0.456	0.611
	XGBoost	**0.768**	**0.749**	**0.437**	**0.590**

Residual drift is harder to predict than peak drift. It depends on the loading path and on hysteretic energy dissipation, which scalar input features cannot fully describe. The error distributions were centred near zero for all models, so no global bias was found; for XGBoost on peak IDR the central half of the errors lay between about −0.25 and

+0.20 $\log_{10}$ units. Taylor-diagram analysis (not shown) indicated that predictions had a smaller spread than observations, so extreme responses should be treated with caution.

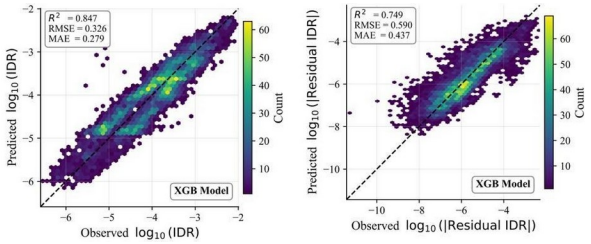


Fig. 3. Observed versus predicted values on the test set for XGBoost: (a) $\log_{10}$ IDR, (b) $\log_{10}$ | residual IDR|. Colour shows point density per hexagonal bin; the dashed line is the 1:1 line.

4.2 Feature Importance

Moment magnitude was the most influential feature for both targets (Table 3, Fig. 4), at 0.43 $\log_{10}$ units for peak IDR, more than 2.5 times the value of the second feature. Larger magnitudes raised the predicted IDR, with contributions spanning roughly −1.0 to +0.75 $\log_{10}$ units (Fig. 4a). A high yield strength lowered the predicted peak IDR.

Table 3. Five largest mean absolute SHAP values ($\log_{10}$ units) for XGBoost.

	Peak IDR		Residual IDR	
Rank	**Feature**	**Mean |SHAP|**	**Feature**	**Mean |SHAP|**
1	M_w	0.43	M_w	0.56
2	D_{epi}	0.16	D_{epi}	0.27
3	H_{foc}	0.13	K_i	0.12
4	F_y	0.08	F_y	0.11
5	V_{s30}	0.07	H_{foc}	0.10

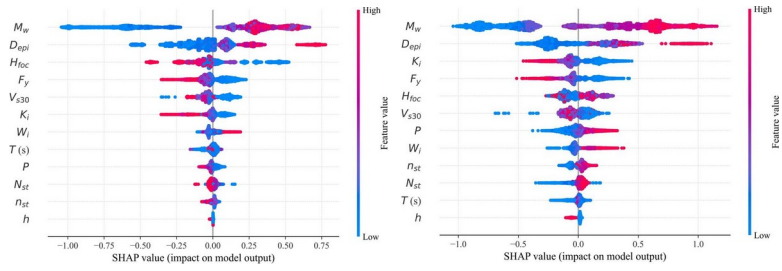


Fig. 4. SHAP beeswarm plots for XGBoost: (a) $\log_{10}$ IDR, (b) $\log_{10}$ |residual IDR|.

For residual IDR, epicentral distance gained weight (0.27) and lateral stiffness rose to third place (0.12). Stiffer structures were predicted to have smaller residual drift, whereas heavier floors raised it (Fig. 4b). In both models, larger epicentral distances

were associated with higher predicted drift. This should not be read as a physical effect: in this record suite larger events lie at longer distances (correlation 0.69 between D_{epi} and M_w), so epicentral distance partly acts as a proxy for source energy. In addition, SHAP describes the behaviour of the fitted model and is not a causal analysis.

4.3 Fragility Curves

Fig. 5 shows the XGBoost-predicted IDR against magnitude for the full dataset, together with the fragility curves from Eq. (2). The curve for the lowest limit (0.01 %) rises first. The curves for 0.05 % and 0.10 % rise more slowly and reach about 0.76 and 0.57 at M_w 9. The curve for 1.00 % stays at zero. This follows from the data and does not demonstrate collapse safety: the largest retained IDR is 9.72×10^{-3}, below 1.00 %, because the 9,602 diverged runs were excluded. The fragility curves are therefore conditional on convergent response.

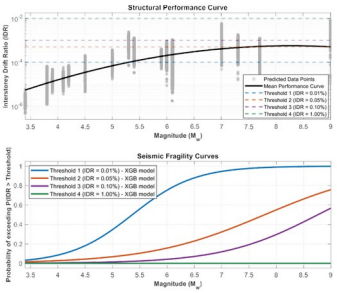


Fig. 5. Performance and fragility curves from XGBoost-predicted IDR for the full dataset. Top: predicted IDR against M_w, mean trend and the four IDR limits. Bottom: fragility curves from the logistic GLM.

The same pipeline was applied to each height class. At M_w 6.5, the exceedance probabilities for the 0.01 %, 0.05 % and 0.10 % limits were 88 %, 31 % and 9 % for the 6-storey class and 89 %, 23 % and 4 % for the 15-storey class, so the dependence on height is not monotonic in these results, and explaining it requires period-dependent intensity measures.

5 Discussion and Limitations

The results support tree-based surrogates for peak and residual drift within the space covered by the database, whose scope is bounded in six ways. First, all models are variants of one 15-storey building idealised as shear-type systems, so the findings do not extend to other structural systems or to detailed frame models; generalisation to other buildings is not claimed. Second, storey-level rows were split at random, so storeys of the same model and record can appear in both training and test sets; the test scores therefore describe interpolation within the sampled models and records, and validation that holds out whole models or records is needed. Third, collapse-type cases

were removed, so the fragility curves describe only convergent response, and they were derived from predicted responses only, without comparison against curves from simulated responses (not yet derived). Fourth, magnitude is the only intensity measure in Eq. (2). Fifth, the four IDR limits (0.01–1.00 %) differ from the damage-state thresholds of 0.2, 0.5, 1.5 and 4.0 % commonly tabulated for RC frames, so the curves show exceedance of drift levels, not of damage states, until a code-based set is adopted. Sixth, computational savings were not measured.

6 Conclusions

This paper developed and tested an ML framework for predicting seismic drift response and deriving fragility curves for RC moment-resisting frames from 40,017 valid NLTHA observations. XGBoost outperformed Decision Tree and Random Forest, with test R^2 of 0.847 for peak IDR and 0.749 for residual IDR. SHAP showed that moment magnitude and epicentral distance dominate both targets, and that stiffness and yield strength matter more for residual drift. Predicted peak IDR was converted into magnitude-based fragility curves through a logistic GLM. Future work should add comparison with curves from simulated responses, other structural systems, group-wise validation, several intensity measures and measured computing cost.

7 References

1. Krawinkler, H., Miranda, E.: Performance-based earthquake engineering. In: Earthquake Engineering. CRC Press (2004)
2. Vamvatsikos, D., Cornell, C.A.: Incremental dynamic analysis. Earthq. Eng. Struct. Dyn. 31(3), 491–514 (2002)
3. Thai, H.-T.: Machine learning for structural engineering: a state-of-the-art review. Structures 38, 448–491 (2022)
4. Dabiri, H., Faramarzi, A., Dall'Asta, A., Tondi, E., Micozzi, F.: A machine learning-based analysis for predicting fragility curve parameters of buildings. J. Build. Eng. 62, 105367 (2022)
5. Kazemi, F., Asgarkhani, N., Jankowski, R.: Machine learning-based seismic fragility and seismic vulnerability assessment of reinforced concrete structures. Soil Dyn. Earthq. Eng. 166, 107761 (2023)
6. Ruiz-García, J., Miranda, E.: Evaluation of residual drift demands in regular multi-storey frames for performance-based seismic assessment. Earthq. Eng. Struct. Dyn. 35(13), 1609–1629 (2006)
7. Peña, F., Ramos, J.: Dynamic characteristics of a 1950s heritage building: a comparison of original design methods and modern techniques. Buildings 14(9), 2944 (2024)
8. Pheresi: THAMDOF: time history analysis of multi-degree-of-freedom systems, https://github.com/pheresi/THAMDOF, last accessed 2026/10/25
9. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. ACM (2016)
10. Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30, pp. 4765–4774 (2017)

Machine Learning-Based Prediction of the Nonlinear Seismic Response of Reinforced Concrete Frames

Yassamine Zahia Klouche [1][0009-0006-3159-531X], Zamila Harichane[1][0000-0003-2581-0344] and Halima Belbahi[2][0009-0001-8307-6224]

[1] Geomaterials Laboratory, Department of Civil Engineering, Faculty of Civil Engineering and Architecture, Hassiba Benbouali University of Chlef, Chlef, Algeria
[2] Laboratory of Acoustics and Civil Engineering, Djilali Bounaama University of Khemis Miliana, Algeria
z.harichane@univ-chlef.dz

Abstract. Nonlinear time-history analysis (NLTHA) remains the benchmark for assessing the seismic performance of buildings, yet its computational burden restricts its use in regional risk studies and in design optimisation. This paper develops machine-learning surrogates that estimate three engineering demand parameters of reinforced concrete (RC) moment-resisting frames: the peak inter-storey drift ratio (IDR), the residual IDR and the shear force. A simulation database was assembled by subjecting four archetypes of 3, 6, 10 and 15 storeys to 148 recorded ground motions, which produced 67,044 storey-level results; after the removal of numerically diverged runs, 60,824 valid observations were retained. Eleven seismic and structural features feed Decision Tree, Random Forest and XGBoost regressors, whose hyperparameters were tuned with Optuna and five-fold cross-validation. XGBoost was the most accurate model, reaching test R^2 values of 0.859 for the IDR, 0.815 for the residual IDR and 0.915 for the shear force, with only a small gap between training and test performance. SHAP values, together with ICE and partial dependence plots, show that moment magnitude is the leading predictor of every target, followed by source-to-site and site parameters, whereas the influence of the structural features differs from one response quantity to another. The trained surrogates offer a rapid and physically coherent means of estimating seismic demands for probabilistic risk assessment and performance-based design.

Keywords: Reinforced concrete frames, Seismic response, Machine learning, XGBoost, Inter-storey drift, Nonlinear time-history analysis, SHAP interpretability.

1 Introduction

Performance-based earthquake engineering (PBEE) judges a building through engineering demand parameters (EDPs) such as the peak inter-storey drift ratio (IDR), the residual drift and the storey shear force. Nonlinear time-history analysis (NLTHA) is the most dependable way to obtain them [1], but its computing cost becomes

prohibitive when thousands of analyses are required, as in incremental dynamic analysis, regional loss estimation or design optimisation.

Machine learning (ML) can move this cost offline: a model trained once on simulations returns the response of an unseen building–record pair almost instantly. Kazemi et al. [2] built 92,400 data points from incremental dynamic analyses of 165 reinforced concrete (RC) frames and showed that XGBoost predicted the maximum IDR better than improved neural networks; related work includes a database of steel-frame responses [1] and deep-learning models of response histories [3]. Two gaps remain: many predictors are black boxes whose physical plausibility is rarely checked, and most studies address a single response quantity, whereas PBEE needs peak drift, residual drift and force demands together. This paper therefore (i) builds a database of RC moment frames from 67,044 storey-level NLTHA outcomes under 148 recorded accelerograms, (ii) compares Decision Tree, Random Forest and XGBoost models tuned by Bayesian optimisation for the three EDPs, and (iii) examines with SHAP, ICE and partial dependence plots whether the learnt relationships are physically sound.

2 Database Generation

2.1 Structural Archetypes and Simulation

Four RC moment-frame archetypes with 3, 6, 10 and 15 storeys were adopted (Fig. 1). All have a storey height of 3.30 m; the lateral stiffness (1,128–1,500 N/cm) and the floor weight (392–4,413 tf) were varied to produce 5, 8, 12 and 18 variants, respectively (43 in total) and 453 storey-level cases, each analysed under 148 ground motions, which gives the 67,044 rows of the raw database. The analyses were run in THAMDOF [4], a MATLAB platform that idealises a building as a lumped-mass shear structure. The equations of motion are integrated with Newmark's constant-average-acceleration method and Newton–Raphson iterations; a tri-linear hysteretic law reproduces yielding, stiffness degradation and post-yield softening, and P–Δ effects are included.

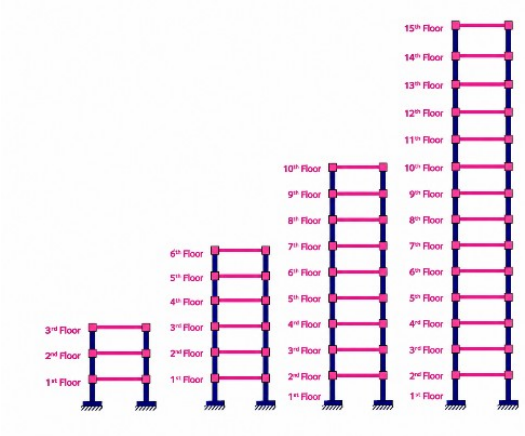


Fig. 1. RC moment-frame archetypes of 3, 6, 10 and 15 storeys used to build the database.

2.2 Ground Motions, Input Features and Quality Control

The excitation consists of 148 real horizontal accelerograms from the KiK-net network [5]. The suite spans moment magnitudes M_w = 3.4–9.0, epicentral distances D_{epi} = 14.9–277 km, focal depths H_{foc} = 5–50 km and shear-wave velocities V_{s30} = 237–585 m/s. The peak IDR of a storey is given by Eq. (1)

$$IDR = \Delta / h \tag{1}$$

where Δ is the relative displacement between consecutive floors and h is the storey height; a value above 1.0 is geometrically impossible and can only stem from numerical failure. The residual IDR is the permanent drift remaining after shaking, normalised by h, and the shear force is derived from the storey forces of the analysis. The eleven predictors are the four seismic parameters above and seven structural parameters of each storey: the fundamental period T, the total number of storeys N_{story}, the storey number n_{st}, the floor seismic weight W_i, the gravity load P, the storey lateral stiffness K_i and the storey yield strength F_y.

In the raw database, 6,188 runs (9.23%) have an IDR above 1.0, with a maximum of 6.24×10^{304}, which reveals divergence of the nonlinear solver. They concentrate in the strongest events (85.7% belong to records with $M_w \approx 9.0$) and in tall buildings (the 15- and 10-storey archetypes hold 61.6% and 29.5% of them). The residual IDR is missing in 5,634 rows, of which 5,602 coincide with diverged runs; the other 32 are near-collapse cases that cannot be imputed on a physical basis. Both groups were removed, leaving 60,824 valid observations (90.72%). Negative residual drifts were retained (53.2% of the records), because the sign only gives the direction of the permanent displacement.

2.3 Target Transformation and Partitioning

The targets are strongly skewed and span many orders of magnitude, so a $\log_{10}$ transformation was applied; it compresses the range, approximates the lognormal form assumed in probabilistic seismic demand analysis and linearises the feature–target relations, and predictions are back-transformed with 10^x. The data were split into a training set (48,659 rows, 80%) and a test set (12,165 rows, 20%) by stratified sampling on magnitude bin and archetype; a Kolmogorov–Smirnov test on $\log_{10}$(IDR) (statistic 0.008, p = 0.556) shows that both subsets share the same distribution.

3 Machine Learning Methodology

3.1 Algorithms and Hyperparameter Optimisation

Three tree-based regressors were compared: the Decision Tree (DT), which applies CART with variance-reducing binary splits; the Random Forest (RF) [6], which averages trees grown on bootstrap samples with random feature subsets; and XGBoost [7], which adds trees sequentially, each correcting the errors of the current ensemble, with a regularised objective, second-order gradient information, shrinkage and

subsampling. Hyperparameters were tuned with Optuna [8]. For each of the nine algorithm–target pairs, the objective was the mean RMSE of the $\log_{10}$ target over five-fold cross-validation on the training set; each study was limited to 500 trials with a median pruner, and the number of XGBoost trees (maximum 2,000) was set by early stopping. The optimal settings differ among targets: for example, leaf-wise growth (8 leaves, learning rate 0.05) for the IDR and shallow depth-wise trees (depth 3) with strong L1 regularisation for the shear force. Each model was finally refitted on the whole training set.

3.2 Metrics and Interpretability Tools

Accuracy was measured with the coefficient of determination R^2, the root mean square error (RMSE) and the mean absolute error (MAE) in $\log_{10}$ space; learning curves complemented this assessment. The coefficient of determination is given by Eq. (2)

$$R^2 = 1 - \Sigma(y_i - \hat{y}_i)^2 / \Sigma(y_i - \bar{y})^2 \quad (2)$$

where y_i, $\hat{y}_i$ and $\bar{y}$ are the observed, predicted and mean observed values. For interpretation, SHAP values [9] split each prediction into additive feature contributions, and the global importance of feature i is given by Eq. (3)

$$\Phi_i = (1 / n)\ \Sigma_j\ |\varphi_i(x^{(j)})| \quad (3)$$

over the n test samples. ICE curves show the marginal effect of a feature for each sample, and their average is the partial dependence plot (PDP).

4 Predictive Performance

4.1 Accuracy on the Test Set

Table 1 and Fig. 2 report the results. XGBoost was the most accurate model for all three EDPs, with test R^2 of 0.859 (IDR), 0.815 (residual IDR) and 0.915 (shear force). For the IDR the three models are close (R^2 = 0.842 for DT and RF), whereas the advantage of boosting is larger for the harder targets: the residual-IDR RMSE falls from about 0.69 to 0.575, and the shear-force RMSE from 0.776 (DT) and 1.016 (RF) to 0.479.

The predictions in Fig. 2 follow the 1:1 line over seven (IDR), ten (residual IDR) and thirteen (shear force) orders of magnitude, and the residual box plots are centred on zero, i.e. without systematic bias but with a larger uncertainty for the residual drift (interquartile range of ±0.3 to ±0.5 log-units, against ±0.25 for the IDR).

The accuracy ranking of the targets, shear force, IDR and residual IDR, follows their physical predictability: the shear force is governed mainly by mass and spectral demand, which the features describe well; the IDR also depends on higher modes and stiffness degradation; and the residual IDR on hysteretic loop closure, which is very sensitive to the record and is reported as the hardest demand parameter to predict [10]. The poor RF result for the shear force (R^2 = 0.618) is probably due to its tuned

configuration, whose strong impurity constraints reduced the diversity of the trees for this high-variance target.

Table 1. Training and test of performance of the Decision Tree (DT), Random Forest (RF) and XGBoost (XGB) models for predicting IDR, residual IDR, and shear

Target	Model	R^2 train	RMSE train	MAE train	R^2 test	RMSE test	MAE test
$\log_{10}$(IDR)	DT	0.842	0.366	0.279	0.842	0.368	0.283
	RF	0.837	0.372	0.281	0.842	0.368	0.280
	XGB	0.865	**0.338**	**0.273**	**0.859**	**0.348**	**0.277**
$\log_{10}$(|Res. IDR|)	DT	0.773	0.636	0.468	0.733	0.691	0.502
	RF	0.741	0.680	0.483	0.731	0.693	0.489
	XGB	0.824	**0.560**	**0.415**	**0.815**	**0.575**	**0.425**
$\log_{10}$(|Shear|)	DT	0.805	0.723	0.372	0.777	0.776	0.384
	RF	0.647	0.973	0.493	0.618	1.016	0.505
	XGB	0.913	**0.484**	**0.305**	**0.915**	**0.479**	**0.305**

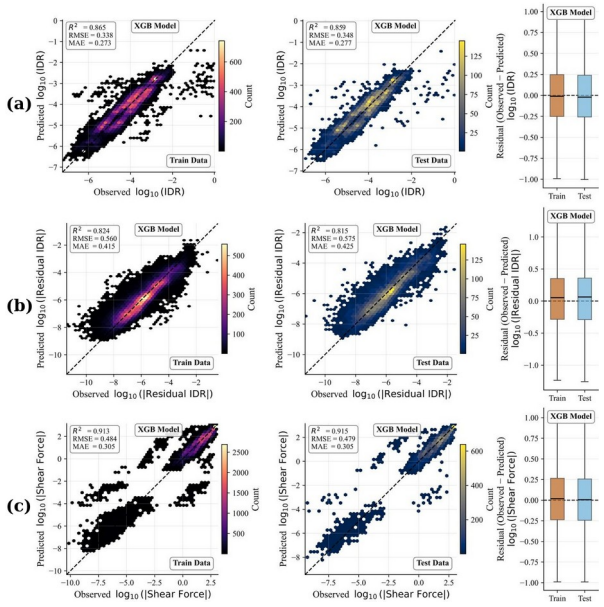


Fig. 2. Observed versus predicted values of the XGBoost models on the training set (left) and test set (centre), with residual box plots (right): (a) $\log_{10}$(IDR); (b) $\log_{10}$(|residual IDR|); (c) $\log_{10}$(|shear force|).

4.2 Generalisation

Overfitting is negligible. The largest drop in R^2 from training to test across the nine models is 0.040 (DT, residual IDR), and it stays below 0.01 for XGBoost for every

target (Table 1). The XGBoost learning curves (Fig. 3) end with small generalisation gaps of +0.018 (IDR), +0.024 (residual IDR) and +0.014 (shear force, curve not shown) log-units. Learning curves are obtained by five-fold cross-validation on the training set (48,659 rows); the maximum training size per fold is therefore 38,927 samples The IDR and shear-force curves flatten at the current database size, whereas the validation curve of the residual IDR is still decreasing beyond 25,000 samples, so this target would benefit most from additional data.

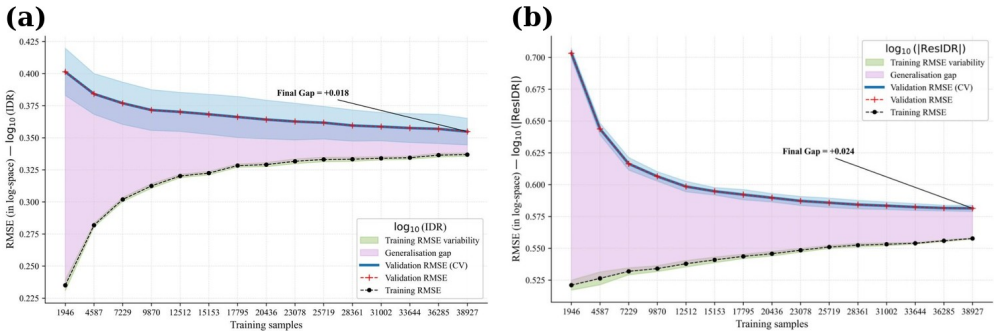


Fig. 3. Learning curves of the XGBoost models: (a) $\log_{10}$(IDR); (b) $\log_{10}$(|residual IDR|). The shaded areas show the generalisation gap and the variability of the cross-validated estimates.

5 Interpretability Analysis

5.1 Global Feature Importance

The SHAP and PDP values below refer to the XGBoost models and are expressed in $\log_{10}$ units. The mean absolute SHAP value Φ of each feature was computed for the three models. Moment magnitude ranks first for every EDP (Φ = 0.528, 0.600 and 0.549 for the IDR, residual IDR and shear force) and carries about half of the total SHAP mass of the IDR and shear-force models. The second tier depends on the target: the path and site variables H_{foc} (Φ = 0.144), V_{s30} (0.117) and D_{epi} (0.109) for the IDR; the gravity load *P* (0.386) and the period *T* (0.228) for the shear force; and D_{epi} (0.298) and *P* (0.202) for the residual IDR. The stiffness K_i and the yield strength F_y are always last (Φ < 0.01), owing to their narrow ranges in the database and because *T* already carries the stiffness information.

5.2 Feature Effects and Limitations

The ICE and PDP curves agree with the SHAP ranking: important features give the steepest PDPs, while K_i and F_y give flat ones. The PDP of M_w rises by about 2.2 log-units for the IDR (Fig. 4a) and saturates above $M_w \approx 5.5$ for the shear force (Fig. 4b). For the shear force the PDP of *T* falls from about 2.0 at 0.05 s to about 1.3 at 0.4 s (Fig. 4c), as stiffer buildings attract larger spectral accelerations. The PDP of V_{s30} decreases with velocity and flattens near 500–600 m/s (Fig. 4d).

Two PDP features should not be over-interpreted. The slight rise of the D_{epi} curve at large distances contradicts attenuation; it is a dataset confound, because distant records belong mostly to large events. The step-like jumps in other PDPs stem from the discrete set of magnitudes and archetypes, which a continuously sampled database would smooth. The surrogates are valid for the four regular archetypes and the lumped-mass idealisation, and back-transformation from log space introduces a bias that should be corrected when median values in physical units are needed.

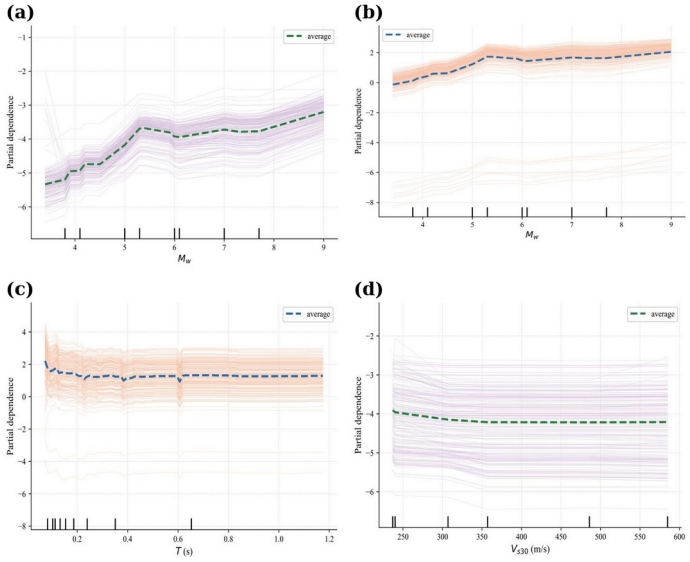


Fig. 4. ICE curves (thin lines) and PDP (dashed line): (a) M_w for the IDR; (b) M_w for the shear force; (c) T for the shear force; (d) V_{s30} for the IDR.

6 Conclusions

ML surrogates of nonlinear time-history analysis were developed and interpreted for RC moment frames from 60,824 valid observations. Three conclusions can be drawn. (1) XGBoost outperformed DT and RF for all targets (test R^2 = 0.859, 0.815 and 0.915 for the IDR, residual IDR and shear force) with negligible overfitting. (2) Predictability decreases from the shear force to the IDR and then to the residual IDR, matching the physics of these quantities, and the residual-IDR model would gain most from more data. (3) SHAP, ICE and PDP analyses identify moment magnitude as the dominant predictor, followed by path and site parameters, while the influence of the structural features depends on the EDP and stiffness and yield strength are negligible within the database ranges. Once trained, the models predict in a fraction of the time of a nonlinear analysis, which suits rapid demand estimation in risk assessment and performance-based design. Future work will extend the framework to irregular structures, richer databases and uncertainty quantification.

References

1. Guan, X., Burton, H., Shokrabadi, M.: A database of seismic designs, nonlinear models, and seismic responses for steel moment-resisting frame buildings. Earthquake Spectra 37(2), 1199–1222 (2021).
2. Kazemi, F., Asgarkhani, N., Jankowski, R.: Machine learning-based seismic response and performance assessment of reinforced concrete buildings. Arch. Civ. Mech. Eng. 23, 94 (2023).
3. Damikoukas, S., Lagaros, N.D.: MLPER: machine learning-based prediction model for earthquake response. Appl. Sci. 13(19), 10622 (2023).
4. Heresi, P.: THAMDOF: time-history analysis of multi-degree-of-freedom systems. https://github.com/pheresi/THAMDOF.
5. National Research Institute for Earth Science and Disaster Resilience (NIED): Strong-motion seismograph networks (K-NET, KiK-net). http://www.kyoshin.bosai.go.jp/.
6. Breiman, L.: Random forests. Mach. Learn. 45(1), 5–32 (2001).
7. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. ACM, New York (2016).
8. Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.: Optuna: a next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 2623–2631. ACM, New York (2019).
9. Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30, pp. 4765–4774. Curran Associates, Red Hook (2017).
10. Ruiz-García, J., Miranda, E.: Evaluation of residual drift demands in regular multi-storey frames for performance-based seismic assessment. Earthq. Eng. Struct. Dyn. 35(13), 1609–1629 (2006).

Multi-Layer Neural Network Classifier for Earthquake Building Damage Prediction

Halima Belbahi[1][0009-0001-8307-6224], Hichem Noura[1][0000-0003-0725-0937] and Mohamed Abed[2]

[1] Laboratory of Acoustics and Civil Engineering, Djilali Bounaama University of Khemis Miliana, Algeria
[2] Saad Dahleb University of Blida 1, Blida, Algeria
h.belbahi@univ-dbkm.dz

Abstract. Fast and dependable estimates of earthquake damage to buildings support both risk mitigation and emergency response. This paper presents a data-driven classifier that predicts the damage state of a building from only five inputs: construction age, number of storeys, a behaviour–quality ratio (R/Q), peak ground acceleration (PGA) and the distance to the rupture (R_{jb}). The model is a multi-layer neural network whose hyperparameters were tuned by Bayesian optimization and which was trained on 18,628 buildings surveyed after the 2003 Boumerdes (Algeria) earthquake and assigned to five ordered damage states, from negligible (D1) to collapse (D5). On an independent test set it reaches an accuracy of 0.93 and a macro-averaged F1-score of 0.80 (macro precision 0.81, macro recall 0.80), with comparable values on the training and validation subsets. Most errors occur between neighbouring damage states, in line with the gradual nature of seismic damage. A SHAP analysis ranks R/Q as the most influential input, followed by PGA and R_{jb}, so that the learned behaviour agrees with engineering expectations on structural capacity and seismic demand. Finally, the trained network is embedded in a graphical user interface that returns the predicted damage state together with class probabilities for decision support.

Keywords: Seismic vulnerability, Damage assessment, Boumerdes earthquake, Neural networks, SHAP analysis.

1 Introduction

Earthquakes are among the most destructive natural hazards. The damage a building suffers depends on the interplay between its structural configuration, construction quality and age and the intensity of the ground motion [1,2], so quantifying seismic vulnerability is a precondition for prioritizing retrofitting and planning the emergency response. Vulnerability methods are usually divided into analytical ones, based on structural models and nonlinear analyses, and empirical ones, based on damage observed after real earthquakes [3].

Machine learning (ML) can learn nonlinear relations among many interacting variables directly from data and has been applied to damage classification, fragility analysis and rapid screening [4,5]. Applications to North African building stocks remain comparatively scarce [6–8]. In addition, many models need long lists of descriptors that are rarely available after an earthquake, interpretability is often neglected, and few studies deliver the model as a tool for practitioners.

To address these gaps, this paper uses the survey database of the 21 May 2003 Boumerdes earthquake [9,10] to develop a neural network that relies on five physically meaningful inputs, is tuned by Bayesian optimization, is explained with SHAP and is embedded in a graphical user interface (GUI). Section 2 presents the data and methodology, Section 3 the network, Section 4 its performance, Section 5 the interpretation, Section 6 the GUI and Section 7 the conclusions.

2 Data and Methodology

Fig. 1 summarizes the workflow: data preparation, partitioning, tuning and training, evaluation, interpretation and deployment in the GUI.

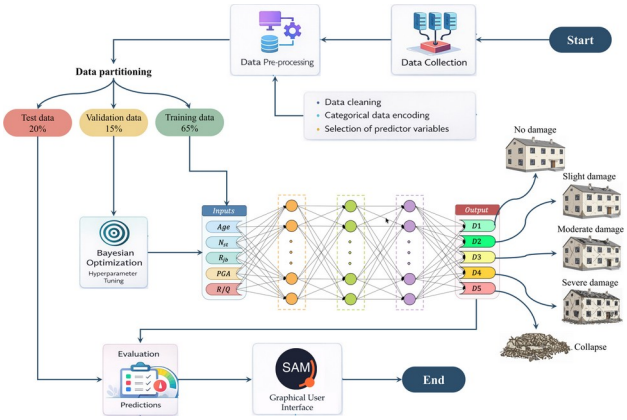


Fig. 1. Overall workflow of the study, from survey data to the graphical user interface.

2.1 Boumerdes Damage Survey

The data come from the survey carried out by the Algerian earthquake engineering research center (CGS) after the Mw 6.8 Boumerdes earthquake of 21 May 2003. Engineers inspected buildings visually with standardized forms [11] and recorded structural attributes, the position relative to the source and an overall damage grade, giving a database of 18,628 buildings.

2.2 Input and Output Variables

Five inputs were retained because they govern seismic response, are reliably available in the survey and are not redundant. Construction age (Age) is a proxy for material condition and for the seismic code in force at the time of construction [12]. The number of storeys (N_{st}) measures height, which affects inertia forces and damage mechanisms such as soft storeys [13]. The behaviour–quality ratio R/Q divides the behaviour coefficient R of the Algerian code RPA 99/2003 [14], which reflects ductility and energy dissipation, by the quality factor Q, which penalizes irregularity, poor redundancy and execution defects; the ratio condenses structural capacity into one indicator [15]. Seismic demand is described by PGA, a scalar measure of shaking intensity, and by R_{jb}, the shortest horizontal distance between the site and the surface projection of the rupture, which accounts for attenuation and source-to-site geometry. The output is the global damage state, an ordinal variable from D1 (negligible damage) to D5 (collapse).

2.3 Exploratory Analysis

Table 1 describes the inputs. The stock is dominated by relatively recent low-rise buildings (median age 20 years, median height one storey). R/Q is bimodal, with clusters near 2.17 and 4.78, because it derives from the discrete values of R and Q. R_{jb} spans 0.35–38.1 km and PGA 0.07–0.26 g, which gives enough variability in seismic demand.

Table 1. Descriptive statistics of the input variables (18,628 buildings).

Variable	Unit	Mean	Std	Min	25%	50%	75%	Max
Age	years	22.54	16.82	1.00	12.00	20.00	27.00	96.00
N_{st}	–	1.41	0.82	1.00	1.00	1.00	2.00	13.00
R/Q	–	3.61	1.38	2.17	2.17	4.78	4.78	6.50
R_{jb}	km	12.63	9.77	0.35	5.29	10.25	18.83	38.10
PGA	g	0.15	0.05	0.07	0.11	0.13	0.17	0.26

The damage classes are strongly imbalanced: D2 accounts for 60.26% of the buildings (11,225), D3 for 20.10% (3,744), D4 for 11.18% (2,083), D5 for 5.36% (999) and D1 for only 3.10% (577). The imbalance is handled through stratified splitting, class weighting and macro-averaged metrics.

2.4 Data Partitioning and Scaling

The data were split in a stratified way, so that every damage class keeps its relative frequency, into 12,040 training, 2,953 validation and 3,635 test buildings (64.6%, 15.9% and 19.5% of the data, respectively) The test subset served only for the final evaluation. The inputs were standardized as according to Eq. (1)

$$x_{\text{scaled}} = (x - \mu) / \sigma \quad (1)$$

where μ and σ are the mean and standard deviation of each input, estimated on the training subset only and applied unchanged to the other subsets to prevent information leakage.

3 Neural Network Classifier and Optimization

3.1 Architecture and Training

The classifier is a fully connected feedforward network that maps the input vector (Age, N_{st}, R/Q, R_{jb}, PGA) to five class scores. Each hidden layer applies an affine transformation followed by the Gaussian Error Linear Unit (GELU) activation [16], defined in Eq. (2)

$$\text{GELU}(z) = z\,\Phi(z) \quad (2)$$

where Φ is the cumulative distribution function of the standard normal law. Unlike ReLU, GELU is smooth and non-monotonic. The output layer has five neurons with a softmax activation [17], given by Eq. (3)

$$y_i = \exp(z_i) / \Sigma_{k=1}^{5} \exp(z_k) \quad (3)$$

so that the outputs can be read as class probabilities. The weights were learned by backpropagation, minimizing the sparse categorical cross-entropy, defined in Eq. (4)

$$L = -(1/N)\,\Sigma_{i=1}^{N} \log p_{i,ci} \quad (4)$$

where N is the number of samples and $p_{i,ci}$ is the probability assigned to the true class c_i of sample i. To counter the imbalance, each class was weighted inversely to its frequency (D1: 6.49, D2: 0.33, D3: 0.98, D4: 1.77, D5: 3.64). Batch normalization [18] follows every hidden layer to stabilize training. Dropout and weight decay add further regularization.

3.2 Bayesian Hyperparameter Optimization

Hyperparameters must be fixed before training; random search is more efficient than grid search when only a few matter [19]. We used Optuna [20] with its Tree-structured Parzen Estimator sampler, a sequential model-based (Bayesian) strategy that learns from the trials already evaluated. The objective was the macro-averaged F1-score on the validation subset, which weights all damage classes equally and is therefore better suited than accuracy to rare but critical states. We ran 500 trials (200 random start-up trials), each for at most 2,000 epochs with early stopping (patience of 100 epochs, best weights restored) and a fixed seed. Table 2 gives the search space and the optimum.

The selected network has three hidden layers of 94, 180 and 252 neurons with GELU activation and batch normalization, and dropout (rate 0.4) after the last hidden layer only.

Table 2. Hyperparameter search space and optimum found by Bayesian optimization.

Hyperparameter	Search space	Optimum
Batch size	{32, 64, 128, 256, 512, 1024}	1024
Number of hidden layers	1–4	3
Units per hidden layer	20–256	94, 180, 252
Dropout after each layer	on / off	off, off, on (0.4)
Activation function	GELU	GELU
Optimizer	Adam	Adam
Learning rate	10^{-3}–0.3 (log scale)	0.029
Weight decay	10^{-6}–0.3 (log scale)	2.26×10^{-5}

4 Performance Evaluation

Performance was measured on the independent test subset after retraining with the optimal settings. Accuracy, precision, recall and the F1-score were computed per class in a one-vs-rest manner and then averaged with equal weights (macro averaging), where TP, FP and FN denote true positives, false positives and false negatives. Confusion matrices and one-vs-rest ROC curves (with the area under the curve, AUC) complete the evaluation. Table 3 summarizes the results. On the test subset the network reaches an accuracy of 0.93 and a macro-averaged F1-score of 0.80, with macro precision of 0.81 and macro recall of 0.80. The small gap between training (0.94) and test (0.93) accuracy indicates limited overfitting.

Table 3. Performance of the optimized network (precision and recall and F1-score are macro-averaged).

Subset	Samples	Accuracy	Precision	Recall	F1
Training	12,040	0.94	0.85	0.84	0.95
Validation	2,953	0.93	0.83	0.80	0.94
Test	3,635	0.93	0.81	0.80	0.95

The confusion matrices in Fig. 2 are strongly diagonal. On the test subset, recall is 95.7% for D1, 85.1% for D2, 82.6% for D3, 64.9% for D4 and 70.0% for D5, while precision is 97%, 93%, 58%, 83% and 76%, respectively. Errors concentrate between neighbouring grades, as expected from the gradual nature of damage (e.g. 12.8% of D2 buildings are predicted as D3 and 22.1% of D4 as D3). Confusion between non-adjacent classes (differing by two or more grades) is limited for D1 and D2 (at most 2.6%, for D1 predicted as D3 but part of the collapsed buildings is assigned to lower grades (23.2% of D5 to D3 and 4.2% to D2), so the most severe damage is occasionally underestimated.

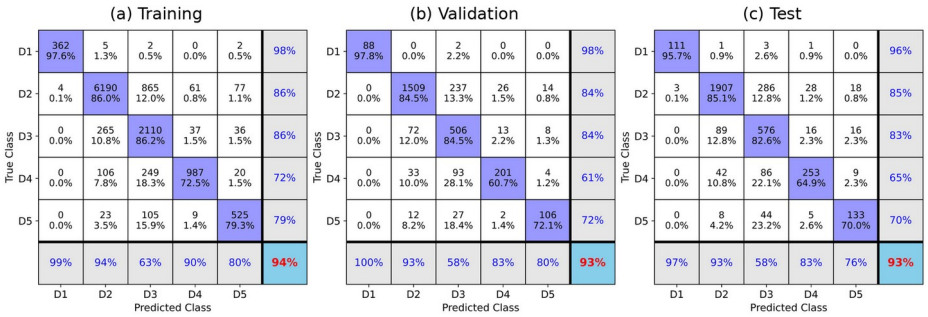


Fig. 2. Confusion matrices of the optimized network for the (a) training, (b) validation and (c) test subsets.

The ROC curves in Fig. 3(a) confirm the discriminative ability of the network, with AUC values of 1.00 (D1), 0.96 (D2), 0.93 (D3), 0.95 (D4) and 0.98 (D5). The lowest value corresponds to D3, whose responses overlap most with those of its neighbours.

5 Interpretation with SHAP

SHapley Additive exPlanations (SHAP) [21] attribute a prediction to the inputs through their marginal contribution over all feature combinations. Fig. 3(b) ranks the inputs by their mean absolute SHAP value over all samples and classes.

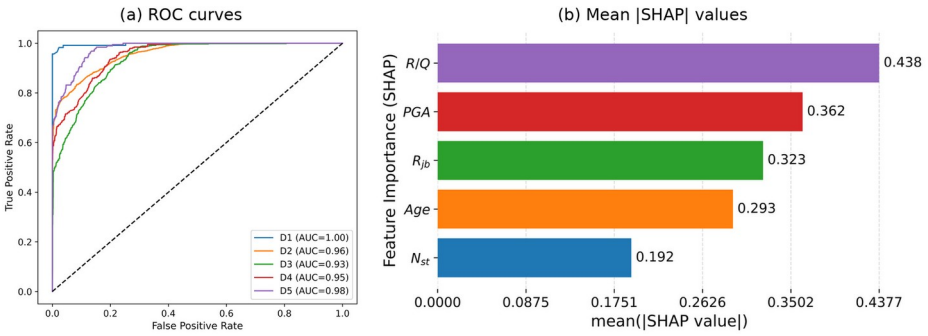


Fig. 3. Test-set results: (a) one-vs-rest ROC curves; (b) global feature importance given by the mean absolute SHAP value.

R/Q is the most influential input (0.438), followed by PGA (0.362), R_{jb} (0.323), age (0.293) and the number of storeys (0.192). The dominance of R/Q indicates that ductility and configuration quality are the main drivers of damage in this stock. R_{jb} adds attenuation and source-to-site information that PGA alone does not carry.The number of storeys matters least, which is plausible since most buildings have one or two storeys. The ranking agrees with the capacity–demand logic of seismic design.

6 Graphical User Interface

The model was embedded in a desktop GUI (Fig. 4). The input panel holds the five variables; the output panel shows the predicted damage state, its probability, a qualitative severity level and a bar chart of the five class probabilities. In the example of Fig. 4, a two-storey building is assigned to D2 with a probability of 41.3%, closely followed by D3 with 40.4%. The full distribution reveals the uncertainty between adjacent states that a single label would hide.



Fig. 4. Graphical user interface: input panel, predicted damage state and class probabilities.

7 Conclusions

A neural network classifier for post-earthquake building damage was developed from the Boumerdes survey using only five inputs describing structural capacity and seismic demand. After Bayesian optimization it reached an accuracy of 0.93 and a macro F1-score of 0.80 on unseen data, consistently across the three subsets. Errors mostly occur between adjacent states, but the recall of D4 and D5 and the precision of D3 remain weak points. SHAP showed that R/Q, PGA and R_{jb} dominate the predictions, in agreement with engineering understanding, and the GUI makes the model available for rapid screening with probabilistic outputs. The results rely on a single earthquake; future work should add site conditions and test transferability to other regions.

Acknowledgements. The authors thank the Earthquake Engineering Research Center (CGS), Algeria, for the survey data, and the Laboratory of Acoustics and Civil Engineering of Djilali Bounaama University of Khemis Miliana for its financial support.

References

1. Coburn, A., Spence, R.: Earthquake Protection. 2nd edn. John Wiley & Sons, Chichester (2002).
2. Rossetto, T., Elnashai, A.: Derivation of vulnerability functions for European-type RC structures based on observational data. Eng. Struct. 25(10), 1241–1263 (2003).
3. Roy, N.S., Choudhury, S.: Seismic vulnerability assessment methods: a review. Struct. Integr. 27, 282–300 (2022).
4. Mangalathu, S., Karthikeyan, K., Feng, D.C., Jeon, J.S.: Machine-learning interpretability techniques for seismic performance assessment of infrastructure systems. Eng. Struct. 250, 112883 (2022).
5. Mangalathu, S., Sun, H., Nweke, C.C., Yi, Z., Burton, H.V.: Classifying earthquake damage to buildings using machine learning. Earthq. Spectra 36(1), 183–208 (2020).
6. Allali, S.A., Abed, M., Mebarki, A.: Post-earthquake assessment of buildings damage using fuzzy logic. Eng. Struct. 166, 117–127 (2018).
7. Noura, H., Mebarki, A., Abed, M.: Post-quake structural damage evaluation by neural networks: theory and calibration. Eur. J. Environ. Civ. Eng. 23(6), 710–727 (2019).
8. Belbahi, H., Noura, H., Abed, M.: Bayesian-optimized machine learning classifiers for post-earthquake building damage prediction. Asian J. Civ. Eng., 1–17 (2025).
9. Laouami, N., Slimani, A., Bouhadad, Y., Nour, A., Larbes, S.: Analysis of strong ground motions recorded during the 21 May 2003 Boumerdes, Algeria earthquake. CSEM Newsl. (2003).
10. Mebarki, A., Boukri, M., Laribi, A., Farsi, M., Belazougui, M., et al.: Seismic vulnerability: theory and application to Algerian buildings. J. Seismol. 18(2), 331–345 (2014).
11. Boukri, M., Farsi, M.N., Mebarki, A., Belazougui, M.: Development of an integrated approach for Algerian building seismic damage assessment. Struct. Eng. Mech. 47(4), 471–493 (2013).
12. Scala, S.A., Del Gaudio, C., Verderame, G.M.: Influence of construction age on seismic vulnerability of masonry buildings damaged after 2009 L'Aquila earthquake. Soil Dyn. Earthq. Eng. 157, 107199 (2022).
13. Alothman, A., Mangalathu, S., Al-Mosawe, A., Alam, M.M., Allawi, A.: The influence of earthquake characteristics on the seismic performance of reinforced concrete buildings in Australia with varying heights. J. Build. Eng. 67, 105957 (2023).
14. DTR: Règles Parasismiques Algériennes, RPA99/Version 2003. Centre National de Recherche Appliquée en Génie Parasismique (CGS), Algeria (2003).
15. Noura, H., Abed, M., Belbahi, H.: Assessment of the effect of regular and irregular buildings on structural damage using neural networks. J. Build. Rehabil. 11, 147 (2026).
16. Hendrycks, D., Gimpel, K.: Gaussian error linear units (GELUs). arXiv:1606.08415 (2016).
17. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge (2016).
18. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167 (2015).
19. Bergstra, J., Bengio, Y.: Random search for hyper-parameter optimization. J. Mach. Learn. Res. 13, 281–305 (2012).
20. Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.: Optuna: a next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 2623–2631 (2019).
21. Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30, pp. 4765–4774 (2017).

ABSTRACT Track 2 · Structural Engineering Innovations

Structural Design of Mid-Rise RC Buildings in Abuja Subjected to Abuja-Specific Wind Loads

Ifeyinwa Ijeoma Obianyo[a*], Fisayo Praise Olatunji[a], Anthony Muoka[a], Sanusi Abdulganiyu[a], Abdulhameed Danjuma Mambo[a]

[a] *Department of Civil Engineering, Nile University of Nigeria, Abuja, Nigeria*
ifeyinwa.obianyo@nileuniversity.edu.ng

Wind loading is a major environmental action which affects the safety, stability and serviceability of buildings of reinforced concrete. Nigerian structural designs are frequently based on generalized or non-localized wind parameters which may not be a representative of the wind characteristics of specific sites, such as Abuja. In this study, the wind loading on a building structure was assessed, and its effect on the structural design and performance of an eight-storey reinforced concrete building structure in Abuja was evaluated using the provisions of Eurocode. The daily wind speed and wind direction data for Abuja for a period of 10 years (2015–2024) were obtained from the Nigerian Meteorological Agency (NiMet) and analyzed using descriptive statistical method and Gumbel Extreme Value Analysis to estimate the design wind speed for 50 years return period. A design wind speed of about 13 m/s was achieved for structural design. The design wind pressures were then calculated according to the Eurocode 1, and applied to a three-dimensional structural model created using ProtaStructures software. The model was analyzed and its structural response has been evaluated in terms of stability, lateral displacement, storey drift and reinforcement requirements. There was an increase in both the reinforcement requirement and costs by about 3.77% when wind loading was included in the model analyzed, compared with the model analyzed without wind loading. There was also a noticeable increase in roof displacement with the maximum displacement as 9.443 mm, with the maximum inter-storey drift between the 8^{th} and 9^{th} storey. This study indicates that the use of local wind parameters in the design of structures offers a more realistic wind effect and improves the safety and reliability of the reinforced concrete structures.

Keywords: *Wind loading, reinforced concrete building, Abuja, Eurocode, Gumbel Extreme Value Analysis, ProtaStructures, structural design.*

ABSTRACT Track 2 · Structural Engineering Innovations

Assessing the Accuracy of the Bi-Objective Optimization-Based SonReb Method for Nondestructive In-Situ Concrete Strength Estimation: A Real Case Study

Faiza Neggaz[1], Khoudja Ali-Benyahia[2] and Mohamed Ghrici[1]

[1] *Geomaterials Laboratory, Hassiba Benbouali University of Chlef, 02000 Chlef, Algeria*
[2] *Department of Civil Engineering and Architecture, University of Khemis-Miliana, Algeria*
neggazfaiza1@gmail.com

The compressive strength of concrete in existing structures is commonly estimated using non-destructive testing (NDT). When combined with destructive tests such as core sampling, NDT can provide a reliable assessment of in-situ concrete strength. The rebound hammer (RH) and ultrasonic pulse velocity (UPV) techniques are considered complementary and are often combined in an approach commonly known as the SonReb method to improve the accuracy of strength estimation. Several SonReb empirical models have been developed in the literature, generally based on linear regression. While some studies report that combining NDT techniques improves the accuracy of strength estimation, others reach opposite conclusions. In this work, the combined NDT approach is formulated and solved using a bi-objective optimization method, and its performance is compared with that of conventional linear regression. A thorough statistical analysis of a real case study from the literature is carried out to demonstrate the potential of bi-objective optimization for improving the SonReb method and to explain the divergence in researchers' conclusions regarding the effectiveness of the combined technique.

Keywords: *Concrete compressive strength, Non-destructive testing, SonReb method, Linear regression, Bi-objective optimization.*

TRACK 3

Emerging Trends and Technologies in Naval Architecture, Marine Engineering and Maritime Sciences

SPECIAL TRACK ORGANIZERS

Barış Barlas Piri Reis University, Türkiye

Gülçin Çivi Bilir Istanbul Technical University, Türkiye



KEYNOTE SPEAKER

Assist. Prof. Dr. Utku Cem Karabulut

Bandırma Onyedi Eylül University, Türkiye

Hydrodynamic Control and Diving Performance in the Conceptual Design of a Semi-submersible Naval Platform

Dr. Utku Cem Karabulut is an Assistant Professor at the Department of Naval Architecture and Marine Engineering, Faculty of Maritime Studies, Bandırma Onyedi Eylül University, Türkiye. He received his B.Sc. in Naval Architecture and Marine Engineering from Karadeniz Technical University in 2015, and his M.Sc. and Ph.D. from Istanbul Technical University in 2019 and 2026, respectively.

With 10 years of academic experience, his research focuses on ship hydrodynamics, design and optimization of marine vehicles, and Computational Fluid Dynamics (CFD). His recent scientific work includes full-scale ship resistance prediction for high-speed planing hulls, mesh structure effects on resistance prediction for marine vehicles, design prioritization for semi-submersible naval ships, and IDDES-based roughness characterization, with publications in international journals and conference proceedings.

TRACK TOPICS

- Autonomous Uncrewed Systems
- Green Shipping
- Green Shipyards and Green Production
- Ship Production and Ship Repair
- Maritime Security
- Maritime Operational Research
- Autonomous and AI Integration
- Vessel Decarbonization
- Alternative Marine Fuels
- Recent Technology Hull Coatings and Materials
- Acoustic Signature Reduction
- Underwater Drones
- Advanced Hull Materials
- Encryption in Underwater Communication Systems
- Circular Economy Applications in Maritime
- Lifecycle and Environmental Impact Management
- Logistics and Port Management
- Maritime Law and Governance
- Human Factors in Maritime Environments

The keynote lecture is included in this section as a full paper (see Table of Contents).

Hydrodynamic Control and Diving Performance in the Conceptual Design of a Semi-submersible Naval Platform

Utku Cem Karabulut[1][0000-0002-7588-0132] and Baris Barlas[2][0000-0002-5846-2369]

[1] Bandirma Onyedi Eylul University, Bandirma, Balıkesir 10200, TR
[2] Piri Reis University, Tuzla, İstanbul 34940, TR

Abstract. This study presents the conceptual design of a semi-submersible naval platform with a specific focus on hydrodynamic control and diving performance. Unlike conventional surface vessels or fully submerged submarines, this hybrid platform must maintain stability and maneuverability during both high-speed surface navigation and low-speed submerged operation. A mission-driven methodology is adopted, directly translating operational priorities into engineering decisions. The paper develops analytical models for two critical aspects often overlooked in early-stage hybrid design: (i) the force requirements for bow fins to achieve underwater equilibrium, solved via static force and moment balance using empirical lift coefficients, and (ii) the estimation of submergence time using a simplified differential equation model of ballast tank flooding. These models are applied to a 15.75 m, nine-personnel platform. Results indicate that the required downward force on the bow fins is 1.85 kN, and the platform can achieve full submergence within approximately 62 seconds under idealized conditions. The proposed approach provides a reproducible, control-oriented framework for conceptual design of semi-submersible naval platforms.

Keywords: Semi-Submersible Naval Platforms; Hydrodynamic Control Surfaces; Diving Performance; Conceptual Design Methodology.

1 Introduction

The increasing strategic importance of underwater operations has led to a growing interest in naval platforms capable of operating both at the surface and below the free surface (Karabulut et al. 2025). Considering that a significant portion of the Earth's surface is covered by oceans and that modern naval warfare increasingly relies on stealth, surveillance, and rapid response capabilities, platforms that combine the advantages of surface vessels and submarines offer a strategically significant and operationally flexible solution (Ma et al. 2025). Semi-submersible naval platforms, which can operate in both domains, represent a hybrid class of vehicles designed to achieve high surface speed while maintaining low detectability during submerged operations.

Historically, underwater vehicles have evolved from early conceptual designs to highly sophisticated submarines and special-purpose submersibles. However, fully

submerged platforms are generally limited in speed and maneuverability at the surface, while conventional high-speed vessels lack stealth capabilities and survivability in hostile environments (Wang et al. 2019). This dichotomy has motivated the development of hybrid platforms that integrate the hydrodynamic advantages of planing hulls with the stealth characteristics of underwater vehicles. Semi-submersible infiltration craft and special operations vehicles represent early examples of such systems, although their designs are often not systematically documented in the open literature (Matveev 2022).

Despite their operational potential, the number of comprehensive studies on the design of semisubmersible naval platforms remains limited. Existing literature primarily focuses on either high-speed surface craft hydrodynamics or underwater vehicle performance, with relatively few studies addressing the coupled design problem. Classical works on planing hull hydrodynamics provide fundamental insights into high-speed surface performance (Tavakoli et al. 2024), while submarine design methodologies emphasize buoyancy control, stability, and submerged resistance characteristics (Joubert 2004; Moonesun et al. 2016). However, these approaches are rarely integrated into a unified design framework suitable for hybrid platforms.

Recent advances in computational fluid dynamics (CFD) and experimental techniques have significantly improved the ability to analyze complex hydrodynamic phenomena associated with both surface and submerged operations. Studies focusing on free-surface effects, resistance prediction, and scale effects have demonstrated the importance of accurate modeling in predicting vessel performance. In particular, CFD-based approaches have been widely used to evaluate resistance, trim, and flow characteristics of high-speed vessels and underwater vehicles under various operating conditions. In this context, recent studies have investigated the hydrodynamic behavior of semi-submersible naval platforms and proposed novel methodologies for full-scale resistance prediction of high-speed planing hulls (Karabulut and Barlas 2025a, 2025b). These findings are complemented by experimental and numerical analyses on shallowly submerged underwater vehicles, which highlight the complex interaction between free surface effects and hydrodynamic forces (Karabulut and Barlas 2026).

Although these studies contribute significantly to the understanding of hydrodynamic performance, the conceptual design stage, where operational requirements are translated into a feasible platform configuration, remains insufficiently addressed. Ship design is inherently a multidisciplinary and iterative process in which multiple subsystems interact, and decisions made at early stages have a profound impact on overall performance (Papanikolaou 2014). Treating individual design elements in isolation may lead to suboptimal solutions, emphasizing the need for a holistic and system-oriented design approach (Papanikolaou et al. 2022).

A key challenge in the conceptual design of semi-submersible platforms lies in balancing conflicting requirements. High-speed surface performance demands lightweight structures and planing hull forms, whereas submerged operations require sufficient displacement, buoyancy control, and reduced hydrodynamic resistance. In addition, operational considerations such as maneuverability, endurance, payload capacity, and stealth impose further constraints on the design process. These competing requirements necessitate a structured methodology capable of prioritizing design

parameters and guiding decision-making. To address this challenge, the present study adopts a mission-driven design approach in which operational requirements are first identified and prioritized based on expert evaluations. A systematic prioritization study was conducted using inputs from experienced naval officers, revealing that high surface performance and maneuverability, along with low acoustic detectability during submerged operations, are among the most critical design drivers (Karabulut et al. 2024). These findings form the foundation for the subsequent conceptual design process.

Based on these prioritized requirements, a semi-submersible naval platform with an overall length of approximately 15 m and a crew capacity of nine personnel is developed. However, unlike conventional conceptual design studies that focus primarily on resistance and propulsion power, the present work specifically emphasizes two underdeveloped aspects critical for hybrid operation: (i) the force balance on hydrodynamic control surfaces (bow fins) required to maintain steady submerged equilibrium, and (ii) the estimation of submergence time using a simplified ballast tank flooding model. These aspects are often oversimplified or omitted in early-stage designs, yet they directly determine underwater controllability and mission responsiveness.

The primary objective of this study is to present a structured and reproducible conceptual design methodology for semi-submersible naval platforms, with a specific focus on hydrodynamic control and diving performance. The paper presents the design requirements, conceptual design methodology, control surface and ballast analyses, followed by discussion and conclusions.

1 Conceptual design methodology

The conceptual design of a semi-submersible naval platform is inherently multidisciplinary, requiring integration of hydrodynamics, naval architecture, propulsion, and operational constraints. In this study, a structured, mission-driven methodology is adopted to systematically translate prioritized operational requirements into a feasible platform configuration. The overall design approach is illustrated in Fig. 1.

The design requirements are established based on naval operations involving rapid deployment, covert approach, and short-duration submerged missions. Key performance criteria include high surface speed, enhanced maneuverability, low acoustic detectability during submerged operation, adequate payload capacity, and sufficient operational range. A comprehensive prioritization of these requirements was previously conducted using expert input from experienced naval personnel and applying the Fast Decision Method. The results indicated that surface performance parameters (maximum speed and maneuverability) are the most critical drivers, followed by low detectability and stability during submerged operations. The present study directly utilizes these previously established weighting factors to guide design

decisions, ensuring consistency between operational expectations and engineering implementation.

The methodology consists of successive stages: initial sizing, hull form development, general arrangement, propulsion system selection, and preliminary performance assessment. These stages are interconnected through iterative feedback loops rather than being strictly sequential. The process begins with initial sizing, where principal dimensions are estimated based on crew capacity (nine personnel), mission duration, and performance targets. Empirical relations and reference vessel data are used to maintain a balance between displacement, internal volume, and hydrodynamic efficiency.

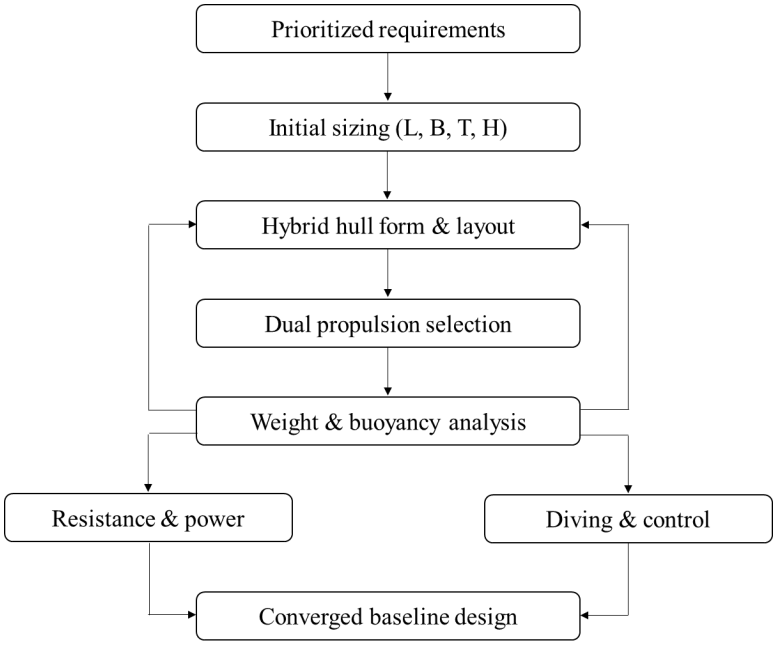


Fig. 1. Flowchart of the mission-driven, iterative conceptual design methodology for the semi-submersible naval platform.

The hull form is developed as a hybrid configuration combining planing characteristics for high-speed surface operation with streamlined underwater geometry to minimize submerged resistance. Lifting surfaces (bow and stern fins) are incorporated to enhance hydrodynamic control, particularly during diving and resurfacing maneuvers. The general arrangement ensures efficient spatial distribution of command, propulsion, and accommodation spaces, which is critical for weight distribution and center of gravity control.

For propulsion, a dual system is adopted: diesel-driven waterjets for surface navigation (high speed, maneuverability) and electric motors for submerged operation (low acoustic signature). The integration of these systems requires careful consideration

of power requirements, energy storage, and system compatibility within the available volume.

Preliminary performance assessments focus on two key aspects that are often oversimplified at the conceptual stage: (i) the force balance on bow fins to maintain steady submerged equilibrium, and (ii) the estimation of submergence time based on ballast tank flooding dynamics. These analyses are carried out using simplified analytical approaches (static equilibrium, empirical fin lift coefficients, and differential equations for ballast flow) appropriate for conceptual design. The outcomes are fed back into earlier design stages through iterative loops, allowing adjustments in dimensions, hull geometry, and system configuration. This holistic perspective ensures that the final design converges toward a balanced solution that meets both performance and operational criteria.

2 Conceptual design framework

Based on the prioritized operational requirements and the methodology presented in the previous section, a semi-submersible naval platform is developed. The conceptual design aims to achieve a balanced configuration capable of high-speed surface navigation and efficient submerged operation, while maintaining stability, maneuverability, and operational functionality.

2.1 Hull form

The computer-aided design (CAD) model of the proposed semi-submersible naval platform is depicted in Fig. 2. The hull form is developed as a hybrid configuration that integrates the geometric characteristics of high-speed planing craft with those of conventional underwater vehicles. This approach satisfies the fundamental design requirement of achieving both high surface performance and low-resistance submerged operation.

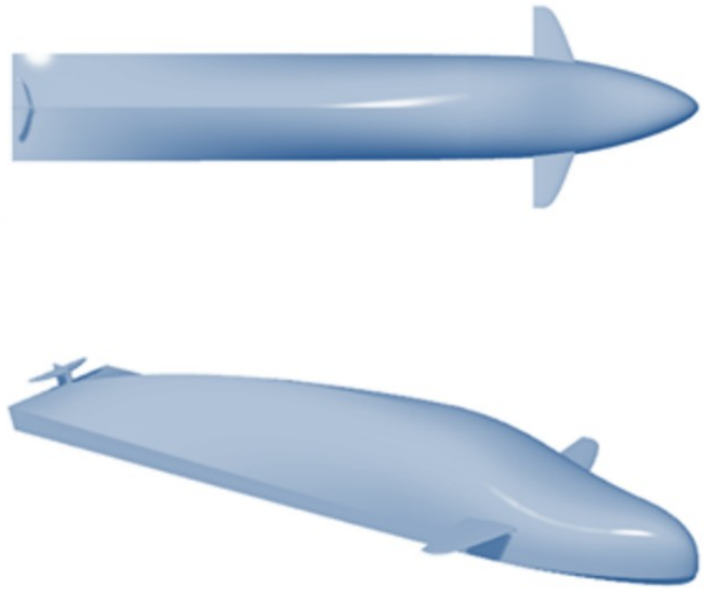


Fig. 2. CAD Plan of the semi-submersible naval platform.

Table 1 presents the principal dimensions and general characteristics. The overall length is 15.75 m, with a beam of 2.50 m and a draft of 0.78 m in surface condition. Design displacement is approximately 10.6 tons at the surface and increases to 57.0 tons during submerged operation due to ballast intake. These dimensions ensure adequate internal volume while maintaining favorable hydrodynamic characteristics.

Table 1. General design features of the platform.

Length overall	$L_{OA}\ (m)$	15.75
Length of waterline	$L_{WL}\ (m)$	14.78
Length between perpendiculars	$L_{BP}\ (m)$	14.78
Beam	$B\ (m)$	2.50
Draught	$T\ (m)$	0.78
Height	$H\ (m)$	3.00
Displacement (surface mode)	$\Delta\ (t)$	10.6
Displacement (underwater mode)	$\Delta\ (t)$	57.0
Wetted area (surface mode)	$S\ (m^2)$	35.8

Wetted area (underwater mode)	$S\ (m^2)$	117.1
Design speed (surface mode)	$V_A\ (knot)$	30.0
Design speed (underwater mode)	$V_S\ (knot)$	5.00
Longitudinal center of bouyancy (surface mode)	$LCB\ (m)$	5.202
Longitudinal center of bouyancy (underwater mode)	$LCB\ (m)$	7.750
Depth of submergence	$D\ (m)$	4.00

The Surface performance is governed by planing behavior. The hull geometry is shaped to generate sufficient hydrodynamic lift at high speeds (30 knots design speed), reducing the wetted surface area and consequently the resistance. For submerged operation (5 knots), the lower hull geometry is streamlined to minimize drag. This dual-characteristic design enables the platform to transition between operating modes without compromising overall performance.

To enhance hydrodynamic control and operational stability, lifting surfaces are incorporated at both the bow and stern. These control surfaces, designed as fins, play a critical role in regulating diving angle and maintaining course stability during submerged operation. Their contribution is particularly significant during transition phases requiring precise control of pitch and depth. The stern and bow fins of the semi-submersible naval platform is depicted in Fig.3.

The platform accommodates nine personnel and integrates mission, propulsion, and support systems within a compact layout. Diesel engines and waterjet units are positioned to support high-speed surface capability, while electric motors and batteries are located in the lower part of the hull for low-noise underwater propulsion. Ballast, trim, and compensation tanks are integrated to control buoyancy and maintain equilibrium during diving operations. Operational equipment such as a periscope and echo sounder are also incorporated. Overall, the hull form and general arrangement represent an integrated design solution addressing hydrodynamic performance, operational functionality, and system compatibility.

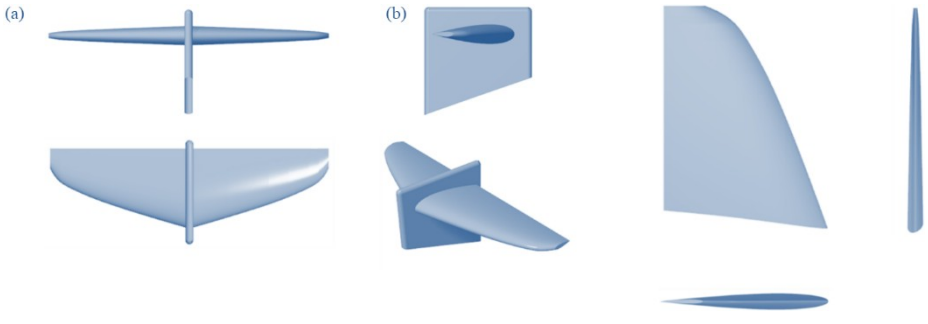


Fig. 3. Lifting surfaces to enhance hydrodynamic control and operational stability: (a) stern fins, (b) bow fins.

2.2 Control surfaces and ballast concept

Underwater operation of a semi-submersible platform requires active control of pitch, depth, and course stability, as well as a reliable means of submerging and surfacing. In the present design, control surfaces are implemented as lifting surfaces (fins) placed at both the bow and stern. These fins generate vertical forces that balance the hull's hydrodynamic pitch moment and maintain a steady submerged trim. The bow fins, in particular, are critical for counteracting the nose-down moment induced by the stern propellers during submerged transit. Their required lift force is determined from static equilibrium of the hull, as briefly discussed in the next section.

For diving and surfacing, a ballast system is integrated using dedicated tanks. When water is admitted through an orifice at the bottom of a ballast tank, the increased weight causes the platform to sink. The transient process from surface to fully submerged condition depends on the tank cross-sectional area, orifice diameter, and initial draft. A simplified dynamic model based on Bernoulli's equation and Newton's second law is used to estimate the submergence time under idealized conditions (no waves, constant tank area, negligible hull resistance during diving). This model is presented in the next section and provides a rapid conceptual-level assessment before more detailed dynamic simulations are undertaken. The combination of actively controlled bow/stern fins and a passive ballast system enables the platform to transition between surface and submerged modes with acceptable stability and responsiveness.

2.3 Performance Analysis

Bow Fin Force Requirement. To ensure both the propulsion force required for underwater operations and course stability, the propellers to be installed at the stern and the lifting surfaces to be installed at the bow must be designed by considering their hydrodynamic drag, lift, and torque. As a result of the CFD analysis, the calculated hydrodynamic resistance force $F_R = 2.89\ kN$, lift force $F_L = 3.9$, and torque $M = 5.19\ kN \cdot m$ were determined while traveling at a speed of 5 knots at a depth of 4 meters. Fig. 4 schematically illustrates the forces and moments acting on the vessel (free-body diagram).

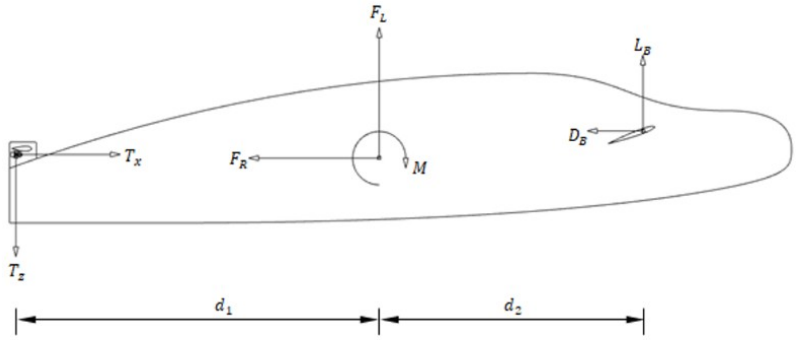


Fig. 4. Forces and moments acting on the hull.

F_R: Drag
F_L: Lift
M: Torque
T_x: Horizontal component of propulsive force
T_z: Vertical component of propulsive force
D_B: Drag on the bow fins
L_B: Lift generated by the bow fins
d_1: The horizontal distance between the center of gravity and the propellers
d_2: The vertical distance between the center of gravity and the propellers

Using these forces and moments, we can write the following static equilibrium equations:

$$\sum F_x = 0 \rightarrow T_x = F_R + D_B \tag{1}$$

$$\sum F_z = 0 \rightarrow T_z = F_L + L_B \tag{2}$$

$$\sum M = 0 \rightarrow M = T_z d_1 + L_B d_2 \tag{3}$$

In this form, there are 6 unknowns in the equations. Considering the design constraints and selecting $d1 = 7.3\ m$ and $d2 = 5.3\ m$, and since $D_B \ll F_R$, the resistance of the bow thruster can be neglected:

$$T_x = F_R = 2.89\ kN \tag{4}$$

$$T_z = F_L + L_B\ or\ T_z - L_B = 3.9\ kN \tag{5}$$

$$7.3T_z + 5,3L_B = 5.19\ kN \bullet m \tag{6}$$

Solving the equations yields:

$$T_x = 2.89\ kN \tag{7}$$

$$T_z = 2.05\ kN \tag{8}$$

$$L_B = -1.85\ kN \tag{9}$$

In this case, the propellers must generate a total thrust of $T = 3.55\ kN$, and the bow rudders must generate a downward force of 1.85 *kN*. The force generated by the rudder

can be calculated using the empirical equations presented by Whicker and Fehlner (1958). Accordingly,

$$C_{LS} = \frac{\partial C_{LS}}{\partial \beta} \frac{\beta}{57.3} + \frac{C_{D_C}}{AR_e} \left(\frac{\beta}{57.3} \right)^2 \tag{10}$$

Here, β is the angle of attack, $\frac{\partial C_{LS}}{\partial \beta}$ is the slope of the lift force curve, AR_e is the effective aspect ratio, and C_{D_C} is the drag coefficient. $\frac{\partial C_{LS}}{\partial \beta}$ is calculated from (Whicker and Fehlner 1958):

$$\frac{\partial C_{LS}}{\partial \beta} = \frac{1.8\pi AR_e}{1.8+\cos\Lambda\sqrt{4+\frac{AR_e{}^2}{\cos^4\Lambda}}} \tag{11}$$

Here, Λ represents the sweep angle. Depending on the lateral area (A_{LS}) and span (b), AR_e is defined as follows (Whicker and Fehlner 1958):

$$AR_e = \frac{2b^2}{A_{LS}} \tag{12}$$

Finally, the C_{D_C} coefficient is given as follows (Whicker and Fehlner 1958):

$$C_{D_C} = 0.1 + 1.6\lambda \tag{13}$$

where λ represents the conicity ratio. Conicity denotes the geometric taper or inclination of conical shell sections of the hull, particularly in the bow and stern regions. At the stern, this tapering defines how the hull narrows toward the aft end, contributing to reduced hydrodynamic drag and improved propulsive efficiency. A well-designed taper minimizes wake formation, supports a smooth flow attachment.

Diving time calculations. In addition to achieving steady-state equilibrium, the platform must also be capable of submerging within a reasonable timeframe. The following analysis provides an estimate of the diving time based on a simplified ballast model. Diving and surfacing velocities are also key considerations for semi-submersible naval platforms. This subsection outlines a practical and efficient approach for estimating diving speed under ideal conditions, assuming a simplified ballast tank geometry with a constant cross-sectional area.

Fig. 5 schematically illustrates a cylindrical tank on a floating vessel (Mungan and Lipscombe 2019). It has a flat bottom with an area of A_{boat} and vertical side walls of height H. The vessel contains a vertical tank (or compartment) of the same height H and cross-sectional area A_{tank}, and this tank is assumed to be horizontally centered within the boat. The zero time corresponds to the moment when water begins to enter through a hole at the center of the tank's base. The initial water levels outside the vessel and inside the tank are $h_1(0) < H$, and $h_2(0) = 0$, respectively, and the vessel is floating in a stationary state. That is, if the boat has a mass m_{boat} while the ballast tank is empty, the downward gravitational force must balance the upward buoyant force.

$$m_{boat} g = \rho g A_{boat} h_1(0) \tag{14}$$

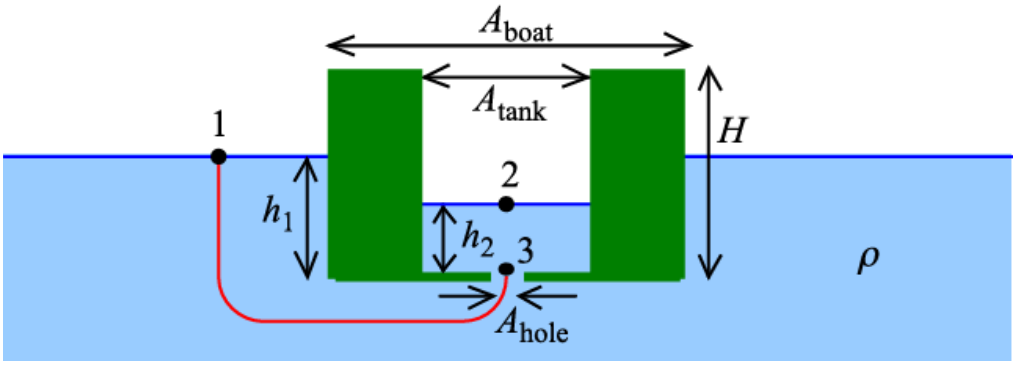


Fig. 5. Cross-section of a cylindrical tank on a floating boat (in green) with a hole in the center of its flat bottom (Mungan and Lipscombe 2019).

As the boat sinks at a rate of $\frac{dh_1}{dt} = v_1$ relative to the water surface outside the hull, the water inside the tank will rise at a rate of $\frac{dh_2}{dt} = v_2$ relative to the hull. Here, v_2 can be related to the velocity v_3 of the water jet passing through the hole using the continuity equation.

$$A_{hole}v_3 = A_{tank}v_2 \tag{15}$$

If the Bernoulli equation is written for points 1–3 (Fig. 5),

$$P_{atm} + \rho g h_1 = P_{atm} + \rho g h_2 + \tfrac{1}{2}\rho {v_3}^2 \tag{16}$$

If the P_{atm} terms on both sides of the equation are simplified and the v_3 term is eliminated using equation (15),

$$\rho g h_1 = +\rho g h_2 + \tfrac{1}{2}\rho \left(\frac{A_{tank}}{A_{hole}}\right)^2 {v_2}^2 \tag{17}$$

If the transformations $R = \left(\frac{A_{tank}}{A_{hole}}\right)^2$, $x = \frac{h_1}{h_1(0)}$, $y = \frac{h_2}{h_1(0)}$, and $T = \sqrt{\frac{2g}{h_1(0)}}t$ transformations are applied in sequence, the following differential equation is obtained (Mungan and Lipscombe 2019).

$$\left(\frac{dx}{dT}\right)^2 + x = R\left(\frac{dy}{dT}\right)^2 + y \tag{18}$$

This equation is a two-variable differential equation derived using the work-energy principle. A second equation for the variables x and y can be derived using Newton's second law of motion (Mungan and Lipscombe 2019).

$$\vec{F} = \frac{d}{dt}\left(m_{boat}\vec{v}_{boat} + m_{water\ inside}\vec{v}_{water\ inside}\right) \tag{19}$$

For the vertical component of the motion, this equation takes the following form:

$$\frac{d}{dt}\left[A_{boat}h_1(0)v_1 + A_{tank}h_2(v_1 - v_2)\right] = g\left[A_{boat}h_1(0) + A_{tank}h_2 - A_{tekne}h_1\right] \tag{20}$$

Let $r = \frac{A_{tank}}{A_{tekne}}$; this equation can be written as follows:

$$\frac{1}{g}\frac{d}{dt}\left[\frac{dh_1}{dt}+\frac{rh_2}{h_1(0)}\left(\frac{dh_1}{dt}-\frac{dh_2}{dt}\right)\right]=1+\frac{rh_2}{h_1(0)}-\frac{h_1}{h_1(0)} \tag{21}$$

If the transformations $x=\frac{h_1}{h_1(0)}$, $y=\frac{h_2}{h_1(0)}$, and $T=\sqrt{\frac{2g}{h_1(0)}}t$ are applied in sequence, the following differential equation is obtained (Mungan and Lipscombe 2019):

$$2\frac{d}{dT}\left[\frac{dx}{dT}+ry\left(\frac{dx}{dT}-\frac{dy}{dT}\right)\right]=1+ry-x \tag{22}$$

Equations (18) and (22) form a system of two-variable nonlinear differential equations, and the corresponding initial conditions are given below:

$$x(0)=1,\ y(0)=0,\quad \frac{dx}{dT}\Big|_{T=0}=0,\quad \frac{dy}{dT}\Big|_{T=0}=\sqrt{\frac{1}{R}} \tag{23}$$

This system of equations can be expressed as follows using the substitution method and can be integrated as a standard initial value problem:

$$\frac{d^2y}{dT^2}=\frac{1}{2R\frac{dy}{dT}}\left[2\frac{dx}{dT}\frac{d^2x}{dT^2}+\frac{dx}{dT}-\frac{dy}{dT}\right] \tag{24}$$

$$\frac{d^2x}{dT^2}=\frac{\left[1+ry-\frac{ry}{2R}-x-r\frac{dy}{dT}\frac{dx}{dT}+\frac{ry\frac{dx}{dT}}{2R\frac{dy}{dT}}\right]}{\left[(1+ry)-\frac{ry\frac{dx}{dT}}{R\frac{dy}{dT}}\right]} \tag{25}$$

Equations (24) and (25) were integrated using the Euler method with a time step of $\Delta T = 0.01$ s, and the results are presented in Fig. 6. The values $r = 0.75$ and $R = 2500$ were used in the calculations. The time-dependent change in the vessel's draft is shown in Fig. 7. According to this, the time required until the platform is completely submerged is 62 seconds.

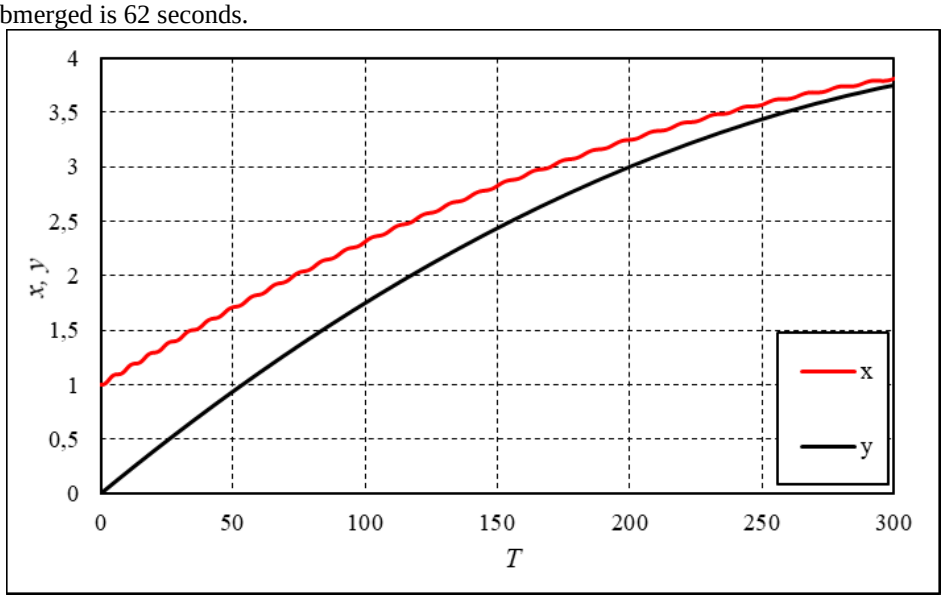


Fig. 6. The variation of dimensionless water levels inside and outside the platform as a function of dimensionless time.

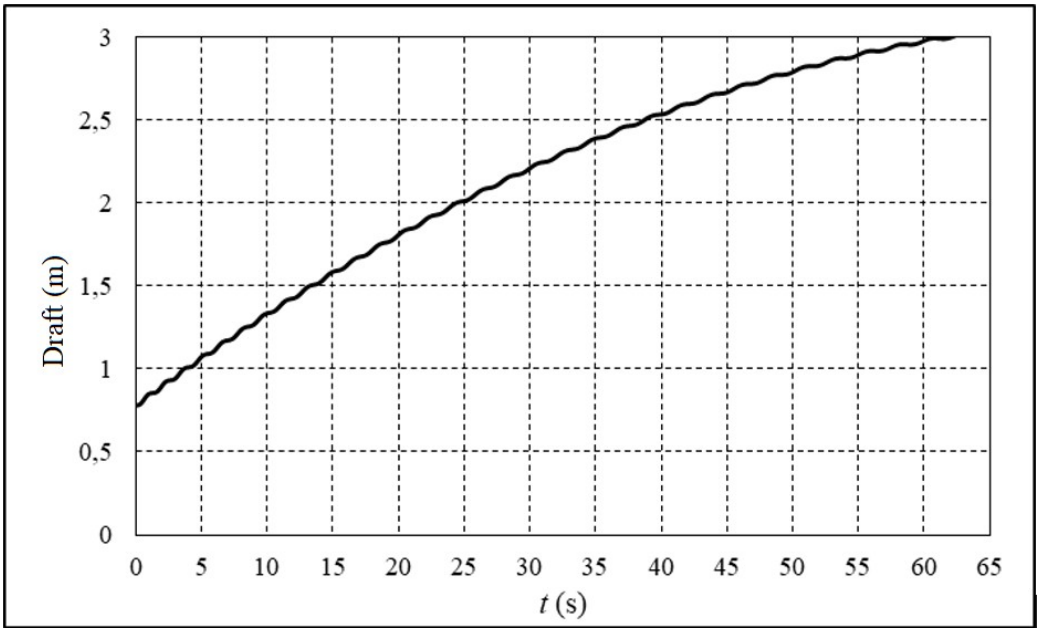


Fig. 7. Draft of the platform with respect to time.

The resulting platform configuration can therefore be summarized as follows: a 15.75 m hybrid semi-submersible platform equipped with dual diesel-waterjet propulsion for surface navigation and electric propulsion for submerged operation, capable of achieving approximately 30 knots at the surface and 5 knots underwater with a crew capacity of nine personnel.

4. Conclusions

This study presents the conceptual design of a semi-submersible naval platform, with a particular focus on hydrodynamic control and diving performance. A mission-driven, system-oriented methodology was adopted to translate prioritized operational requirements into a feasible engineering configuration.

The first key finding concerns the bow fins, which are essential for maintaining steady submerged equilibrium. Using static force and moment balance, together with empirical lift coefficients, the analysis shows that the fins must generate a downward force of 1.85 kN at a submerged speed of 5 knots, while the propellers provide a total thrust of 3.55 kN. This analytical approach offers a reproducible method for sizing control surfaces at the conceptual stage, addressing a gap in hybrid platform design where such force requirements are often overlooked.

The second contribution is the diving time estimation model. A simplified differential equation system based on ballast tank flooding, incorporating Bernoulli's equation and Newton's second law, was developed. Under idealized conditions (constant tank area, no wave action, negligible hull resistance during diving), the

calculated time to full submergence is approximately 62 seconds. This model enables rapid assessment of ballast tank sizing and orifice diameter before more detailed dynamic simulations are undertaken, thus supporting early-stage decision-making.

The second contribution is a diving time estimation model. A simplified system of differential equations based on ballast tank flooding (incorporating Bernoulli's equation and Newton's second law) was developed. Under idealized conditions (constant tank cross-sectional area, no wave action, and negligible hull resistance during diving), the calculated time to full submergence is approximately 62 seconds. This model enables rapid assessment of ballast tank sizing and orifice diameter prior to more detailed dynamic simulations, thereby supporting early-stage design decisions.

The third outcome relates to the integrated methodology itself. The proposed framework successfully combines hybrid hull form development, dual propulsion (diesel–waterjet for surface operation and electric propulsion for submerged operation), and weight distribution analysis. The general arrangement and center-of-gravity calculations confirm favorable stability characteristics for the 15.75 m platform designed for nine personnel, demonstrating that a balanced configuration can be achieved despite the conflicting requirements of surface and submerged operation.

Several limitations must be acknowledged. The analyses are based on simplified methods suitable only for conceptual design. High-fidelity CFD, experimental validation, structural design, seakeeping in rough sea states, and acoustic signature analysis are required for detailed design phases. Future work should also integrate active control systems for pitch and depth regulation, as well as optimize energy storage for extended submerged endurance.

Overall, this study contributes a structured, control-oriented framework for the conceptual design of semi-submersible naval platforms, addressing two critical but often overlooked aspects: bow fin force requirements and diving time estimation. The proposed methodology can serve as a foundation for further research and development of advanced hybrid naval systems.

References

Denny SB, Feller AR. 1979. *Waterjet Propulsor Performance Prediction in Planing Craft Applications*, Technical Report No: DTNSRDC SPD/0905-01, https://apps.dtic.mil/sti/citations/ADA075690

Joubert PN. 2004. Some aspects of submarine design: part 1: Hydrodynamics (U), Technical Report No: DSTO-TR-1622, *Australian Government Department of Defence*. http://www.dsto.defence.gov.au/ corporate/reports/DSTO-TR-1622.pdf

Karabulut UC, Barlas B. 2025a. Computational analysis on the hydrodynamics of a semisubmersible naval ship. *Journal of Marine Science and Application*, *24*(2), 331-344.

Karabulut U, Barlas B. 2025b. A novel approach to full-scale ship resistance predictions: experimental and numerical study of a high-speed planing hull. *Journal of Marine Science and Technology*, *30*(4), 939-963.

Karabulut UC, Barlas B. 2026. Experimental and Numerical Analyses of the Hydrodynamics of a Shallowly Submerged Underwater Vehicle. *IEEE Journal of Oceanic Engineering*.

Karabulut UC, Barlas B, Baykal, MA. 2024. Design prioritization study for a semi-submersible naval ship based on fast decision method. *Journal of Naval Sciences and Engineering*, *20*(1), 3-19.

Karabulut UC, Barlas B., Helvacıoğlu Ş, Baykal MA. 2025. Su altı ve su üstünde seyir yapabilen hibrit bir otonom deniz platformunun su altı hidrodinamiğinin incelenmesi. *Akıllı Ulaşım Sistemleri ve Uygulamaları Dergisi*, *8*(2), 197-216.

Ma D, Li Y, Ma T, Pascoal AM. 2025. The state of the art in key technologies for autonomous underwater vehicles: A review. *Engineering*, https://doi.org/10.1016/j.eng.2025.08.002.

Matveev KI. 2022. Hydrodynamic Modeling of Fast Semi-Submersible Vehicle in Deep and Shallow Water. In *Proceedings of OCEANS Conference 2022, Hampton Roads, VA, USA* (pp. 1-6). DOI: 10.1109/OCEANS47191.2022.9977281

Moonesun M, Mahdian A, Korol YM, Dadkhah M, Javadi, MM. 2016. Concepts in submarine shape design. *Indian Journal of Geo-Marine Sciences*, *45*(1), 100-104.

Mungan CE, Lipscombe TC. 2019. How quickly can a submarine dive? *The Physics Educator*, *1*(04), 1950013.

Papanikolaou A. 2014. *Ship design: methodologies of preliminary design* (Vol. 9). Berlin/Heidelberg, Germany: Springer.

Papanikolaou A, Harries S, Hooijmans P. et al. 2022. A holistic approach to ship design: Tools and applications. *Journal of Ship Research*, *66*(01), 25-53.

Tavakoli S, Zhang M, Kondratenko AA, Hirdaris S. 2024. A review on the hydrodynamics of planing hulls. *Ocean Engineering*, *303*, 117046.

Wang Z, Liu X, Huang H, Chen Y. 2019. Development of an autonomous underwater helicopter with high maneuverability. *Applied Sciences*, *9*(19), 4072.

Whicker LF, Fehlner LF. 1958. *Free-stream characteristics of a family of low-aspect-ratio, all-movable control surfaces for application to ship design*, DTMB Technical Report No: 933, https://apps.dtic.mil/sti/citations/ADA014272.

AN INVESTIGATION OF FORCES AND MOMENTS ACTING ON SEMI-STANDARDIZED AND NON-STANDARDIZED CARGO UNITS ABOARD SHIPS

Hakan Akyıldız[1] and Fatih Tuncay Kolçak[2]

[1]Department of Shipbuilding, Maritime Higher Vocational School, Piri Reis University, Istanbul, Türkiye; hakyildiz@pirireis.edu.tr
[2]Department of Shipbuilding and Ocean Engineering, Istanbul Technical University, Istanbul, Türkiye; kolcak@itu.edu.tr

Abstract. This study investigates forces and moments acting on semi-standardized and non-standardized cargoes transported on the weather decks of vessels. These cargo types, often irregular in geometry and variable in securing requirements, are subject to a combination of gravitational, inertial, aerodynamic, and hydrodynamic forces under varying severe weather and sea conditions. Lashing forces are derived from cargo accelerations resulting from ship motions. The most intense forces, driven by peak accelerations, typically occur at the ship's forward, rear, and uppermost stowage areas on both port and starboard sides. The inadequacy of conventional securing guidelines for such cargoes frequently necessitates reliance on crew judgment, introducing inconsistencies in safety margins. This study presents analytical methods, as described in the CSS Code Annex 13, to evaluate the stability assessments of transverse sliding, transverse tipping and longitudinal sliding in terms of the balance of forces and moments. Two case studies are presented using same lashing configurations to compare outcomes from the balance of forces and moments approach and the method horizontal securing angles involved. The results reveal that lashings with large horizontal and vertical angles contribute minimally to securing efficiency unless they are pretensioned and continuously monitored. This study also recommends improvements in simulation accuracy, standardization of friction-enhancing materials, automation in lashing applications, and the future integration of AI-based decision-support tools to enhance the safety and reliability of securing semi- standardized and non- standardized cargoes

Keywords: Semi-standardized cargo, non-standardized cargo, CSS code, Cargo securing devices, Transverse sliding, Transverse tipping, Longitudinal sliding, Lashing Forces

1 Introduction

The sheer volume of goods crossing oceans each day attests to maritime shipping's foundational role in global trade architecture, transforming ports into pulsing gateways of commerce and the sea into a vital economic corridor. In 2023, maritime cargo

transport constitutes approximately 80% of global trade by volume, highlighting its critical and enduring role in sustaining international supply chains. United Nations Conference on Trade and Development (UNCTAD) forecasts that, after 2024, maritime trade will grow at an average annual rate of 2.4% from 2025 to 2029, with containerized trade expected to expand at a slightly higher rate of 2.7% annually [1].

During maritime transportation on the weather deck of ocean-going vessels, cargoes are exposed to a range of interacting forces. These include the gravitational force due to their own mass, the frictional resistance between the cargo and the deck surface or any dunnage material, additional buoyant forces resulting from seawater intermittently washing over the deck, aerodynamic forces such as wind pressure, hydrodynamic forces arising from wave impact, and inertial forces induced by vessel motions in turbulent sea states.

The detrimental effects of severe weather conditions on deck cargo can manifest in the following ways:

- In turbulent sea conditions, unsecured or inadequately stowed deck cargo is highly susceptible to displacement by powerful wave forces or gusty winds, resulting in losses overboard.
- When cargo experiences physical damage, whether from impact or shifting, it can lead to the contamination of neighboring cargoes.
- Cargo exposed to prolonged atmospheric exposure or sea spray may suffer from water ingress into the cargo holds or stowage spaces.
- Intense ship motion in rough seas can cause cargo to collapse, break, shift, or even fall from the deck, leading to structural failure of the cargo and potential damage to the ship’s own infrastructure [2].

Excessive loads on deck cargo during wave encounters - arising from ship motions, slamming, whipping effects, green water overtopping, and direct wave impacts - must be carefully considered when evaluating the safety [3]. The movement of the ship generates inertial forces, wind loads, and splash impacts, all of which act upon the cargo unit. Among these, the inertial force plays a dominant role in initiating the displacement of the cargo [4]. Precise understanding of the acceleration distribution throughout the vessel would enhance the safety of container securing system designs and help prevent cargo displacement [5]. In conditions where substantial wave action leads to the partial or complete immersion of the ship’s deck, transient yet considerable buoyant forces may exert upward pressure on the stowed cargo. This phenomenon is particularly pronounced in containers or units featuring sealed or air-filled voids, which enhance their buoyancy. As a result, the frictional interface between the cargo and the deck can be significantly diminished, potentially undermining the stability of the cargo and impairing the reliability of the securing arrangements [6]. Repeated exposure to fluctuating wave loads can weaken cargo securing arrangements and structural elements due to fatigue, especially over long voyages [7].

Large rolling motion and high accelerations often cause cargo to shift on board, leading to a range of problems or even serious stability accidents, exacerbated by the harmful effects of waves and wind. As a result, it is crucial to incorporate both inertia and external forces, including wind pressure and sea sloshing, into the calculations [8]. According to statistical data, approximately 11% of severe stability accidents in vessels

are primarily triggered by the shifting of cargo and/or the impactful forces of waves and wind, with these types of incidents being particularly prevalent and consequential on bulk carriers and ro-ro vessels [9]. Although, the magnitude of the wind force exerted on deck cargo is highly dependent on the geometrically prominent exposed frontal area of the cargo and the dynamically fluctuating intensity - i.e., the gustiness - of the ambient wind, both of which significantly influence aerodynamic loading in maritime environments. The precise determination of wind-induced forces and moments presents a significant challenge, primarily because of its far-reaching consequences for numerous critical analyses, such as assessing ship stability, estimating ship speed, evaluating maneuvering capabilities, and ensuring effective station-keeping and mooring operations [10].

To ensure stability during maritime transport, especially under abrupt and severe weather phenomena such as intense winds and sudden wave action, it is essential that both vehicles and cargo - regardless of whether they are positioned on deck or in the cargo hold - are rigorously secured. Despite this necessity, current regulatory documents provide only limited provisions for the stowage of irregular or non-standardized cargo types. Consequently, in operational practice, securing procedures are predominantly informed by the subjective judgment and experiential knowledge of the crew, often in the absence of formalized verification protocols. This approach frequently leads to inconsistent restraint applications, characterized by either over-securing or insufficient stabilization, ultimately compromising cargo safety and operational efficiency [11]. Depending on the coefficient of friction between the cargo and the ship, inadequately secured cargo may slide or tip over. Beyond that, it could even lead to the sinking of the ship [12].

By utilizing sea state spectra specific to a given region, it is possible to generate time series data representing the nonlinear motions of the ship, which can then be applied in simulation analyses. These resulting time histories of lashing forces can be condensed into load spectra, enabling fatigue assessments of the lashing components while incorporating the statistical variability of sea state conditions [13]. The coefficient of friction between the cargo and the ship is critically important, as it can significantly reduce the maximum lashing loads expected from the securing systems. Inadequate friction often results in cargo sliding, which can subsequently lead to inconsistent and uneven forces within the securing elements. Achieving sufficiently high coefficient of friction to prevent sliding depends substantially on several factors, including the amplitude of roll and pitch motions, the pitch period, the orientation of the cargo on the deck, and the vertical distance of the cargo from the ship’s baseline [14].

The regulatory framework governing the carriage of semi-standardized and non-standardized cargo on ship decks remains a dynamically evolving field, intricately influenced by international conventions, national regulations, and classification society guidelines. Predominantly, instruments such as the International Maritime Organization’s (IMO) International Convention for the Safety of Life at Sea (SOLAS) and the Code of Safe Practice for Cargo Stowage and Securing (CSS Code) meticulously stipulate essential requirements. In 1991, the IMO implemented CSS Code through Resolution A.714(17). It is also recommended that the code be implemented by governments without delay. The task of monitoring the

implementation of the CSS Code and enhancing it based on accumulated experience was entrusted to the Maritime Safety Committee (MSC). The CSS Code evolved into its present form through the following historical revisions: MSC/Circ.664, MSC/Circ.691, MSC/Circ.740, MSC/Circ.812, MSC/Circ.1026, MSC.1/Circ.1352, MSC.1/Circ.1352/Rev.1, and MSC.1/Circ.1623 [15]. Industry goals emphatically prioritize enhancing operational safety, promoting environmental sustainability, and standardizing securing methods without stifling the flexibility necessary for diverse cargo types [16]. Semi-standardized cargo, exemplified by modular industrial components, occupies an intermediate classification, whereas non-standardized cargo, such as large machinery or bespoke structures, demands highly individualized securing strategies [17]. Classification societies like Lloyd's Register and Bureau Veritas continuously refine their notations and rules to accommodate these distinctions, seeking greater harmonization. Looking forward, the industry aims to integrate advanced simulation technologies, real-time monitoring systems, and AI-driven risk assessments into regulatory practices [18]. The future direction necessarily involves updating design and strength criteria for deck fittings, lashings, and cargo units to reflect the increasing complexity and variability of transported items. Nevertheless, extensive research is urgently needed to better understand dynamic load interactions, fatigue behavior under non-linear conditions, and the resilience of securing systems during extreme weather events. Recent developments have seen the emergence of lashing software tools [19]. Bridging these gaps will undoubtedly require collaborative international initiatives and innovative experimental methodologies. Ultimately, a more robust and adaptive regulatory landscape is indispensable for ensuring the safe, efficient, and environmentally responsible transport of these complex cargoes across global oceans.

In this study, some cargo lashing units are defined and the forces and moments that may occur on the cargo during the cruise are examined. Sample calculations are given about the adequacy of cargo lashing systems according to the forces acting on the cargo units. In addition, the requirements for cargo lashing have been examined within the framework of international rules. Since it is not possible to calculate the maximum forces and moments that will affect the cargo units in the worst sea conditions, the basic rule should be to design a cargo lashing system that will provide sufficient strength. There are different loads other than the load units examined here and there may be different forces and moments that can act on these cargo loads. Therefore, it will be understood that different examinations should be carried out both based on international rules and when calculations are considered.

2 Safety Requirements of Cargo Units and Lashing Systems

The 1974 SOLAS Convention (Regulations VI/5 and VII/5) decreed that cargo units and cargo transport units be secured, stowed, and loaded throughout the voyage in accordance with a Cargo Securing Manual (CSM), which is a ship-specific document providing detailed instructions and procedures for the safe stowage and securing of cargo during sea transport. Containerships must have a Cargo Safe Access Plan (CSAP) that ensures personnel can safely access areas for securing containers [20]. The

information presented in this paper will be highly beneficial in the preparation and updating of these two documents.

During the design phase for containers stowed on deck, CSAP should be developed, considering the factors listed below:

- Risks associated with slips, trips, and falls must be considered.
- The danger of falling from elevated positions should be evaluated.
- Potential injuries resulting from manual handling of lashing equipment must be addressed.
- The risk of being struck by falling lashing gear or other objects needs to be assessed.
- The likelihood of damage caused by container handling operations must be identified, and high-risk zones should be clearly defined to implement suitable protective measures or prevention strategies.
- Possible hazards from nearby electrical sources, such as connections for refrigerated container units, should be considered.
- The sufficiency of safe access to all areas required for the execution of container securing tasks must be ensured.
- The ergonomics related to the handling of lashing gear, including the size and weight of the equipment, should be evaluated to minimize injury risks.
- Special attention should be given to the challenges and safety implications of securing 2.9 meters high containers or higher, as well as mixed stowage configurations involving 12.2 m and 13.7 m containers [21].

Devices used to fasten cargo on ships must meet specific strength requirements, identified by the term Maximum Securing Load (MSL), which represents their upper load capacity. In place of MSL, the Safe Working Load (SWL) can also be applied, but only if it meets or surpasses the load strength specified by the MSL. The calculation of forces acting on securing elements is a complex problem involving the elasticity, angles, lengths, and pre-tension of each individual element. However, this complexity is simplified by using the term Calculated Strength (CS), which is derived from the MSL reduced by an appropriate safety factor, as expressed in the following manner [22]:

$$CS = \frac{MSL}{safety\ factor} \tag{2.1}$$

During the manufacturing process, permanently fixed securing devices shall be subjected to a test load equivalent to 1.25 times the MSL for a minimum duration of one minute. This load shall be applied under vertical securing angles of 10°, 45°, and 80°, and horizontal securing angles ranging from 0° to 360°, without exceeding 45° intervals. The resulting deformations must not exceed the following limits:

- For socket type circular securing fittings, maximum deflection shall not exceed 3 mm.
- For D-ring and crinkle bar type fittings, maximum elongation in any direction shall not exceed 2%.
- The maximum indentation shall be limited to 1.5 mm.
- No permanent deformation shall be present that could impair cargo operations [23].

SWL of fixed cargo securing devices, shall be included in the cargo stowage plan, expressed in tonnes, as a design parameter. Such devices are welded directly to hull with adequate local structural reinforcements. Formula to calculate design force as a multiple of SWL, representing both static and dynamic load conditions as follows:

$$F_{SWL} = 1.25\ x\ SWL\ x\ g \tag{2.2}$$

For RO/RO vessels, MSL assigned to each individual lashing point on board shall not be less than any of the following criteria:

- $k.Q.g$
- $0.5\ x\ P_m$
- 100 kN for heavy cargo
- 15 kN for car only decks

where

k : $\frac{n}{r}$ when $r \neq 1$, k multiplied with 1.1

P_m: breaking strength of cargo securing device

r : number of usable lashings on either side of the vehicle, relative to the axle group of size n

Q : maximum axle load given in tons [24].

As another semi-standardized cargo, timber, when carried on deck, must comply with specific rules to prevent damage to the vessel or the cargo itself. The following precautions must be observed:

- Timber deck cargo should be stowed fore and aft whenever possible.
- A minimum of two transverse lashings is required per package at the ship's side, with spacing set at 3 m for cargo up to 4 m high and 1.5 m for taller stowage.
- Web lashings may be utilized; however, fabricated web lashings shall not be used together with chains or wires.
- Chains or wire rope lashings should have a minimum breaking strength of 133 kN.
- Daily checks and adjustments of lashing and securing systems are essential while at sea. Following partial discharge at intermediate ports, the residual cargo should be leveled and secured again as required by the regulatory Codes.
- When an eye is formed with four grips around a thimble, it must retain at least 90% of the wire rope's Minimum Breaking Load (MBL). Without a thimble, the holding power drops to approximately 70%
- In winter zones, the maximum height of timber deck cargo must not exceed one third of the vessel's overall breadth [25].

For portable tanks, vertical angles used to counteract sliding must remain below 25°, whereas those resisting tipping should be set between 45° and 60° [26].

The strategic placement of timber dunnage in proximity to securing elements contributes to cargo stability: when positioned near vertically oriented lashings, it effectively aids in counteracting tipping moments, whereas its use adjacent to horizontally aligned lashings enhances resistance against sliding forces. Standard dunnage timbers are typically limited to rough planks measuring 150 × 25 mm, but heavier cargo requires more substantial support. In such cases, the timber should be no less than 50 mm thick and 150 mm wide, with a preferred dimension of 75 × 225 mm.

When utilizing 75 mm thick dunnage, it is crucial to select straight-grained wood of the greatest possible width and to ensure that the grain is oriented horizontally and parallel to the deck to maintain structural effectiveness. With timber dunnage, a friction coefficient of 0.3 results in sliding at around 17°, while a coefficient of 0.7 increases the sliding threshold to 35°. Steel-on-steel contact, however, may lead to sliding at inclinations as low as 6° [27].

3 Methodology of Forces and Moments Acting on Cargo Units

In the completely revised Annex 13 of the CSS Code (MSC.1/Circ.1623), two methodologies are described for calculating the safety of loads acting on transported cargoes and securing systems. The first is the Rule-of-Thumb method, which employs a simplified and approximate approach based on basic logic like the combined maximum securing load of securing devices on each side of a cargo should match its weight. The second is referred to as the Advanced Calculation Method, which provides a more precise and comprehensive analytical framework [22].

3.1 Advanced Calculation Method

External forces acting on the cargo item are resolved along three axes and are expressed by the following formula:

$$F_{(x,\,y,\,z)} = m.a_{(x,\,y,\,z)} + F_{w(x,\,y)} + F_{s(x,\,y)} \tag{3.1}$$

where

$F_{(x,\,y,\,z)}$: respectively longitudinal, transverse and vertical forces
m : mass of the cargo item
$a_{(x,\,y,\,z)}$: longitudinal, transverse and vertical accelerations
$F_{w(x,\,y)}$: longitudinal and transverse forces exerted by wind
$F_{s(x,\,y)}$: longitudinal and transverse forces exerted by sea

The acceleration values to be used in the formula are provided below in Figure 1, based on specific assumptions in Table 1:

Table 1. Assumptions for acceleration values.

Parameter	Value
Operational Area	Unrestricted
Operational Period	Year-round
Ship Length	100 m
Service Speed	15 knots
Stability Ratio (B/GM)	≥ 13

For every condition beyond the values presented in Table 1, the accelerations provided in Figure 1 must be adjusted by multiplying them with a suitable correction factor to ensure accurate revision.

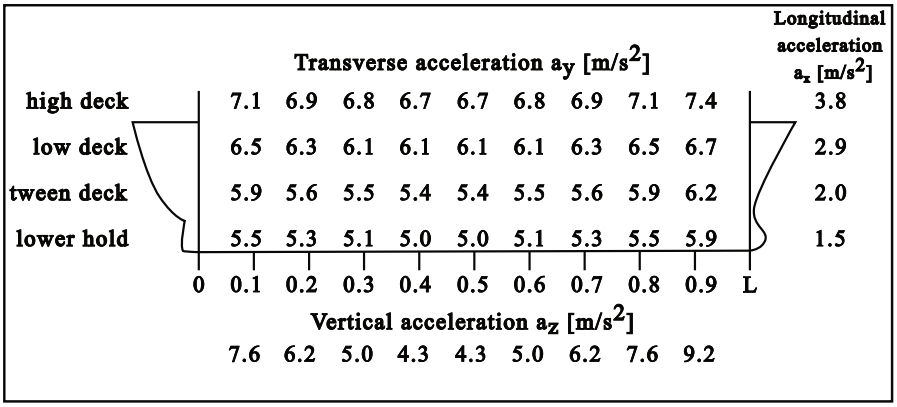


Fig. 1. Longitudinal, transverse and vertical accelerations.

For operations within a restricted area, weather forecasts indicating significant wave height are required to calculate reduction factor defined below:

$$f_R = 1 - (H_S - 13)^2 / 240 \qquad (3.2)$$

For ships with a length differing from 100 meters and a service speed other than 15 knots, the correction factors to be applied to the acceleration values are provided in Table 2.

Table 2. Correction factors for length and speed.

Length (m) / Speed (kn)	50	60	70	80	90	100	120	140	160	180	200
9	1.20	1.09	1.00	0.92	0.85	0.79	0.70	0.63	0.57	0.53	0.49
12	1.34	1.22	1.12	1.03	0.96	0.90	0.79	0.72	0.65	0.60	0.56
15	1.49	1.36	1.24	1.15	1.07	1.00	0.89	0.80	0.73	0.68	0.63
18	1.64	1.49	1.37	1.27	1.18	1.10	0.98	0.89	0.82	0.76	0.71
21	1.78	1.62	1.49	1.38	1.29	1.21	1.08	0.98	0.90	0.83	0.78
24	1.93	1.76	1.62	1.50	1.40	1.31	1.17	1.07	0.98	0.91	0.85

For vessels with a length between perpendiculars ranging from 50 to 300 meters, which are not listed in the table, the calculation shall be performed using the following formula:

$$c.f. = \left(0.345\frac{v}{\sqrt{L}}\right) + \left(\frac{58.62\ L-1034.5}{L^2}\right) \tag{3.3}$$

For vessels with B/GM ratio less than 13, the lateral acceleration values should be multiplied by the correction factors provided in Table 3.

Table 3. Correction factors for B/GM ratio.

B/GM	3	4	5	6	7	8	9	10	11	12	≥ 13
high deck	2.64	2.28	1.98	1.74	1.56	1.40	1.27	1.19	1.11	1.05	1.00
low deck	2.18	1.93	1.72	1.55	1.42	1.30	1.21	1.14	1.09	1.04	1.00
tween deck	1.62	1.51	1.41	1.33	1.26	1.19	1.14	1.09	1.06	1.03	1.00
lower hold	1.24	1.23	1.20	1.18	1.15	1.12	1.09	1.06	1.04	1.02	1.00

Unless a specific condition is stated, a simplified approach may be adopted whereby the wind pressure and sea sloshing effect acting per square meter can each be considered as 1 kN [22].

3.2 Balance of forces and moments

Stability assessments are conducted for transverse sliding, transverse tipping, and longitudinal sliding. In cases of symmetrical lashing arrangements, it is sufficient to perform calculations for either the port or starboard side, or for either the fore or aft direction. The most effective measure against sliding is to increase the coefficient of friction between the cargo and the surface on which it is transported. The frictional forces between various material pairings are provided in Table 4. In practical operations, it must be noted that the values in the table can vary significantly due to factors such as dust, frost, moisture, ice, greasy dirt, or snow.

Table 4. Friction coefficient between material pairs.

Material pair	**Friction Coefficient (μ)**
timber - timber (wet or dry)	0.4
steel - timber, steel - rubber	0.3
steel - steel (dry)	0.1
steel - steel (wet)	0.0

Transverse sliding

The balance equation is expressed as follows:

$$F_y \leq \mu.m.g + CS_1.f_1 + CS_2.f_2 + \ \ldots\ldots\ + CS_n.f_n \tag{3.4}$$

where

F_y : force that transverse lashing must withstand [kN] (Figure 3).

μ : friction coefficient
m : mass of the cargo item [t]
g : gravitational acceleration = 9.81 [$\frac{m}{s^2}$]
CS: calculated strength = $\frac{MSL}{1.5}$ [kN]
f : referencing Table 5, depends on μ and vertical securing angle (α)

$$f = \mu.sin\alpha + cos\alpha$$

n : number of lashings

Table 5. f values for corresponding α and μ.

α / μ	-30°	-20°	-10°	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°
0.3	0.72	0.84	0.93	1.00	1.04	1.04	1.02	0.96	0.87	0.76	0.62	0.47	0.30
0.1	0.82	0.91	0.97	1.00	1.00	0.97	0.92	0.83	0.72	0.59	0.44	0.27	0.10
0.0	0.87	0.94	0.98	1.00	0.98	0.94	0.87	0.77	0.64	0.50	0.34	0.17	0.00

Lashings with a vertical securing angle greater than 60° or a horizontal securing angle exceeding 30° provide a negligible contribution to the overall securing effectiveness and should therefore be excluded from force calculations. Their effectiveness may only be considered in cases where a constant tipping force acts in one direction, or when the lashings are pre-tensioned and this pre-tension is reliably maintained through periodic inspections during the voyage. Definitions of vertical and horizontal angles are presented in Figure 2.

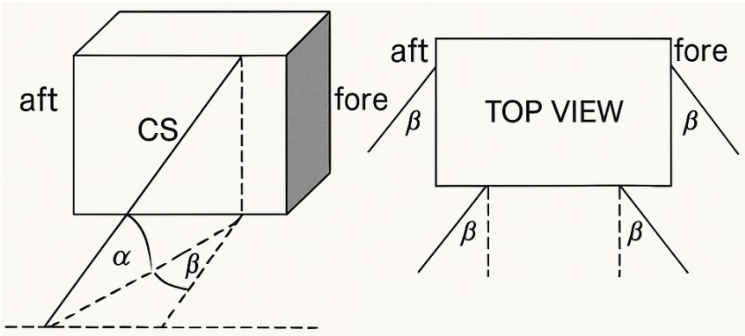


Fig. 2. Vertical α and horizontal β securing angles.

Transverse tipping
The balance equation is expressed as follows:

$$F_y.a \leq b.m.g + CS_1.c_1 + CS_2.c_2 + \ ..... \ + CS_n.c_n \qquad (3.5)$$

where

a : tipping lever arm [m]

b : stableness lever arm [m]

c : lever arm of securing force [m] (Figure 3).

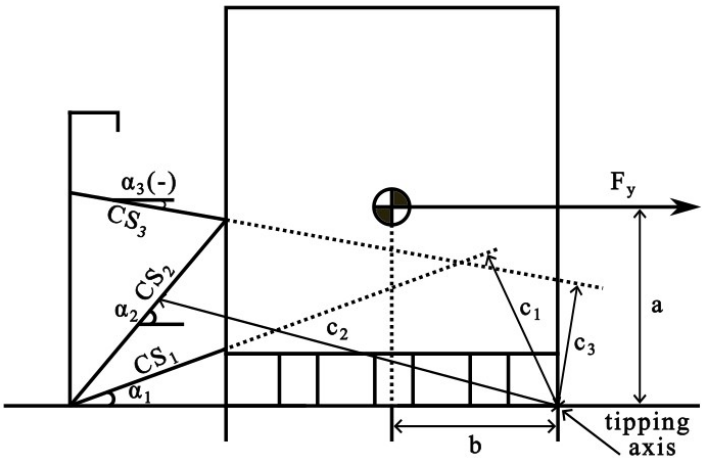


Fig. 3. Balance of transverse forces and moments.

Longitudinal sliding

Transverse lashings and frictional resistance are generally sufficient to prevent longitudinal sliding. However, to ensure reliability, the following calculation may be performed:

$$F_x \leq \mu.(m.g - f_z.F_z) + CS_1.f_1 + CS_2.f_2 + \ ..... \ + CS_n.f_n \quad (3.6)$$

where

F_x : longitudinal external force [kN]

F_z : vertical external force [kN]

f_z : correction factor depending on friction, see Table 6.

Table 6. Correction factor for vertical external force.

μ	0.0	0.1	0.2	0.3	0.4	0.6
f_z	0.20	0.50	0.70	0.80	0.85	0.90

The longitudinal contribution of transverse securing arrangements shall be limited to a maximum of 0.5 times CS [22].

3.3 Balance of forces - alternative method

A securing device positioned with a horizontal angle β, generates restraining forces in both longitudinal and transverse directions. CS of it, is calculated by applying the corresponding directional coefficients $\boldsymbol{f_x}$ and $\boldsymbol{f_y}$ which they are read from corresponding table of friction coefficient (0, 0.1, 0.2, 0.3 and 0.4). The table for μ = 0.3 is Table 7.

Table 7. f_x and f_y coefficients for μ = 0.3.

β for f_y	α														β for f_x
	-30	-20	-10	0	10	20	30	40	45	50	60	70	80	90	
0	0.72	0.84	0.93	1.00	1.04	1.04	1.02	0.96	0.92	0.87	0.76	0.62	0.47	0.30	90
10	0.70	0.82	0.92	0.98	1.02	1.03	1.00	0.95	0.91	0.86	0.75	0.62	0.47	0.30	80
20	0.66	0.78	0.87	0.94	0.98	0.99	0.96	0.91	0.88	0.83	0.73	0.60	0.46	0.30	70
30	0.60	0.71	0.80	0.87	0.90	0.92	0.90	0.86	0.82	0.79	0.69	0.58	0.45	0.30	60
40	0.51	0.62	0.70	0.77	0.81	0.82	0.81	0.78	0.75	0.72	0.64	0.54	0.43	0.30	50
50	0.41	0.50	0.58	0.64	0.69	0.71	0.71	0.69	0.67	0.64	0.58	0.50	0.41	0.30	40
60	0.28	0.37	0.44	0.50	0.54	0.57	0.58	0.58	0.57	0.55	0.51	0.45	0.38	0.30	30
70	0.15	0.22	0.28	0.34	0.39	0.42	0.45	0.45	0.45	0.45	0.43	0.40	0.35	0.30	20
80	0.00	0.06	0.12	0.17	0.22	0.27	0.30	0.33	0.33	0.34	0.35	0.34	0.33	0.30	10
90	-0.15	-0.10	-0.05	0.00	0.05	0.10	0.15	0.19	0.21	0.23	0.26	0.28	0.30	0.30	0

The balance equations are slightly different:
Transverse sliding

$$F_y \le \mu.m.g + f_{y1}.CS_1 + f_{y2}.CS_2 + \ ..... \ + f_{yn}.CS_n \qquad (3.7)$$

Transverse tipping

$$F_y.a \le b.m.g + 0.9\ x\ (CS_1.c_1 + CS_2.c_2 + \ ..... \ + CS_n.c_n) \qquad (3.8)$$

Longitudinal sliding

$$F_x \le \mu.(m.g - f_z.F_z) + f_{x1}.CS_1 + f_{x2}.CS_2 + \ ..... \ + f_{xn}.CS_n \qquad (3.9)$$

The contribution of securing devices to transverse tipping stability should be disregarded in cases where their vertical angle is less than 45°, and their horizontal angle exceeds 45°, within the context of the formula.

Safety factor when calculating CS is taken as 1.35 for this method [22].

4 Sample Calculations

Two explanatory examples, each solved using a different method, are presented herein. Both examples represent two distinct approaches operating under the framework of advanced calculation method. Balance of forces and moments approach utilized in the first example enables a calculation that is sufficiently reliable from a safety perspective; however, balance of forces - alternative method approach employed in the second example incorporates the horizontal lashing angles into the analysis, thereby providing a more accurate and comprehensive assessment.

4.1 Case 1 - Balance of forces and moments

A generator, the specifications of which are provided below, will be transported by sea. The safety of the voyage will be evaluated based on the adequacy of the existing securing arrangements. For an unrestricted and year-round operation, the characteristics of the vessel are presented in Table 8.

Table 8. Specs of the ship.

Ship	Length (m)	Beam (m)	GM (m)	Speed (knot)
specs	190	30	2.5	18

The characteristics of the cargo are presented in Table 9.

Table 9. Specs of the cargo.

Cargo	Length (m)	Width (m)	Height (m)	Mass (ton)	at 0.4 L
specs	5.8	2.3	2.4	6	on high deck

The cargo's front view is illustrated in Figure 4, and its top view is presented in Figure 5.

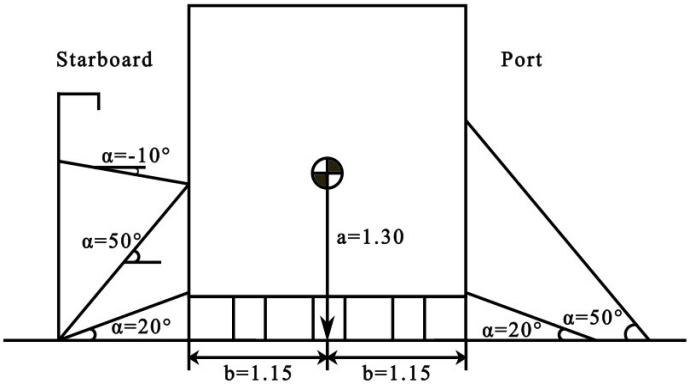


Fig. 4. Front view of the cargo.

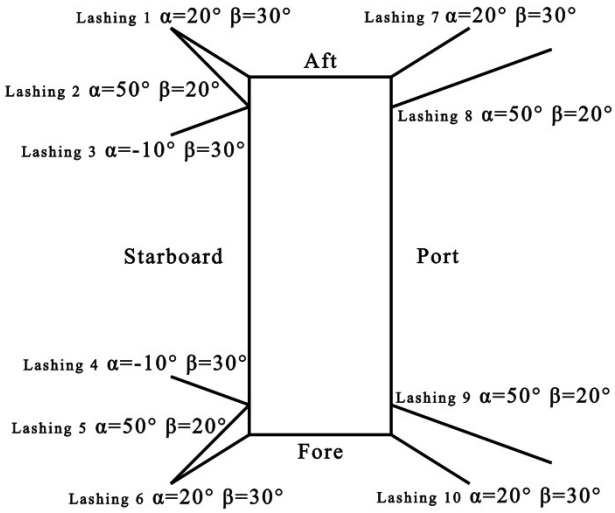


Fig. 5. Top view of the cargo.

Lashing equipment:
Shackles, deckeyes, turnbuckles Breaking strength 65 kN MSL = 32.5 kN
Wire rope (single use) Breaking stren gth 44 kN MSL = 35.2 kN
Timber dunnage μ = 0.3 (timber-steel) CS = 32.5/1.5 = 21.67 kN
Calculated Strength is calculated accourding to the weakest unit in line.

Lashing arrangement:

side	n	CS [kN]	α	f	n x CS x f
Starboard	2	21.67	-10	0.93	40.31
Starboar	2	21.67	50	0.87	37.71

d					
Starboard	2	21.67	20	1.04	45.07
Port	2	21.67	50	0.87	37.71
Port	2	21.67	20	1.04	45.07

External forces:

using formula 3.3. correction factor for ships length and speed = 0.73

F_x = 3.8 x 0.73 x 6 + (2.3 x 2.4 x 1) + (2.3 x 1.2 x 1) = 24.92 kN

F_y = 6.7 x 0.73 x 6 + (5.8 x 2.4 x 1) + (5.8 x 1.2 x 1) = 50.23 kN

F_z = 4.3 x 0.73 x 6 = 18.83 kN

Balance of forces, Starboard check:

$$F_y \leq \mu.m.g + CS_1.f_1 + CS_2.f_2 + \ldots\ldots\ldots + CS_n.f_n$$

0.3 x 6 x 9.81 + 40.31 + 37.71 + 45.07 = 140.75 kN

140.75 > 50.23 **lashing is safe.**

Balance of forces, Port check:

$$F_y \leq \mu.m.g + CS_1.f_1 + CS_2.f_2 + \ldots\ldots\ldots + CS_n.f_n$$

0.3 x 6 x 9.81 + 37.71 + 45.07 = 100.44 kN

100.44 > 50.23 **lashing is safe.**

Balance of moments, Tipping check:

$$F_y.a \leq b.m.g + CS_1.c_1 + CS_2.c_2 + \ldots\ldots\ldots + CS_n.c_n$$

1.15 x 6 x 9.81 = 67.69 kNm

50.23 x 1.30 = 65.30 kNm

67.69 > 65.30 **Safe** even without the contribution of the lashing.

4.2 Case 2 - Balance of forces, alternative method

In this second example, a more precise calculation has been conducted by also taking into account the horizontal angles.

Lashing arrangement:

No.heading	side	MSL	CS	α	β	f_x	f_y
1.aft	Starboard	32.5	24.07	20	30	0.57	0.92
2.aft	Starboard	32.5	24.07	50	20	0.45	0.83
3.fwd	Starboard	32.5	24.07	-10	30	0.44	0.80
4.aft	Starboard	32.5	24.07	-10	30	0.44	0.80
5.fwd	Starboard	32.5	24.07	50	20	0.45	0.83
6.fwd	Starboard	32.5	24.07	20	30	0.57	0.92
7.aft	Port	32.5	24.07	20	30	0.57	0.92
8.aft	Port	32.5	24.07	50	20	0.45	0.83
9.fwd	Port	32.5	24.07	50	20	0.45	0.83
10.fwd	Port	32.5	24.07	20	30	0.57	0.92

External forces:

F_x = 24.92 kN, F_y = 50.23 kN, F_z = 18.83 kN f_z = 0.8

Transverse balance, Starboard check:

$$F_y \leq \mu.m.g + f_{y1}.CS_1 + f_{y2}.CS_2 + \ldots\ldots\ldots + f_{yn}.CS_n$$

0.3 x 6 x 9.81 + (24.07 x 0.92) + (24.07 x 0.83) + (24.07 x 0.80) + (24.07 x 0.80) + (24.07 x 0.83) + (24.07 x 0.92) = 140.42 kN

140.42 > 50.23 **lashing is safe.**

Transverse balance, Port check:

$$F_y \leq \mu.m.g + f_{y1}.CS_1 + f_{y2}.CS_2 + \ldots\ldots\ldots + f_{yn}.CS_n$$

0.3 x 6 x 9.81 + (24.07 x 0.92) + (24.07 x 0.83) + (24.07 x 0.83) + (24.07 x 0.92) = 101.90 kN

101.90 > 50.23 **lashing is safe.**

Longitudinal balance, Forward check:

$$F_x \leq \mu.(m.g - f_z.F_z) + f_{x1}.CS_1 + f_{x2}.CS_2 + \ldots\ldots + f_{xn}.CS_n$$

0.3 x (6 x 9.81 – 0.8 x 18.83) + (24.07 x 0.44) + (24.07 x 0.45) + (24.07 x 0.57) +

(24.07 x 0.45) + (24.07 x 0.57) = 72.83 kN

72.83 > 24.92 **lashing is safe.**

Since the lashings are longitudinally (fore-aft) symmetrical, a unidirectional calculation is sufficient for verification purposes.

Transverse tipping check:

In the absence of specific information, the cargo's center of gravity may be reasonably assumed to lie at half its vertical height and at half its horizontal width. Furthermore, it is a reasonable simplification to take the lever arm **c**, defined as the distance from the tipping axis to the lashing point, as equal to the cargo's overall width (Figure 6).

$$F_y.a \leq b.m.g + 0.9\ x\ (CS_1.c_1 + CS_2.c_2 + \ldots\ldots\ldots + CS_n.c_n)$$

1.15 x 6 x 9.81 + 0.9 x 2.30 x (24.07 + 24.07 + 24.07 + 24.07) = 267.00 kNm

50.23 x 1.30 = 65.30 kNm

267.00 > 65.30 **lashing is safe.**

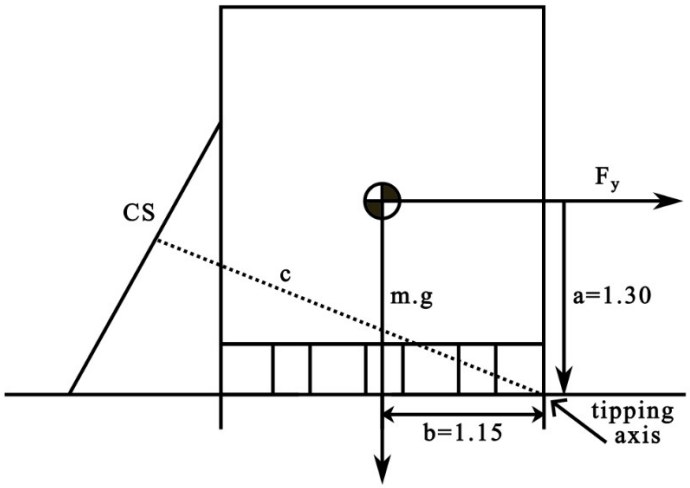


Fig. 6. Forces acting on transverse tipping.

5 Results and Discussion

Upon examining the calculations, it is evident that both methods yield results that are sufficiently close to be used interchangeably for evaluating transverse sliding. An adequate securing force has been established to effectively prevent transverse sliding. In cases with small horizontal lashing angles, as illustrated in the example, the f-factor - determined by the vertical lashing angle and the frictional force - does not significantly influence the outcomes to the extent of creating major discrepancies between the two approaches. This observation is explicitly addressed in the thoroughly revised Annex 13 of the CSS Code, which states: In cases where the horizontal lashing angle exceeds 30 degrees, the contribution of such lashings to transverse sliding resistance may be disregarded, given their relatively limited effectiveness. This situation is clearly observable in Table 7, where the coefficient ‘fy’ shows a striking decline as the ‘β’ angle increases - particularly beyond 30 degrees.

Furthermore, when the vertical lashing angle exceeds 60 degrees, the securing effect against sliding becomes increasingly inefficient. This degradation in effectiveness arises from the elongation of the lashing element as its length increases, thereby diminishing its ability to counteract movement. Lashings with such big angles may only be employed effectively if they are pre-tensioned and periodically re-tightened during the voyage to maintain their securing capacity. To reduce the adverse effects of elasticity in lashings primarily intended for sliding resistance, they should be kept short. Conversely, lashings designed to counteract tipping should be relatively longer in order to amplify the factor-c, as defined in Figure 3.

On the other hand, since the first method does not incorporate horizontal lashing angles, the contribution of transverse lashings to longitudinal sliding cannot be accounted for. In general, a longitudinal tipping analysis is not required, as most cargo units are inherently resistant to tipping longitudinally due to their form. However, for wide and tall non-standard cargo items, longitudinal tipping must be assessed using moment equilibrium calculations. In transverse tipping analyses, the alternative method that includes horizontal lashing angles applies a reduction factor of 0.9. This conservative adjustment accounts for the inefficiencies and uncertainties introduced by deviations from ideal transverse alignment.

According to the CSS Code, lashings with a vertical angle less than 45 degrees and at the same time horizontal angle greater than 45 degrees should not be used in transverse tipping calculations. Ideally, a lashing system should be oriented in a way that counteracts the acting force or moment directly, without any angular deviation. As horizontal angles increase, the lashing begins to share its load-bearing capacity along the longitudinal axis, thereby reducing its effectiveness in resisting transverse loads.

The real-time monitoring of dynamic loads acting on securing systems under varying severe wheather and sea conditions throughout the voyage should be considered a significant subject for future research. Computational methods must be enhanced through more advanced mathematical models. Instead of evaluating individual acceleration components in isolation, as currently practiced, the coupled motion analysis approach - such as that employed in Second Generation Intact Stability Criteria - will offer a more accurate estimation of loads acting on securing systems. Additionally, material-focused studies on securing elements could lead to improvements in their elasticity properties and fracture resistance, thereby enhancing the overall reliability of the securing arrangements. Standardization of friction-enhancing materials and coatings, along with performance validation under diverse climatic conditions, should be pursued to ensure repeatable safety margins. Automation in the application and tensioning of lashings, particularly for semi-standardized and non-standardized cargoes, may reduce human error and injury risk. Future revisions of the CSS Code could also benefit from incorporating AI-based decision support tools for voyage-specific securing recommendations.

6 Conclusions

The cargo is secured according to accepted principles, taking into account the dynamic forces that may occur during maritime transport and the harshest weather and sea conditions. In this study, some cargo lashing units are defined and the forces and moments that may occur on the cargo during the cruise are examined. Sample calculations are given about the adequacy of cargo lashing systems according to the forces acting on the cargo units. In addition, the requirements for the cargo lashing have been examined within the framework of international rules. Since it is not possible to calculate the maximum forces and moments that will affect the cargo units in the worst sea conditions, the basic rule should be to design a cargo lashing system that will provide sufficient strength. When designing a cargo lashing system, the ship's six-degree-of-freedom movements and the dynamic accelerations that will occur must be taken into account.

Lashing forces are derived from cargo accelerations resulting from ship motions. The highest accelerations—and consequently the most intense forces—typically occur at the front, rear, and uppermost stowage locations on either side of a vessel. In general, the loads acting on securing systems are resolved along the ship’s primary axes: longitudinal, transverse, and vertical. Among these, the longitudinal and transverse components are particularly critical from a lashing perspective, as lashings are primarily intended to prevent cargo units from sliding or tipping in those directions. Since transverse accelerations increase in direct proportion to the GM value, excessive accelerations should be avoided during cargo stowing and distribution. Under extreme sea conditions, it is difficult to precisely quantify the peak forces that may be encountered. Lashings tend to perform best when installed at angles ranging between 30° and 60° relative to the deck. If such optimal angles cannot be achieved, it may be

necessary to apply additional securing elements to ensure stability. It is clear that there are different loads other than the load units examined here and that there may be different forces and moments that can act on these cargo loads. Therefore, it will be understood that different examinations should be carried out both on the basis of international rules and when calculations are taken into account.

References

1. United Nations Conference on Trade and Development. Review of maritime transport 2024: Navigating maritime chokepoints. UNCTAD. Retrieved March 7, 2025, from https://unctad.org/publication/review-maritime-transport-(2024).
2. Melnyk, O., Onyshchenko, S., Onishchenko, O., Koskina, Y., Lohinov, O., Veretennik, O., & Stukalenko, O. Fundamental concepts of deck cargo handling and transportation safety. European Transport / Trasporti Europei, (98), Paper No. 1. https://doi.org/10.48295/ET.2024.98.1 (2024).
3. Shigunov, V., Rathje, H., & El Moctar, B. Towards safer container shipping. Ship Technology Research, 60(1), 3–12. https://doi.org/10.1179/str.2013.60.1.003 (2013)
4. Ling, D.-C., Tian, B.-J., & Liu, Y.-G. Amendments to the methods of non-standardized cargo unit securing assessing. In 2020 IEEE 5th International Conference on Intelligent Transportation Engineering (ICITE) (pp. 123–127). IEEE. https://doi.org/10.1109/ICITE50838.2020.9231357 (2020)
5. Acanfora, M., Montewka, J., Hinz, T., & Matusiak, J. On the estimation of the design loads on container stacks due to excessive acceleration in adverse weather conditions. Marine Structures, 53, 105–123. https://doi.org/10.1016/j.marstruc.2017.01.003 (2017)
6. OCIMF. (2019). Deck cargo management onboard offshore vessels. Oil Companies International Marine Forum. Retrieved March 13, 2025, from https://safety4sea.com/wp-content/uploads/2019/01/OCIMF-Deck-Cargo-management-Onboard-Offshore-Vessels-2019_01.pdf
7. Corigliano, P., Frisone, F., Chianese, C., Altosole, M., Piscopo, V., & Scamardella, A. Fatigue overview of ship structures under induced wave loads. Journal of Marine Science and Engineering, 12(9), 1608. https://doi.org/10.3390/jmse12091608 (2024)
8. Pérez-Canosa, J. M., & Orosa, J. A. A new methodology for optimization of lashing lines in the securing arrangement of non-standardized cargo on ships. Applied Sciences, 14(23), 11442. https://doi.org/10.3390/app142311442 (2024)
9. Matusiak, J. Dynamics of cargo shift onboard a ship in irregular beam waves. International Shipbuilding Progress, 47(449), 77-93. Helsinki University of Technology, Ship Laboratory, Finland (2000).
10. Prpić-Oršić, J., Valčić, M., & Čarija, Z. A hybrid wind load estimation method for container ship based on computational fluid dynamics and neural

networks. Journal of Marine Science and Engineering, 8(7), 539. https://doi.org/10.3390/jmse8070539 (2020)
11. Li, M., Wang, G., Liu, K., Lu, Y., & Wang, J. Experimental and numerical analysis of supporting forces and lashing forces in a ship cargo securing scheme. Journal of Marine Science and Engineering, 12(1), 158. https://doi.org/10.3390/jmse12010158 (2024)
12. Sternsson, M., & Björkenstam, U. Statistical study of lashing forces measured on-board a PCTC vessel. Marine Structures, 16(4), 345–354. https://doi.org/10.1016/S0951-8339(03)00020-0 (2003)
13. Kreuzer, E., Schlegel, V., & Stache, F. Multibody simulation tool for the calculation of lashing loads on RoRo ships. Multibody System Dynamics, 18(1), 73–80. https://doi.org/10.1007/s11044-007-9061-7 (2007)
14. Jia, J., & Ringsberg, J. W. Numerical and experimental investigation of dynamics of vehicle/ship-deck interactions. Marine Technology and SNAME News, 45(1), 28–41. https://doi.org/10.5957/mt1.2008.45.1.28 (2008)
15. International Maritime Organization. Code of Safe Practice for Cargo Stowage and Securing (CSS Code). Retrieved March 7, 2025, from https://www.imo.org/en/OurWork/Safety/Pages/CSS-Code.aspx
16. International Maritime Organization. Resolution A.1173(33) – Strategic plan for the Organization for the six-year period 2024 to 2029 (2023, December 11).
17. The London P&I Club. Reducing the risk of damage to or loss of non-standardized cargo: Operational guidance for securing during transportation. The London Steam-Ship Owners' Mutual Insurance Association Limited. https://www.londonpandi.com/Media/2942/6705-reducing-the-risk-update-v3-mar11-2022.pdf (2022)
18. Durlik, I., Miller, T., Kostecka, E., & Tuński, T. Artificial intelligence in maritime transportation: A comprehensive review of safety and risk management applications. Applied Sciences, 14(18), 8420. https://doi.org/10.3390/app14188420 (2024)
19. American Bureau of Shipping. Container Securing System Non-Linear Analysis (ABS Eagle C-LASH) https://ww2.eagle.org/en/rules-and-resources/engineering-software/marine-engineering-software.html
20. International Maritime Organization. Revised guidelines for the preparation of the cargo securing manual (MSC.1/Circ.1353/Rev.2) (2020, December 7).
21. International Maritime Organization. Amendments to the Code of Safe Practice for Cargo Stowage and Securing (CSS Code) (MSC.1/Circ.1352/Rev.1) (2014, December 15).
22. International Maritime Organization. Amendments to the Code of Safe Practice for Cargo Stowage and Securing (CSS Code) (MSC.1/Circ.1623) (2020, December 7).
23. International Organization for Standardization. BS ISO 23575:2022 - Ships and marine technology - Marine securing devices for ro-ro cargoes. British Standards Institution (2022).

24. DNV AS. December; amended 2025, March). Rules for classification: Ships - Part 6, Chapter 1: Structural strength and integrity (DNV-RU-SHIP Pt.6 Ch.1) (2024).
25. UK P&I Club. Carefully to carry: Consolidated edition 2023 (Chapter 14: Timber Deck Cargoes). Witherby Publishing Group (2023).
26. The London P&I Club, TMC Marine, and Bureau Veritas. Reducing the risk of damage to or loss of non-standardized cargo: Operational guidance for securing during transportation (2022).
27. UK P&I Club. Carefully to carry: Consolidated edition 2023 (Chapter 54 - Lashing and securing deck cargoes). Witherby Publishing Group (2023).

A Fuzzy DEMATEL Approach for Benefit–Risk Assessment in Sports and Recreational Activities

Ayhan Mentes[1], Hakan Akyildiz[2], Sebnem Helvacioglu[2]

[1] Istanbul Technical University, İstanbul 34469, Turkey
mentes@itu.edu.tr
[2] Piri Reis University, İstanbul, 34940, Turkey
hakyildiz@pirireis.edu.tr
shelvacioglu@pirireis.edu.tr

Abstract. The terms “sports” and “recreation” encompass a range of activities and the necessary equipment for users across all age ranges and abilities. In line with their nature, recreational activities contain a degree of tolerable risks and present different challenges for all ages. These potential risks must be balanced to maximize the overall social utility and public good. Therefore, utilizing a risk assessment method in recreational activities is beneficial. Benefit–risk assessment has emerged as a critical decision-support tool for mitigating accidents, damage, and hazardous events in sports and recreational settings. It facilitates the development of more realistic and robust safety strategies.

To improve safety in recreational activities, it is essential to identify risk factors and make safety-based decisions logically and confidently. Risk modeling and decision-making tools need to be developed and applied in a practical environment. The benefit–risk assessment methodology requires the identification of external and internal contexts for the user(s). Benefit and risk evaluation can be carried out by the interrelationship of generic, site-specific, and dynamic benefit–risk assessments, which is developed as an integrated methodology.

This study introduces an integrated methodology that combines fuzzy set theory and the Decision-Making Trial and Evaluation Laboratory (DEMATEL) technique to identify and evaluate key contributing factors, including physical context and failure modes associated with adverse outcomes. A case study focusing on unintentional incidents, the most common causes of recreational accidents, has been presented. This case study aims to quantify interrelationships among the identified risk factors by utilizing an integrated methodology. The results provide a set of targeted recommendations to enhance safety and reduce the risk.

Keywords: Recreation activity, Fuzzy set theory, DEMATEL, Risk, Safety.

1 Introduction

To maximize safety in recreational sectors, risk factors must be modeled, and safety-based decisions must be made logically and confidently. Risk modeling and decision-making tools need to be developed and applied in a practical environment. The study aims to facilitate risk modeling for cleaner and safer sports. To facilitate this, an integrated benefit–risk assessment methodology will be developed. The proposed benefit–risk assessment approach combines fuzzy set theory (FST) and decision-making trial and evaluation laboratory technique (DEMATEL). The proposed methodology resolves some of the problems in traditional methods of evaluation and it has several advantages over strictly numerical methods: (1) ambiguous, qualitative or imprecise information, as well as quantitative data can be used in assessment and they are handled consistently; (2) the risk associated with item failure modes can be evaluated directly using linguistic terms that are employed in making the criticality assessment; (3) it gives more flexible structure for combining the severity and occurrence parameters; (4) It considers the indirect relationships in relation analyze. This method allows the analyst to evaluate the risk associated with item failure modes directly using linguistic terms that are employed in making the criticality assessment. Combining quantitative and qualitative methods makes risk analysis more flexible and more representative of the real-world decision-making problems.

As the recreational sectors have specific characteristics and face different challenges for all ages and abilities, there is an expectation that the user is making an implicit trade-off between the benefits and the inherent risks of the activity, including the potential for dangerous situations, in achieving a balance that maximizes the overall social utility and public good. These are the considerations that bring rewards but also dangers.

Risks and benefits, based on the context of a project scope, should include the objectives of injury and lethality (risk) and the improvement of health and wellbeing in a broad sense (benefit), as required by law.

The terms “sports” and “recreation” describe diverse activities and the necessary equipment for users of all ages and abilities to assess the acceptable risks and the minimum benefits to be achieved. The differentiation of sports and recreational activities is designed for the public good, including the social, physical, and psychological health and welfare of participants and society. It involves the varied interests of the different stakeholders, such as designers, manufacturers, and recreation equipment and facilities. Therefore, a single system of hazard identification and risk assessment is not feasible.

Participation in recreational activities involves exposure to risk based on the age of participants when assessing benefits and risks. In order to be balanced with the exclusion of unforeseen or unrecognizable sources of harm, exposure to risk in daily life can reduce fear and improve the development of human competency [1].

The risks and benefits are complicated by the fact that they often cannot be measured on a uniform scale for subjects across different domains of health status and individuals [2]. Benefit–risk assessment for sports and recreational activities should be applied by different types of assessment: the generic risk assessment, the site-specific risk assessment, and the dynamic risk assessment. A generic risk assessment technique can

be used to analyze the risks related to sports and recreational activities. A site-specific risk assessment, on the other hand, can be used by the owner/operator to evaluate the characteristics of the site, while the participant and instructor dynamically assess the risk during the activity.

This study specifies methods for benefit–risk assessment for recreational sports and facilities, including equipment. It also provides requirements on benefit–risk assessment within this field, including examples for injury thresholds.

2 Benefit–risk assessment

In the benefit-risk process, the first task performed when conducting a benefit–risk analysis should be the context developed in terms of the requirement to assess a product, facility/location, or activity during its conduct. It is of great importance to define the context based on internal and external factors for effective risk assessment and appropriate management of the activity. Understanding the external and internal environment where the activity takes place results in understanding the nature of the facility or activity, and the social acceptance of the risks related to the activity. Therefore, the context should include the type of sports and recreational experience, the nature of the activity itself and the inherent benefits and risks created by that activity, the geographic location or other relevant factors. Benefit assessment is directly connected to the type and geographic location of the activity and the social aspects. Risk management takes place in the context of objectives and sports and recreational activities.

In this paper, the establishment of context consists of defining the context, identifying the user(s), and evaluating the risk and benefits (see Fig. 1).

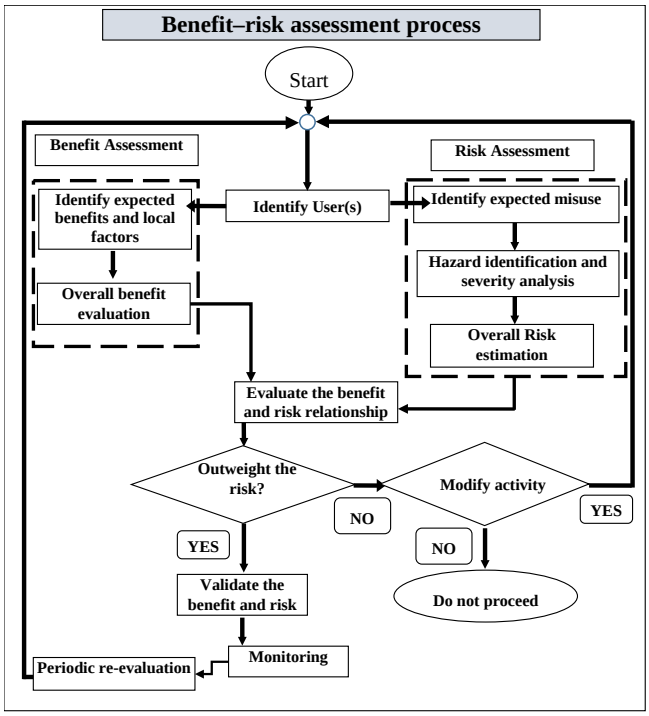


Fig. 1. Benefit–risk assessment process.

2.1 External and internal context

The context of the benefit and risk assessment provides an understanding of the benefit assessment and the risk assessment regarding the legal limitations, the limits for risk tolerability, and other considerations. Cultural and legal aspects affecting the level of risk appetite shall be part of establishing the context and determination of the factors that need to be included and involve communication between stakeholders.

Examination of the external context may include, but is not limited to:

- the environment (e.g., urban, rural, remote),
- the social, cultural, political, legal, regulatory, financial, educational, technological, economic, and environmental factors, whether international, national, regional, or local,
- community values,
- types of users/participants,
- user needs.

Examination of the internal context may include, but is not limited to:

- vision, mission, and values of the organization providing the product, facility, or activity,
- the governance, organizational structure, roles, and accountabilities,
- strategy, objectives, and policies,

- the organization's culture,
- standards, guidelines, and models adopted by the organization,
- capabilities, understood in terms of resources and knowledge (e.g., capital, time, people, intellectual property, processes, systems, and technologies),
- data, information systems, and information flows,
- relationships with internal stakeholders, considering their perceptions and values,
- contractual relationships and commitments,
- interdependencies and interconnections,
- environment (urban, rural, remote),
- local policy, community values,
- types of users,
- user needs.

2.2 Identification of user(s)

Benefit–risk assessment should be conducted by a team of at least the minimum number of participants necessary to include all relevant expertise. Team leaders should have training and experience, including the development of consensus. Team members should also be chosen from the group of product and facility designers, site managers, and activity leaders.

Benefit and risk assessment as parallel processes are required for concluding with a final evaluation of the benefit and risk balance and ongoing maintenance activities. Then the next step is to determine the expected users of sports and recreational activities, considering as follows:

- Age,
- physical size,
- skills,
- capacity for judgment,
- capacity for supervision.

Once the user is identified, the process moves to two parallel processes determined by the assessment team. The identification and evaluation of the benefits from the activity begins with the identification of the expected benefits. At this point, a benefit analysis can be performed, including identification of any precedents and comparisons. The process continues with hazard identification and the development of an understanding of the potential severity of the injury or specified harm. At this point, an estimation of the risk can be performed, resulting in an overall evaluation of the risk.

2.3 Benefit and risk evaluation

This process results in a decision as to whether the risk is tolerable or whether the benefits outweigh the risks or both. The decision, based on the benefits outweighing the risks, either results in the activity moving forward or being re-evaluated for further risk reduction. At this stage, the team has the option of:

- returning to the hazard identification and determining if the risk(s) have been adequately reduced for the project to proceed, or
- returning to the beginning of the entire process, re-evaluating both the benefits and the risks with different starting parameters, such as the types of users or other parameters.

Where the activity moves forward, the decision shall be validated and documented. Periodic monitoring and re-evaluation of the benefits and risks shall be conducted, which means performing the entire process again (see Fig. 1).

2.4 Interrelationship of generic, site-specific and dynamic benefit–risk assessments

The generic risk assessment, the site-specific risk assessment, and the dynamic risk assessment are dependent on the organization conducting the benefit–risk assessment. In general, the generic benefit and risk assessment method can be used for products and facilities, while for site operators, some form of site-specific benefit and risk assessment is preferred, and dynamic benefit–risk assessment is preferred for activity leaders and team members. In Figure 2, designers should not only use generic benefit and risk assessment but also should be aware of the benefits of their designs. Site owners or operators should understand the benefits of the equipment on their site and the capabilities of the users they intend to attract, and activity leaders or team members should have some awareness of product design issues, and which products are fit for the intended purpose.

Generic benefit–risk assessment is considered generic at the product design stage by the product designer or manufacturer, because it is conducted without specific knowledge of the location where the product is used or any personal details about the users. This process focuses on products, and equipment designs due to the probability aspects, including references to intended use.

Site-specific benefit–risk assessment is conducted by the site designer, the installer, the owner/operator, or others at the installation and operational stage. The process considers the benefits and risks associated with the intended use(s) and the reasonably foreseeable misuse(s) of a specific sport and/or recreation environment. Both the physical site and the activity(s) provided on the site by the owner/operator should be assessed according to the policy of the organization. The site designer shall also review the compliance with the manufacturer’s installation specification, established equipment standards, and local regulations.

Site-specific benefit–risk assessment can be used for permanent or temporary installations associated with the site selection and layout of all pieces of free-standing equipment. The steps included in the generic-risk assessment technique apply to the site-specific assessment technique. A sport or recreation facility or an activity area can be a formal structure or a location set aside for the activity. The facility policy determines the desired benefits and risk appetites. The facility or activity area should be structured and determined by the facility policies, activity leadership, environmental conditions, and the abilities of the intended and reasonably foreseeable users.

Dynamic benefit–risk assessment is conducted by activity leaders, participants, or others at the activity stage. This process considers the local environmental factors (e.g., weather), and actual user abilities and fitness.

Generally, the relationship between generic and site-specific refers to understanding how an activity, product, or facility is applied to a specific location. The relationship between dynamic and generic is required to blend immediacy with an understanding of the activity, product, or facility. Additionally, the relationship between dynamic and site-specific is also required to blend immediacy with an understanding of the location. In the overall review, all approaches are a blend of all aspects of risk assessment (see in Fig. 2).

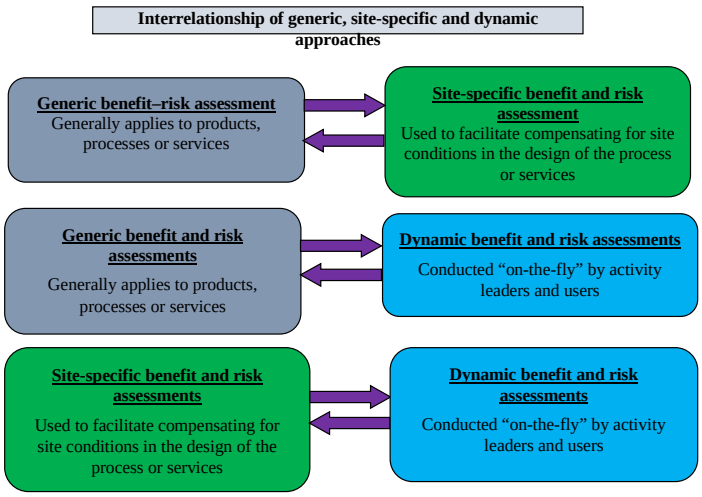


Fig. 1. Interrelationship of generic, site-specific and dynamic benefit–risk assessment.

2.5 Benefits and risks assessments

There are quantitative and qualitative analyses. Quantitative analysis describes statistical analysis techniques that rely upon numeric data. Whenever possible, facility, service, or product providers should gather numeric data relating to the types, severities, and frequencies of injuries suffered by users, so that the data is used for quantitative analysis. Qualitative analysis describes a wide range of methods that, while not mathematically rigorous, provide practitioners with tools that can facilitate decision-making in the absence of numeric data.

These methods can be quite simple or complex. The selection of an appropriate tool by the practitioner is important since many of the available tools come from the occupational health and safety sectors (see in Fig. 1).

Benefit assessment

Benefits can be assessed quantitatively or qualitatively. The public health benefits of sport and recreation are now widely recognized. Benefits can be identified for persons

of all ages and circumstances. The main objective for activity providers, such as sport and recreational facilities, is to maximize the public's gain, which includes physical, mental, emotional, and social wellbeing and enjoyment. The procedure for the assessment of the benefit includes the identification of the expected benefits, local factors and thereafter the application of the benefit analysis method. Whether to attempt a quantitative analysis or to use a qualitative approach is a decision for the local decision-maker.

Risk assessment

Risk assessment aims to gather information to evaluate benefit and risk relationships, considering hazard identification and severity analysis, hazard elimination, risk evaluation, and risk control measures. This process is required to identify intended use and reasonably foreseeable misuse, and risk estimation under the local legislation and regulations. Otherwise, the risk criteria may be set by the responsible person, sports federation, national or regional bodies, based on risk appetite in terms of the external and internal context.

2.6 Benefits and risks comparison

The process of comparing benefits and risks brings the benefit assessment and risk assessment together for further evaluation and conclusion (see Fig. 1). The benefit and risk-related information shall be synthesized to provide a platform upon which to evaluate the balance and to determine whether the project moves forward. Providing information to justify decisions and achieve the objectives of operational intent is of great importance in evaluating the balance between risks and benefits. The quality of information is required to ensure that risks and benefits are comparable. In assessing the applicability of the benefit and risk comparison, they have been identified in the same detail and logic (i.e., for the same facility or activities, user groups, period, surroundings, and other variables) in terms of the external and internal context.

3 Methodology: An Integrated Benefit–Risk Assessment Approach for Recreational Activities

To evaluate the acceptability of a recreational activity based on benefit–risk considerations, a structured decision pathway is applied (see Fig. 1). This process yields two possible outcomes: either the benefits outweigh the risks, rendering the activity viable, or the risks surpass acceptable thresholds, and the associated benefits are deemed insufficient to justify proceeding. When the benefits prevail, the activity may continue as planned or be reassessed to further reduce risks and/or enhance benefits. In such scenarios, the process may follow one of two trajectories:

(1) revisiting the hazard identification phase to verify whether the remaining risks have been adequately mitigated, or

(2) reinitiating the entire benefit–risk assessment procedure, potentially with the inclusion of new user groups or contextual parameters.

Conversely, if the benefit–risk balance is unfavorable, strategies may be implemented to mitigate risks or to increase the perceived or actual benefits. Such improvements may include raising user awareness, promoting nature-based engagement, enhancing the local environment and aesthetics, or generating positive socioeconomic impacts.

Benefit–risk assessments are increasingly acknowledged as effective decision-support mechanisms for addressing accidents, damages, and safety-critical situations in sports and recreational contexts. The integration of complementary analytical techniques facilitates the development of solutions that are not only reliable but also contextually responsive and operationally feasible. To this end, the present study introduces a comprehensive methodological framework for systematically identifying and evaluating key factors—such as geographic conditions at the time of incident occurrence and the failure modes that give rise to hazardous outcomes. Within this framework, an integrated multi-method approach is proposed, combining fuzzy set theory and the Decision-Making Trial and Evaluation Laboratory (DEMATEL) method. Collectively, these methods enable the construction of a robust benefit–risk assessment model that supports informed, balanced, and adaptive decision-making under conditions of uncertainty.

3.1 Fuzzy Set Theory

Fuzzy Set Theory (FST) has emerged as a robust mathematical framework for modeling uncertainty and imprecision in complex systems across industrial, environmental, and social domains. Initially introduced by Zadeh in 1965, FST enables the representation and manipulation of vague and incomplete information, thereby facilitating more intuitive and adaptable reasoning processes in decision-making [3, 4].

FST is particularly effective in modeling situations where classical binary logic fails to capture gradations in membership. For instance, in a referential set comprising individuals who regularly engage in sports, an object (e.g., a person) fully belonging to the subset “daily exercisers” would be assigned a membership degree of 1, while those completely outside the subset would be assigned a membership of 0. Intermediate or borderline cases receive values in the interval [0, 1], enabling nuanced representation through a continuous scale [4].

These degrees of membership can be linguistically expressed as seldom, often, very often, etc., and can be quantified as fuzzy numbers through appropriate transformation scales. Wang et al. [5], for example, proposed such scales in the context of risk-related decision problems.

Several key milestones have marked the development of FST:

1965: Lotfi Zadeh introduced the foundational concept of fuzzy set theory.

1970: Bellman and Zadeh [6] extended FST to decision-making problems characterized by vagueness and uncertainty.

1975: Yager and Basson [7] proposed new decision-making models using FST.

1977: Bass and Kwakernaak [8] established fundamental FMADM (Fuzzy Multi-Attribute Decision-Making) techniques.

1980: Dubois and Prade [4] formalized mathematical models and demonstrated FST applications in diverse scientific domains.

1990s: A surge in literature reviews by scholars such as Zimmermann [9], Lai and Hwang [10], Klir and Yuan [11], and Lootsma [12] enriched the methodological landscape.

2006: Kahraman et al. [13] conducted a comprehensive review consolidating FST applications in engineering and decision sciences.

2009: Wang et al. [5] proposed transformation scales for fuzzy linguistic variables to support decision-making in risk-type problems, significantly improving the usability of fuzzy AHP for real-world applications.

2010: Feng et al. [14] developed a hybrid mathematical approach by integrating soft sets, fuzzy sets, and rough sets, enabling more flexible modeling of uncertainty in multi-criteria decision-making contexts.

2011: Greco et al. [15] introduced a game-theoretic framework for rough sets within a fuzzy context, enhancing decision-theoretic modeling under uncertainty.

2012: Pishvaee et al. [16] proposed a robust possibilistic programming model for socially responsible supply chain network design under uncertainty, applying fuzzy logic to logistics and sustainability planning.

2019: Liu et al. [17] introduced the Fermatean fuzzy linguistic set, expanding the expressive capacity of fuzzy decision models and enabling improved granularity in multi-criteria decision-making problems.

2020: Kahraman and Kaya [18] developed a comprehensive fuzzy multi-criteria decision-making methodology for evaluating energy alternatives, demonstrating practical applications of FST in energy policy and planning.

2021: Aragonés-Beltrán et al [19]. applied a fuzzy AHP–TOPSIS hybrid model to prioritize large-scale infrastructure projects, showcasing how FST can support complex public-sector investment decisions.

2022: Yazdani et al. [20] applied fuzzy decision-making techniques to vaccine allocation problems during the COVID-19 pandemic, demonstrating the adaptability of FST in time-sensitive and uncertain health crises.

2023: Kou et al. [21] integrated hesitant fuzzy linguistic term sets into risk evaluation frameworks for smart cities, advancing urban risk modeling capabilities with enhanced expression of expert judgment.

2024: Zhou et al. [22] introduced a new interval-valued fuzzy TOPSIS method for selecting sustainable suppliers, using FST to evaluate corporate social responsibility and environmental sustainability criteria in supply chains.

A fuzzy number of membership functions can be defined in several methods, such as a trapezoidal (or triangular), where the membership function forms a trapezium (or triangle). A triangular fuzzy number can be defined in Equation (1) where $\mu_A(x)$ is the membership function. A trapezoidal fuzzy number can be denoted as in Equation (2) where $\mu_A(x)$ is the membership function. Chen [23] defined another membership function definition given in Equations (3)-(7). In Chen's method, there are two phases:

in the first one, fuzzy data are converted into numerical scores, and in the second one, a ranking between alternatives is carried out.

$$\mu_A(x) = \left\{ \begin{array}{ll} 0 & x < a_1 \\ \frac{x-a_1}{a_2-a_1} & a_1 \le x \le a_2 \\ \frac{a_3-x}{a_3-a_2} & a_2 \le x \le a_3 \\ 0 & x > a_3 \end{array} \right\} \tag{1}$$

Trapezoidal Fuzzy Numbers

$$\mu_A(x) = \left\{ \begin{array}{ll} 0 & x < a \\ \frac{x-a}{b-a} & a < x \le b \\ 1, & b < x < c \\ \frac{d-x}{d-c}, & c \le x < d \\ 0 & x > d \end{array} \right\} \tag{2}$$

$$\mu_{\max}(x) = \left\{ \begin{array}{ll} x, & 0 \le x \le 1 \\ 0, & otherwise \end{array} \right\} \tag{3}$$

$$\mu_{\min}(x) = \left\{ \begin{array}{ll} 1-x, & 0 \le x \le 1 \\ 0, & otherwise \end{array} \right\} \tag{4}$$

The right score of M can be determined using:

$$\mu_{right}(M) = \sup[\mu_M(x) \cap \mu_{max}(x)] \tag{5}$$

The left score of M can be determined using:

$$\mu_{left}(M) = \sup[\mu_M(x) \cap \mu_{min}(x)] \tag{6}$$

Given the left and right scores, the total score of M can be computed using:

$$\mu_{total}(M) = [\mu_{right}(M) + 1 - \mu_{left}(M)]/2 \tag{7}$$

3.2 DEMATEL Method

Decision-making trial and evaluation laboratory (DEMATEL) is considered an effective method for the identification of cause-and-effect chain components of a complex system. It deals with evaluating interdependent relationships among factors and finding the critical ones through a visual structural model. This method not only converts interdependency relationships into a cause-and-effect group via matrices but also identifies the critical factors of a complex system using an impact relation diagram [24].

The Decision-Making Trial and Evaluation Laboratory (DEMATEL) method is a widely recognized and extensively applied technique for analyzing complex causal relationships among criteria in multi-criteria decision-making (MCDM) problems. Originally developed by Gabus and Fontela at the Battelle Memorial Institute in 1972 and 1973, DEMATEL evolved through integration with various fuzzy and weighted aggregation techniques.

In 2007, Wu and Lee [25] combined fuzzy logic with DEMATEL to classify essential competencies for global managers. Alizadeh et al. [26] further enhanced the methodology by integrating historical data into the construction of fuzzy cognitive maps. Chang [27] extended the approach by incorporating Ordered Weighted Geometric Averaging (OWGA) into DEMATEL to improve Risk Priority Number analysis.

In 2011, Liaw et al [28]. utilized DEMATEL to capture group knowledge and uncover causal relationships among evaluation criteria. The following year, Wang et al. [29] proposed a hybrid model employing DEMATEL and importance-satisfaction analysis to identify key divisions responsible for underperformance in design projects. In 2013, Yang et al. [30] developed an information security risk-control assessment model by integrating VIKOR, DEMATEL, and the Analytic Network Process (ANP).

Later, in 2016, Lin and Jia [31] introduced a risk assessment framework that leveraged the strengths of both the OWGA operator and DEMATEL, offering a more nuanced evaluation of interrelated risk factors. Most recently, in 2022, Mentes and Helvacıoğlu [32] presented a comprehensive hybrid methodology for the safe design of yachts by integrating DEMATEL with OWGA and Grey Theory, thereby illustrating the method's continued evolution and relevance in complex system design.

In the method pair-wise comparison scale may be chosen into four levels the scores of 0, 1, 2, and 3 characterise ‘‘No influence’’, ‘‘Low influence’’, ‘‘High influence’’, and ‘‘Very high influence”.

The primary direct-relation matrix Z is a matrix that is acquired by pair-wise evaluations in terms of influences and directions between measures, in which z_{ij} is represented as the degree to which the principal i penetrates principal j, i.e. $[z_{ij}]_{nxn}$ The normalised direct-relation matrix X can be calculated using Equations (8) and (12) [32]:

$$X = \frac{Z}{s} \tag{8}$$

$$s = \frac{1}{\max_{1\le j\le n}(x\sum_{j=1}^{n} z_{ij}}, \; i, j = 1, 2, \ldots, \; . \tag{9}$$

The total relation matrix T can be acquired by formula (10), in which the ‘I’ is denoted as the identity matrix:

$$T = \lim_{k\to\infty}(X + X^2 + \ldots + X^k = X(I - X)^{-1} \tag{10}$$

The sums of the rows and columns denoted as D_i and R_j respectively, can be obtained through the following two formulas:

$$D = \sum_{j=1}^{n} t_{ij}, \quad (j = 1, 2, \ldots, n) \tag{11}$$

$$R = \sum_{i=1}^{n} t_{ij}, \quad (i = 1, 2, \ldots, n) \tag{12}$$

A causal diagram can be attained by mapping the dataset of (D + R, D − R), where the horizontal axis (D + R) is titled "Prominence", and the vertical axis (D −R) is titled "Relation". When the (D − R) axis is positive, the related element pertains to the cause group. Otherwise, the element pertains to the effect group [32].

In this study, a benefit and risk analysis methodology is proposed to address the high-risk areas that need design or operational attention in an activity, facility, or product.

An integrated benefit and risk assessment methodology has been developed for achieving these goals based on fuzzy set theory (FST) and decision-making trial and evaluation laboratory technique (DEMATEL).

The procedure of the proposed methodology is organized into seven steps and explained as follows:

Step 1 Primarily Analysis: The first step is primarily analysis, which includes the identification of the problem and classification of goals and objects. Accident data is gathered by using accident reports and past experiences in this step.

Step 2 Brainstorming: This step includes the identification of the accident scenarios, the determination of accident sub-criteria, and the investigation of the accidents' causes and effects. A commission is employed to carry out this step.

Step 3 Defuzzification Process: The commission of the occurrence and severity of the failure modes due to the data limitations. The expert judgments were gathered during several workshops and interviews. The experts make judgments by expressing their opinions based on their experience / knowledge, and expertise. Since the experts are not able to appraise the exact frequency and severity of failure modes, they tend to apply linguistic expressions, such as 'very low', low', moderate', high', and 'very high', to describe the probabilities of events. Fuzzy numbers are used by the experts for determination of occurrence and severity (Table 2 & Table 3). Using Chen's [33] method, fuzzy numbers are converted to deterministic numbers. These numbers are used to acquire risk matrix.

Step 4 DEMATEL Method: In this step, the DEMATEL technique is applied to analyze the cause–effect relationships among the identified risk factors. It helps classify the factors into driving (cause) and dependent (effect) groups by calculating their mutual influence levels based on expert evaluations.

Step 5 Risk Control Options: In this step, risk factors are found, and risk control options are given.

Step 6 Cost-Benefit Analysis: In this step, the cost analysis procedure is applied, and the best cost/benefit ratio is calculated.

Step 7 Decision Making: The results are analysed and provided with feasible solutions and suggestions according to risk problems.

4 Case Study on Risk–Benefit Tradeoffs in Recreational Activities

In this study, a hybrid methodology is applied to conduct a comprehensive benefit–risk assessment for sports and recreational activities. The approach is grounded in an extensive literature review, through which 15 critical barriers contributing to hazardous conditions were identified and examined in detail. The goal of the proposed framework is to systematically support inclusive design strategies that enhance safety, maximize user benefit, and mitigate potential harm in recreational settings.

Specifically, the proposed framework is designed to:

- Develop a robust risk evaluation method that enhances safety for all user groups and amplifies the overall social, environmental, and economic benefits of the activity;
- Identify and assess the driving barriers most responsible for creating unsafe conditions;
- Handle ambiguous, imprecise, and qualitative inputs—including linguistic or uncertain expert judgments;
- Allow for a flexible combination of risk parameters such as severity and likelihood;
- Incorporate the DEMATEL technique to analyze direct and indirect relationships among risk factors;
- Prioritize and rank high-risk failure modes based on their systemic influence.

The comprehensive risk assessment process incorporates a diverse set of interrelated factors, as detailed in Table 1. These factors span across multiple domains, reflecting the systemic nature of risk in complex operational environments.

Environmental factors, including international, national, regional, and local conditions—such as climate, daylight variability, and geographical context—directly influence exposure to risk and the design of mitigation strategies. Similarly, user-specific characteristics—such as physical condition, skill level, individual needs, and risk-seeking tendencies—play a critical role in shaping both vulnerability and behavioral responses to risk.

Human factors, particularly the competence, experience, and expertise of facilitators, supervisors, and practitioners, are central to the safe execution of activities. These factors interact closely with educational components, such as training procedures and professional development, which are essential for fostering a sustained awareness of safety and ensuring the correct application of benefit–risk frameworks across stakeholder groups.

Table 1. Barriers for fuzzy DEMATEL analysis.

No	Barriers
1	Environmental Factors (international, national, regional or local).
2	Types of Users, User Profile & User Needs
3	Human Factors (with relevant expertise in an activity; considering experience in the activity).

4	Social, Cultural, Political Factors (international, national, regional or local)
5	Regulatory Factors (international, national, regional or local)
6	Incompleteness and complexity of legislation and procedures for safe operations (Community Values), Incomplete, unfinished or complex legislation, laws and procedures (within the framework of social values).
7	Educational Factors (Training- related procedures)
8	Technological Factors
9	Vision, mission and values of the organization providing the product, facility, or activity; Rules, vision, and mission of facilities that provide equipment or activity opportunities for sports and recreational activities.
10	The governance, organizational structure, roles and accountabilities, Management, responsibilities, liability.
11	Strategy, objectives and policies (international, national, regional or local)
12	Standards, guidelines and models adopted by the organization; Rules and principles guide, organizational models.
13	Data, information systems and information flows
14	Connections and relationships with internal stakeholders
15	Contractual relationships and commitments; Contract terms and obligations.

Social, cultural, and political dynamics, at all jurisdictional levels, shape societal perceptions of risk, tolerance thresholds, and behavioral norms. These dimensions often influence regulatory responses, including the design, implementation, and enforcement of legislation and procedural frameworks. In this regard, regulatory factors, especially when laws and procedures are fragmented, incomplete, or overly complex—can become barriers to consistent risk governance. Thus, the clarity and coherence of the legal framework, aligned with community values, are essential to operational safety.

Technological factors introduce additional layers of complexity. Rapid technological advancement affects equipment design, selection, and monitoring processes, necessitating continuous adaptation to evolving standards and operational practices. Accordingly, organizational standards, guidelines, and models serve as tools for harmonizing internal procedures with international best practices and for ensuring the consistent application of risk-reduction strategies.

The vision, mission, and values of the organization, as well as its governance structure, defined roles, and accountabilities, influence how risks are managed and who holds ultimate responsibility for both risk mitigation and benefit realization. These organizational characteristics must align with strategic objectives and policy frameworks, which should be coherent with legal principles and aim to minimize foreseeable misuse or unintended consequences.

Finally, internal data systems, information flows, and relationships with stakeholders and contractual entities form the informational and institutional backbone of the risk management system. Efficient communication channels and clearly defined contractual

obligations enhance transparency, accountability, and responsiveness in managing emergent risks.

To evaluate and structure the benefit–risk dynamics of sports and recreational activities, a fuzzy DEMATEL-based decision framework was applied. This case study systematically follows the methodological sequence of fuzzy set theory, defuzzification, and DEMATEL matrix computations to quantify interrelationships among the identified risk factors.

The subjective evaluations obtained from domain experts were initially expressed using linguistic variables (e.g., very low, low, medium, high, very high), which were then translated into triangular fuzzy numbers (TFNs) based on the scale presented in Table 2. Each linguistic term was represented as a triangular fuzzy number A (l_i, m_i, u_i) where l_i, m_i, and u_i denote the lower, modal, and upper bounds, respectively.

Table 1. Triangular fuzzy number scale.

		$A(l_i, m_i, u_i)$		
		l_i	m_i	u_i
Abbreviation	**Linguistic Values**	**Low**	**Medium**	**High**
N	None	0.000	0.000	0.100
VL	Very low	0.000	0.100	0.200
L	Low	0.100	0.300	0.500
FL	Fairly low	0.300	0.400	0.500
ML	More or Less Low	0.400	0.450	0.500
M	Medium	0.300	0.500	0.700
MG	More or Less High	0.500	0.550	0.600
FG	Fairly high	0.500	0.600	0.700
G	High	0.500	0.700	0.900
VG	Very high	0.800	0.900	1.000
E	Extremely High	0.900	1.000	1.000

Based on expert inputs, the direct-relation matrix for the 15 risk factors (given previously in Table 1) was constructed in linguistic form. These linguistic evaluations were then converted into TFNs according to Table 2, resulting in a fuzzy direct-relation matrix, denoted as D, as shown in Table 3.

To transition from fuzzy to crisp analysis, Chen's defuzzification method was applied. For a given triangular fuzzy number A(li, mi, ui), the defuzzified (crisp) value was computed. This method ensures greater weight is given to the modal value (most likely value) in the TFN. Applying this to each element in Table 3, the crisp direct-relation matrix was obtained, as shown in Table 4.

Table 1. Binary effect matrix values.

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15
B1	N	M	E	FL	G	G	VG	M	L	ML	ML	FG	G	L	VL
B2	L	N	ML	VL	VL	VL	G	ML	VL	L	VL	FG	E	FL	MG

B3	E	VG	N	M	L	M	E	E	VL	MG	MG	ML	ML	L	L
B4	VL	ML	FG	N	VL	L	FL	VL	L	L	ML	L	VL	ML	ML
B5	M	FL	M	VL	N	VG	G	VG	VL	ML	VL	FL	L	VL	VL
B6	M	FL	FL	VL	VG	N	G	VG	VL	MG	VL	L	L	VL	VL
B7	G	G	E	M	L	FG	N	VG	L	MG	MG	ML	VG	L	VL
B8	VG	M	E	VL	G	G	G	N	L	L	VL	MG	G	L	VL
B9	VL	M	M	L	FL	FL	L	L	N	L	G	G	VL	FL	ML
B10	FL	M	M	VL	M	M	FG	FG	ML	N	FL	FL	L	VL	VL
B11	VL	M	M	VL	L	L	VG	FL	VL	G	N	VL	VL	FL	ML
B12	G	G	VG	MG	FG	FG	G	VG	L	FG	ML	N	ML	MG	ML
B13	VG	FG	VG	VL	M	M	E	E	L	ML	FL	L	N	ML	VL
B14	VL	M	MG	M	FG	FG	G	G	VL	VL	L	ML	ML	N	L
B15	VL	FG	VG	VL	MG	MG	M	ML	L	L	FL	ML	MG	VL	N

The crisp matrix was normalized to ensure that the row sums are within acceptable bounds. The resulting normalized direct-relation matrix and corresponding identity matrix were used to calculate the total relation matrix. The normalized values and intermediate matrices are presented in Table 5, while the identity matrix and difference matrix are provided in Table 6. The inverse of the results was computed to derive the total relation matrix, as shown in Table 7. Each element in the matrix given in Table 7 reflects both the direct and indirect influence of factor *i* on factor *j*.

Subsequently, for each factor *i*, the sum of row values R_i and sum of column values C_i were computed to determine the influential characteristics of the system. These values are interpreted as:

$R_i + C_i$: Total degree of importance (prominence)

$R_i - C_i$: Net influence (positive indicates a causal factor; negative indicates an effect)

These values are presented in Table 8. Finally, based on the calculated Ri and Ci values, a cause-and-effect diagram was plotted (see Figure 2). Factors with $R_i - C_i > 0$ were identified as causal (influencing), while those with $R_i - C_i < 0$ were deemed effect (influenced) factors. This visual mapping facilitates strategic prioritization in the benefit–risk management of sports and recreational activities.

Table 1. Fuzzy total relationship matrix crisp values (Chen (1992) method).

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15
B1	0.0000	0.5000	0.9545	0.4091	0.6667	0.6667	0.8636	0.5000	0.3333	0.4524	0.4524	0.5909	0.6667	0.3333	0.1344
B2	0.3333	0.0000	0.4524	0.1344	0.1344	0.1344	0.6667	0.4524	0.1344	0.3333	0.1344	0.5909	0.9545	0.4091	0.5476
B3	0.9545	0.8636	0.0000	0.5000	0.3333	0.5000	0.9545	0.9545	0.1344	0.5476	0.5476	0.4524	0.4524	0.3333	0.3333
B4	0.1344	0.4524	0.5909	0.0000	0.1344	0.3333	0.4091	0.1344	0.3333	0.3333	0.4524	0.3333	0.1344	0.4524	0.4524
B5	0.5000	0.4091	0.5000	0.1344	0.0000	0.8636	0.6667	0.8636	0.1344	0.4524	0.1344	0.4091	0.3333	0.1344	0.1344
B6	0.5000	0.4091	0.4091	0.1344	0.8636	0.0000	0.6667	0.8636	0.1344	0.5476	0.1344	0.3333	0.3333	0.1344	0.1344
B7	0.6667	0.6667	0.9545	0.5000	0.3333	0.5909	0.0000	0.8636	0.3333	0.5476	0.5476	0.4524	0.8636	0.3333	0.1344
B8	0.8636	0.5000	0.9545	0.1344	0.6667	0.6667	0.6667	0.0000	0.3333	0.3333	0.1344	0.5476	0.6667	0.3333	0.1344
B9	0.1344	0.5000	0.5000	0.3333	0.4091	0.4091	0.3333	0.3333	0.0000	0.3333	0.6667	0.6667	0.1344	0.4091	0.4524
B10	0.4091	0.5000	0.5000	0.1344	0.5000	0.5000	0.5909	0.5909	0.4524	0.0000	0.4091	0.4091	0.3333	0.1344	0.1344
B11	0.1344	0.5000	0.5000	0.1344	0.3333	0.3333	0.8636	0.4091	0.1344	0.6667	0.0000	0.1344	0.1344	0.4091	0.4524
B12	0.6667	0.6667	0.8636	0.5476	0.5909	0.5909	0.6667	0.8636	0.3333	0.5909	0.4524	0.0000	0.4524	0.5476	0.4524
B13	0.8636	0.5909	0.8636	0.1344	0.5000	0.5000	0.9545	0.9545	0.3333	0.4524	0.4091	0.3333	0.0000	0.4524	0.1344
B14	0.1344	0.5000	0.5476	0.5000	0.5909	0.5909	0.6667	0.6667	0.1344	0.1344	0.3333	0.4524	0.4524	0.0000	0.3333
B15	0.1344	0.5909	0.8636	0.1344	0.5476	0.5476	0.5000	0.4524	0.3333	0.3333	0.4091	0.4524	0.5476	0.1344	0.0000

Table 1. Normalized total relationship matrix.

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15
B1	0.0000	0.0528	0.1008	0.0432	0.0704	0.0704	0.0912	0.0528	0.0352	0.0478	0.0478	0.0624	0.0704	0.0352	0.0142
B2	0.0352	0.0000	0.0478	0.0142	0.0142	0.0142	0.0704	0.0478	0.0142	0.0352	0.0142	0.0624	0.1008	0.0432	0.0578
B3	0.1008	0.0912	0.0000	0.0528	0.0352	0.0528	0.1008	0.1008	0.0142	0.0578	0.0578	0.0478	0.0478	0.0352	0.0352
B4	0.0142	0.0478	0.0624	0.0000	0.0142	0.0352	0.0432	0.0142	0.0352	0.0352	0.0478	0.0352	0.0142	0.0478	0.0478
B5	0.0528	0.0432	0.0528	0.0142	0.0000	0.0912	0.0704	0.0912	0.0142	0.0478	0.0142	0.0432	0.0352	0.0142	0.0142
B6	0.0528	0.0432	0.0432	0.0142	0.0912	0.0000	0.0704	0.0912	0.0142	0.0578	0.0142	0.0352	0.0352	0.0142	0.0142
B7	0.0704	0.0704	0.1008	0.0528	0.0352	0.0624	0.0000	0.0912	0.0352	0.0578	0.0578	0.0478	0.0912	0.0352	0.0142

B8	0.0912	0.0528	0.1008	0.0142	0.0704	0.0704	0.0704	0.0000	0.0352	0.0352	0.0142	0.0578	0.0704	0.0352	0.0142
B9	0.0142	0.0528	0.0528	0.0352	0.0432	0.0432	0.0352	0.0352	0.0000	0.0352	0.0704	0.0704	0.0142	0.0432	0.0478
B10	0.0432	0.0528	0.0528	0.0142	0.0528	0.0528	0.0624	0.0624	0.0478	0.0000	0.0432	0.0432	0.0352	0.0142	0.0142
B11	0.0142	0.0528	0.0528	0.0142	0.0352	0.0352	0.0912	0.0432	0.0142	0.0704	0.0000	0.0142	0.0142	0.0432	0.0478
B12	0.0704	0.0704	0.0912	0.0578	0.0624	0.0624	0.0704	0.0912	0.0352	0.0624	0.0478	0.0000	0.0478	0.0578	0.0478
B13	0.0912	0.0624	0.0912	0.0142	0.0528	0.0528	0.1008	0.1008	0.0352	0.0478	0.0432	0.0352	0.0000	0.0478	0.0142
B14	0.0142	0.0528	0.0578	0.0528	0.0624	0.0624	0.0704	0.0704	0.0142	0.0142	0.0352	0.0478	0.0478	0.0000	0.0352
B15	0.0142	0.0624	0.0912	0.0142	0.0578	0.0578	0.0528	0.0478	0.0352	0.0352	0.0432	0.0478	0.0578	0.0142	0.0000

Table 1. I-D values.

	B1	**B2**	**B3**	**B4**	**B5**	**B6**	**B7**	**B8**	**B9**	**B10**	**B11**	**B12**	**B13**	**B14**	**B15**
B1	1.0000	-0.0528	-0.1008	-0.0432	-0.0704	-0.0704	-0.0912	-0.0528	-0.0352	-0.0478	-0.0478	-0.0624	-0.0704	-0.0352	-0.0142
B2	-0.0352	1.0000	-0.0478	-0.0142	-0.0142	-0.0142	-0.0704	-0.0478	-0.0142	-0.0352	-0.0142	-0.0624	-0.1008	-0.0432	-0.0578
B3	-0.1008	-0.0912	1.0000	-0.0528	-0.0352	-0.0528	-0.1008	-0.1008	-0.0142	-0.0578	-0.0578	-0.0478	-0.0478	-0.0352	-0.0352
B4	-0.0142	-0.0478	-0.0624	1.0000	-0.0142	-0.0352	-0.0432	-0.0142	-0.0352	-0.0352	-0.0478	-0.0352	-0.0142	-0.0478	-0.0478
B5	-0.0528	-0.0432	-0.0528	-0.0142	1.0000	-0.0912	-0.0704	-0.0912	-0.0142	-0.0478	-0.0142	-0.0432	-0.0352	-0.0142	-0.0142
B6	-0.0528	-0.0432	-0.0432	-0.0142	-0.0912	1.0000	-0.0704	-0.0912	-0.0142	-0.0578	-0.0142	-0.0352	-0.0352	-0.0142	-0.0142
B7	-0.0704	-0.0704	-0.1008	-0.0528	-0.0352	-0.0624	1.0000	-0.0912	-0.0352	-0.0578	-0.0578	-0.0478	-0.0912	-0.0352	-0.0142
B8	-0.0912	-0.0528	-0.1008	-0.0142	-0.0704	-0.0704	-0.0704	1.0000	-0.0352	-0.0352	-0.0142	-0.0578	-0.0704	-0.0352	-0.0142
B9	-0.0142	-0.0528	-0.0528	-0.0352	-0.0432	-0.0432	-0.0352	-0.0352	1.0000	-0.0352	-0.0704	-0.0704	-0.0142	-0.0432	-0.0478
B10	-0.0432	-0.0528	-0.0528	-0.0142	-0.0528	-0.0528	-0.0624	-0.0624	-0.0478	1.0000	-0.0432	-0.0432	-0.0352	-0.0142	-0.0142
B11	-0.0142	-0.0528	-0.0528	-0.0142	-0.0352	-0.0352	-0.0912	-0.0432	-0.0142	-0.0704	1.0000	-0.0142	-0.0142	-0.0432	-0.0478
B12	-0.0704	-0.0704	-0.0912	-0.0578	-0.0624	-0.0624	-0.0704	-0.0912	-0.0352	-0.0624	-0.0478	1.0000	-0.0478	-0.0578	-0.0478
B13	-0.0912	-0.0624	-0.0912	-0.0142	-0.0528	-0.0528	-0.1008	-0.1008	-0.0352	-0.0478	-0.0432	-0.0352	1.0000	-0.0478	-0.0142
B14	-0.0142	-0.0528	-0.0578	-0.0528	-0.0624	-0.0624	-0.0704	-0.0704	-0.0142	-0.0142	-0.0352	-0.0478	-0.0478	1.0000	-0.0352
B15	-0.0142	-0.0624	-0.0912	-0.0142	-0.0578	-0.0578	-0.0528	-0.0478	-0.0352	-0.0352	-0.0432	-0.0478	-0.0578	-0.0142	1.0000

Table 1. Inverse matrix values.

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15
B1	1.1393	0.1955	0.2697	0.1173	0.1895	0.2027	0.2651	0.2263	0.1012	0.1636	0.1414	0.1743	0.1965	0.1200	0.0844
B2	0.1393	1.1100	0.1821	0.0714	0.1079	0.1173	0.2017	0.1780	0.0666	0.1209	0.0874	0.1461	0.1950	0.1069	0.1072
B3	0.2354	0.2342	1.1859	0.1277	0.1621	0.1907	0.2795	0.2708	0.0857	0.1755	0.1528	0.1665	0.1852	0.1234	0.1063
B4	0.0942	0.1349	0.1653	1.0467	0.0888	0.1154	0.1501	0.1193	0.0749	0.1051	0.1055	0.1047	0.0935	0.0985	0.0916
B5	0.1596	0.1519	0.1871	0.0713	1.0984	0.1914	0.2042	0.2208	0.0668	0.1360	0.0856	0.1311	0.1368	0.0790	0.0656
B6	0.1573	0.1497	0.1763	0.0698	0.1804	1.1061	0.2016	0.2182	0.0661	0.1431	0.0842	0.1226	0.1351	0.0776	0.0644
B7	0.2094	0.2152	0.2754	0.1268	0.1613	0.1978	1.1868	0.2630	0.1038	0.1750	0.1528	0.1647	0.2194	0.1233	0.0869
B8	0.2183	0.1871	0.2607	0.0867	0.1845	0.1963	0.2372	1.1674	0.0973	0.1447	0.1047	0.1651	0.1910	0.1145	0.0789
B9	0.1102	0.1558	0.1772	0.0888	0.1310	0.1396	0.1639	0.1590	1.0487	0.1196	0.1364	0.1503	0.1078	0.1044	0.1002
B10	0.1444	0.1579	0.1819	0.0699	0.1430	0.1521	0.1926	0.1888	0.0962	1.0879	0.1123	0.1286	0.1320	0.0784	0.0667
B11	0.1063	0.1489	0.1698	0.0650	0.1172	0.1263	0.2058	0.1595	0.0606	0.1455	1.0648	0.0934	0.1057	0.0980	0.0931
B12	0.2134	0.2223	0.2758	0.1359	0.1928	0.2063	0.2597	0.2711	0.1072	0.1842	0.1481	1.1258	0.1873	0.1471	0.1217
B13	0.2259	0.2045	0.2642	0.0910	0.1754	0.1883	0.2751	0.2690	0.1019	0.1628	0.1369	0.1519	1.1347	0.1315	0.0833
B14	0.1248	0.1654	0.1960	0.1099	0.1576	0.1681	0.2088	0.2056	0.0678	0.1077	0.1085	0.1375	0.1505	1.0685	0.0902
B15	0.1261	0.1753	0.2249	0.0733	0.1530	0.1629	0.1941	0.1863	0.0872	0.1282	0.1171	0.1379	0.1593	0.0825	1.0566

Table 1. Cause-Effect values.

Abbreviation	Barriers	Ri	Ci	Ri+Ci	Ri-Ci	Affect
B1	Environmental factors	2.5868	2.4039	4.9907	0.1829	Affecting
B2	User profile and needs	1.9377	2.6085	4.5462	-0.6708	Affected
B3	Human Factor	2.6816	3.1922	5.8738	-0.5105	Affected
B4	Social, Cultural, and Political factors	1.5887	1.3515	2.9403	0.2372	Affecting
B5	Regulatory Factors	1.9857	2.2430	4.2287	-0.2573	Affected
B6	Incomplete or complex legislation, laws and procedures	1.9525	2.4612	4.4137	-0.5088	Affected
B7	Educational Factors	2.6615	3.2262	5.8877	-0.5647	Affected
B8	Technological Factors	2.4344	3.1030	5.5374	-0.6686	Affected
B9	Rules, vision and mission values of sports and activities are carried out by providing facilities, equipment or activity opportunities	1.8927	1.2320	3.1246	0.6607	Affecting
B10	Organizational structure and responsibilities	1.9326	2.0997	4.0323	-0.1670	Affected
B11	Strategy, objectives and principles	1.7598	1.7385	3.4983	0.0213	Affecting
B12	Standards, rules and principles guide, organizational models	2.7985	2.1005	4.8991	0.6980	Affecting
B13	Data, information systems and information flow	2.5965	2.3298	4.9263	0.2667	Affecting
B14	Connections and relationships between internal stakeholders	2.0667	1.5536	3.6203	0.5131	Affecting
B15	Contract terms and obligations	2.0647	1.2969	3.3617	0.7678	Affecting

Within the DEMATEL (Decision Making Trial and Evaluation Laboratory) methodology, the analysis of the fifteen factors (B1-B15) reveals a sophisticated web of cause-and-effect relationships that govern the system's internal dynamics (Table 8). The aggregate value $Ri+Ci$ quantifies the overall prominence of each factor within the system, while Ri-Ci serves to differentiate causal (influencing) elements from effect (influenced) ones. Factors with high Ri-Ci values, such as Contract terms and obligations (0.7678), Standards, rules and principles guide, organizational models (0.6980), Rules, vision and mission values of sports and activities carried out by providing facilities, equipment or activity opportunities (0.6607), Connections and relationships between internal stakeholders (0.5131), and Data, information systems and information flow (0.2667), emerge as primary causal variables. These factors act as

initiators of influence across the network and are particularly important because interventions targeting them are likely to produce wide-ranging systemic effects. They represent strategic leverage points, ideal for proactive decision-making and long-term policy planning.

Conversely, factors such as User profile and needs (-0.6708), Technological Factors (-0.6686), Educational Factors (-0.5647), Human Factor (-0.5105), and Incomplete or complex legislation, laws and procedures (-0.5088) have strongly negative Ri-Ci scores and are positioned as effect variables, meaning they are largely shaped by external influences rather than exerting influence themselves. These nodes function as output indicators within the system, reflecting the cumulative impact of upstream causal factors. While they are not optimal targets for strategic intervention, they are critical for monitoring, diagnostics, and performance tracking.

In addition, a subset of nodes including Environmental factors (0.1829), Social, Cultural, Political factors (0.2372), Organizational structure and responsibilities (-0.1670), Strategy, objectives and principles (0.0213), and Regulatory Factors (-0.2573) display near-zero values for, indicating a balanced role in the system. These factors simultaneously act as mediators, both influencing others and being influenced in return. Their positioning suggests they contribute to feedback regulation and the stabilization of dynamic interactions between dominant cause and effect nodes.

When mapped onto the (R+C, R−C) cognitive plane, these relationships produce a visual distribution that reveals not only the hierarchy of importance (R+C) but also the causal propagation structure (R−C). The result is a comprehensive system view, where cause nodes are seen as points of control, effect nodes as points of observation, and balanced nodes as bridges that modulate system behavior. Therefore, prioritizing high R−C causal nodes for intervention is a scientifically grounded strategy, while effect nodes should be leveraged for response sensitivity and feedback. This combined understanding of influence strength and role orientation forms the foundation for evidence-based, system-level decision-making.

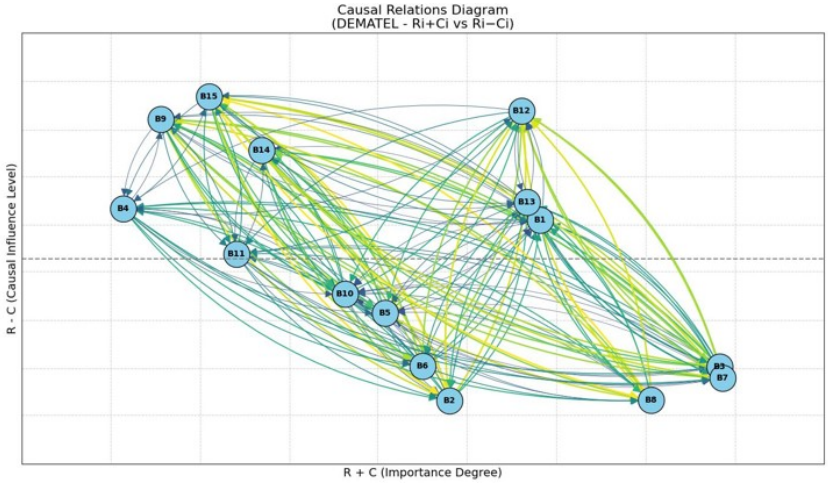


Fig. 1. Cause-Effect diagram.

5 RESULTS AND RECOMMENDATIONS

This paper specifies methods for benefit and risk assessment for sports and recreational activities. It also provides guidance and requirements on benefit–risk assessment within this field. Within sport and recreation, there are many different goals, of which the prevention of injury is but one. Using a benefit–risk approach recognizes the need for making trade-offs in achieving a balance that maximizes overall social utility and public good. These are not the only considerations, though they are important ones. Whether provided on a commercial, not-for-profit, or charitable basis, sport and recreational activities involve an accepted, inherent element of risk and challenge. Taking risks brings rewards but also dangers.

As the sport and recreation sectors have specific characteristics and face different challenges, there is an expectation that the user is making an implicit trade-off between the benefits and the inherent risks of the activity, including the potential for harm. It is up to the user to determine the degree of acceptable risk and the minimum benefits to be achieved regarding the set objectives. Therefore, there is also a requirement for stakeholders, including designers, builders, owners/operators, inspectors, and maintainers of sports and other recreational facilities and equipment, to carry out a risk assessment and, in some cases, to record it. Due to the varied interests of different stakeholders, a single system of hazard identification and risk assessment for the sport and recreation sector is not feasible. This is complicated by the fact that risks and benefits often cannot be measured on a uniform scale. What can be achieved, however, is to identify principles and provide guidance on selecting appropriate techniques.

Every benefit-risk assessment is different, and the context plays an important role. Each benefit–risk assessment shall take into consideration both the opportunities for benefits to a larger group of participants and the risks for that specific product or activity. The risk assessment process described in this study is based on qualitative and

quantitative approaches to risk analysis. Quantitative analysis refers to statistical analysis techniques that rely on numeric data. Whenever possible, facility, service, or product providers should gather numeric data relating to the types, severities, and frequencies of injuries suffered by users so that the data can be used for quantitative analysis.

Qualitative analysis describes a wide range of methods that, while not mathematically rigorous, provide practitioners with tools to facilitate decision-making in the absence of numeric data. These methods can be simple or complex. The selection of an appropriate tool by the practitioner is important, as many available tools come from the occupational health and safety or consumer product safety sectors and may not be suitable for sport and recreation.

In the qualitative approach to risk analysis, including all related equipment and operating procedures, some improvements can be achieved regarding educational factors, human factors, technological factors, environmental factors, and data, information systems, and information flows, as follows:

Human Factors: human error is one of the most common types. Therefore, it is of great importance to reduce human errors that can affect the safety and reliability of the sport and recreation site and its related facilities. Contributing factors include exposure to dangerous situations or stress resulting from making incorrect decisions. Benefit–risk assessments should be conducted by a team of at least three members with relevant expertise in an activity, facility, or product. Teams should be kept to the minimum number of participants necessary to include all relevant expertise. Team membership should be maintained for at least the duration of the first review cycle. The person designated as the team leader should have appropriate training and experience and should be skilled in facilitating effective committee work, including the development of consensus. The team leader should determine the types of expertise needed on the team.

Educational Factors: Where training is required, it should be provided by a competent trainer or training organization. The training program should include documentation of the learning objectives and the scope of the training provided. A certificate may be issued to each trainee following the successful completion of training. Training or experience must enable safe use of the sport or recreation product, site, or activity. Where exposure to hazards is reasonably foreseeable, multiple competencies may be required. Some individuals may have a good awareness of the specific hazards. In contrast, others may have little awareness of those hazards but familiarity with the site, and some may have little awareness of either.

Trained individuals should maintain the currency of their skills and knowledge by attending updated training courses regularly, as appropriate or as required by professional associations. Retraining should be conducted at least once every five years unless a more frequent cycle is determined by a professional organization or the individual’s employer.

A competent person, team, or organization is defined as one with current knowledge, training, skill, education, and experience, and who has successfully demonstrated the ability to identify and resolve issues within the specific scope or subject matter. The competent person is not necessarily required to hold a professional license related to a particular field of study with a defined scope and code of ethics.

Technological Factors: If data suggest a need to correct or adjust the current risk assessment, the overall residual risk should be reevaluated. If the benefit–risk analysis indicates that the benefit does not outweigh the risk, additional control measures must be implemented. In such cases, equipment, activities, or facilities may require design modifications or process adjustments. For equipment already in use, risk control measures must be developed and verified for effectiveness. When user injury data or manufacturer-issued safety notices are available, they should be used to review and update the benefit–risk assessment process at planned intervals to ensure continued effectiveness.

Environmental Factors: Environmental factors can have a wide range of effects that must be considered. Analysts should account for how environmental conditions influence activities or equipment performance, product lifespan, day/night transitions, and how weather changes (rain, snow, heat, cold, etc.) affect user safety and performance.

Data, information systems and information flows: Benefit enhancement involves improving the health and welfare of users. Benefits can be increased through inclusive design that enhances safety for all, including individuals with mobility devices, thereby improving both social benefits and environmental aesthetics. Effective communication with users and consideration of economic impacts are also essential.

Based on the interaction scores of the identified risk factors, and especially the top five critical factors as determined by Ri+Ci values, the following recommendations are proposed to help prevent accidents (referencing [34] and [35]): It must be ensured that there is an experienced and qualified team in the management of the activity and the use of activity areas and related materials.

- Ensure an experienced and qualified team manages the activity areas and associated materials.
- Retraining is recommended at least once every five years, or as determined by professional bodies or certified trainers.
- Define the technical specifications of tools and materials used, ensure proper user training, and verify the effectiveness of risk control measures.
- Adapt materials and procedures to technological developments to reduce risks.
- Operate within the scope of national and international rules regarding environmental and weather conditions.
- Maintain continuous information flow regarding foreseeable hazards, dangerous conditions, and potential environmental harm.
- Utilize effective information systems across all relevant facilities, components, and processes.
- Avoid activities in unsuitable weather unless applicable regulations define mitigation measures.
- Share information related to environmental constraints and ensure local/regional legislation and inspections are implemented.
- Ensure all participants are adequately trained and maintain awareness, especially regarding national/international regulations.
- Use only CE- and ISO-compliant tools and equipment.

- Perform regular inspections and ensure all safety-critical systems meet the legal safety standards.
- Ensure users are appropriately trained, supervised, and accompanied by certified personnel during initial operations.
- Educate users on the impact of fatigue, sleep deprivation, and the dangers of using alcohol, drugs, or medications that impair judgment.
- Review equipment rental agreements in accordance with local standards and ensure compliance with national and international specifications.
- Perform periodic maintenance checks and ensure these are conducted by authorized organizations.
- Develop and implement a comprehensive emergency response plan.

In this study evaluated the results of a comprehensive risk analysis incorporating multiple factors. In particular, the top five most significant risk contributors should be audited by local or regional authorities. Operators must implement the recommendations outlined in this report—especially those related to “Educational Factors,” “Human Factors,” “Technological Factors,” “Environmental Factors,” and “Data, Information Systems, and Information Flow”—by national and international standards. Doing so will improve safety outcomes in individual sports and help prevent accidents or at least mitigate their consequences.

6 CONCLUSIONS

Benefit–risk analysis is a structured and reflective process that makes a significant contribution to the effective operation of sports and recreational activities. Properly conducting a benefit–risk assessment is a crucial prerequisite for achieving organizational goals such as safety, health, welfare, and community cohesion. A competent, multidisciplinary team must carry out this assessment. It requires a thorough understanding of both the principles of benefit–risk evaluation and the specific field of application.

The effectiveness of a benefit–risk assessment should be evaluated using established decision-making frameworks. This evaluation must include verification (e.g., whether the appraisal adequately covers the intended service), validation (e.g., whether all relevant components of the service are considered), and acknowledgment of results (i.e., whether assessors endorse the assessment outcomes).

Verification ensures that the assessment accurately reflects the real-world configuration of the service, including equipment and site-specific conditions. The validity and comprehensiveness of the benefit–risk assessment should be reviewed periodically. The frequency of these reviews must be determined by the responsible parties. Any revisions to the assessment due to new insights or observations should undergo the same level of rigor, control, and documentation as the initial assessment. This includes the identification and evaluation of any new risk control measures, where applicable. Newly observed safety-related issues must be assessed using defined risk criteria.

This paper has demonstrated the successful application of an integrated benefit–risk assessment methodology, combining Fuzzy Set Theory (FST) and the Decision-Making Trial and Evaluation Laboratory (DEMATEL) technique.

Risks and benefits gain significance only about the objectives they serve. In measuring them, any robust benefit–risk analysis should consider at a minimum: injury and lethality risks, improvements in health and well-being (in a broad sense), and compliance with legal requirements. Ultimately, the acceptability of risk and the required minimum benefits are determined by the end-user in relation to the defined objectives.

However, this process is complicated by the fact that risks and benefits often cannot be measured on a common scale. First, risks and benefits may affect different health domains—for instance, physical injury might be accepted in pursuit of a psychological benefit. Second, risks and benefits may impact different individuals—participants often bear the risks, while the benefits may be realized by others, such as future users or the broader community.

This study specifies a methodology for benefit–risk assessment in sports and recreational activities, including facilities, equipment, and injury thresholds. The proposed methodology offers the following advantages:

- It captures both direct and indirect interrelationships among risk factors.
- It quantifies the severity of influence (i.e., $Ri+Ci$ values) between different risk elements.
- It improves the ranking and distinction of failure modes.
- It handles incomplete or imprecise data effectively.
- It delivers more accurate and actionable information for benefit–risk evaluations.

Human error remains one of the most common sources of failure. Contributing factors include exposure to hazardous conditions or stress-induced decision-making errors. A shortage of skilled personnel and insufficient user training are significant safety concerns. Therefore, a mandatory training program addressing regulatory knowledge and proper equipment use is recommended.

Benefit–risk assessments should be conducted by a minimum three-member team with relevant expertise in the activity, facility, or product under review. Team membership should be sustained throughout at least the initial review cycle. The designated team leader must possess experience in benefit–risk assessment and demonstrate competence in facilitating collaborative decision-making. In emergency scenarios, designated safety personnel must be able to lead, assess the situation, and coordinate an adequate response. Hence, human factors and communication play a critical role in ensuring the safety of participants across all ages and abilities.

Sports and recreational activities are designed to serve the public good, which includes physical, mental, and social well-being. Participation inherently involves risk, but risk exposure is not inherently harmful—it can in fact contribute to personal and social development. Thus, critical locations and contexts should be considered when evaluating both benefits and risks, by DEMATEL analysis results.

The benefit–risk analysis revealed that factors such as inspection training, pre-use equipment checks, and observational countermeasures yielded the most substantial

benefits relative to their costs. However, the current safety, environmental performance, and operational efficiency.

For future research, inference rules should be developed to match specific types of risk problems with appropriate decision-making models in fuzzy environments. Additionally, the methodology could be extended to support integrated safety management across interrelated operational divisions.

References

1. Ball-King L. (2022). Benefit-risk assessment: balancing the benefits and risks of leisure, World Leisure Journal, 64(4), s. 383–398.
2. Coleman C. H., (2021). “13 - Risk-Benefit Analysis,” in The Cambridge Handbook of Health Research Regulation, Cambridge: Cambridge University Press, 130–138.
3. Zadeh, L. A., (1965). Fuzzy sets. Information and Control, 8, 338–353.
4. Dubois, D., Prade, H., (1980). Fuzzy Sets and Systems: Theory and Applications, Mathematics in Science and Engineering, vol. 144. Academic Press, Inc., San Diego, New York, Boston, London, Sydney, Tokyo, Toronto.
5. Wang, Y. M., Luo, Y., & Hua, Z. S. (2009). On the extent analysis method for fuzzy AHP and its applications. European Journal of Operational Research, 186(2), 735–747.
6. Bellman, R.E., Zadeh, L.E., (1970). Decision-making in a fuzzy environment. Manage. Sci. 17, 212–223.
7. Yager, R.R., Basson, D., (1975). Decision-making with fuzzy sets. Decision Sci. 6, 590–600.
8. Bass, S.J., Kwakernaak, H., (1977). Rating and ranking of multi-aspects alternatives using fuzzy sets. Automatica 13, 47–58.
9. Zimmermann, H.J., (1996). Fuzzy Set Theory and its Applications. Kluwer, Massachusetts.
10. Lai, YJ., Hwang, CL. (1994). Fuzzy Multiple Objective Decision Making. In: Fuzzy Multiple Objective Decision Making. Lecture Notes in Economics and Mathematical Systems, vol 404. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-57949-3_3
11. Klir, G.J., Yuan, B., (1995). Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice-Hall, Eaglewood Cliffs, NJ.
12. Lootsma, F.A., (1997). Fuzzy Logic for Planning and Decision Making. Kluwer Academic Publisher, Dordrecht.
13. Kahraman, C., Gulbay, M., Kabak, O., (2006). Applications of fuzzy sets in industrial engineering: a topical classification. Studies in Fuzziness and Soft Computing, vol. 201. Springer, Berlin/Heidelberg, pp. 1–55.
14. Feng, F., Li, C. X., Davvaz, B., & Ali, M. (2010). Soft sets combined with fuzzy sets and rough sets: A tentative approach. Soft Computing, 14(9), 899–911
15. Greco, S., Matarazzo, B., & Słowiński, R. (2011). Game-theoretic rough sets: A decision-theoretic rough set model. Fundamenta Informaticae, 101(2), 129–143.
16. Pishvaee, M. S., Razmi, J., & Torabi, S. A. (2012). Robust possibilistic programming for socially responsible supply chain network design under uncertainty. Fuzzy Sets and Systems, 206, 1–20.
17. Liu, D., Liu, Y., & Chen, X. (2019). Fermatean fuzzy linguistic set and its application in multi-criteria decision making. International Journal of Intelligent Systems, 34(8), 1833–1856.

18. Kahraman, C., & Kaya, I. (2020). A fuzzy multicriteria methodology for selection among energy alternatives. Expert Systems with Applications, 157, 113486.
19. Aragonés-Beltrán, P., Pastor-Ferrando, J. P., García-Melón, M., & Montesinos-Valera, J. (2021). An AHP-TOPSIS framework for prioritizing large-scale infrastructure projects under fuzzy environments. Journal of Infrastructure Systems, 27(1), 04020054.
20. Yazdani, M., Chatterjee, P., Zavadskas, E. K., & Turskis, Z. (2022). A fuzzy decision-making framework for COVID-19 vaccine distribution under uncertainty. Computers in Biology and Medicine, 142, 105219.
21. Kou, G., Wu, W., & Yang, H. (2023). Integrating hesitant fuzzy linguistic term sets into risk evaluation in smart cities. Applied Soft Computing, 135, 110035.
22. Zhou, Y., Li, T., & Wang, Y. (2024). A novel interval-valued fuzzy TOPSIS approach for sustainable supplier selection. Expert Systems with Applications, 245, 121421.
23. Chen S.-H., (1999). Ranking Generalized Fuzzy Number with Graded Mean Integration. Proceedings of the 8th International Fuzzy Systems Association World Congress, vol. 2, pp. 899–902, Taipei, Taiwan, Republic of China.
24. Sheng-Li Si, Xiao-Yue You, Hu-Chen Liu, Ping Zhang, (2018). DEMATEL Technique: A Systematic Review of the State-of-the-Art Literature on Methodologies and Applications, Mathematical Problems in Engineering, Volume 2018, Issue 1.
25. Wu, W., Lee, Y.T., (2007). Developing global managers' competencies using the fuzzy DEMATEL method. Expert Syst. Appl. 32 (2), 499–507.
26. Alizadeh, S., Ghazanfari, M., Fathian, M., (2008). Using data mining for learning and clustering FCM. World Acad. Sci. Eng. Technol. 18, 846–853.
27. Chang, B. (2009). Development of a more general risk priority number using fuzzy DEMATEL and ordered weighted geometric averaging. Expert Systems with Applications, 36(3), 6041–6050.
28. Liaw, C.S., Chang, Y., Chang, K.H., Chang, T., (2011). ME-OWA based DEMATEL reliability apportionment method. Expert Syst. Appl. 38, 9713–9723.
29. Wang, W., Lin, Y.H., Lin, C.H., Chung, C.H., Lee, M.T., 2012. DEMATEL-based model to improve the performance in a matrix organization. Expert Syst Appl.
30. Yang, Y.P., Shieh, H.M., Tzeng, G.H., (2013). A VIKOR technique based on DEMATEL and ANP for information security risk control assessment. Inform. Sci. 232, 482–500.
31. Lin X. and Jia W., (2016). A Risk Assessment Method Based on OWGA Operator and Fuzzy DEMATEL Algorithm, Proceedings of the 2nd International Conference on Electronics, Network and Computer Engineering (ICENCE 2016).
32. Mentes, A., Helvacioglu, S., (2022). An integrated methodology for enhancing safety assessment in yacht system design, Ships and Offshore Structures, volume 17, issue 8, 1852–1862.
33. Chen, S.-J., Hwang, C.-L., (1992). Fuzzy Multiple Attribute Decision Making Methods and Applications. Berlin, Springer.
34. BS EN ISO 21853: (2020). Kite Boarding-Release System-Safety Requirements and Test Methods, BSI Standarts Publication.
35. BS ISO 4980: (2023). Benefit-Risk Assessment for sports and recreational facilities, activities and equipment, BSI Standarts Publication.

Evidence to Investigation Report: An IMO-Standards-Grounded Human-in-the-Loop Multi-Agent Framework for Marine Casualty Investigation

Serdar Yıldız[1*[0000-0002-3340-5819]], Özkan Uğurlu[2[0000-0002-3788-1759]], and Barış Barlas[3[0000-0002-5846-2369]]

1 Department of Maritime Transport, Sharjah Maritime Academy, Sharjah, United Arab Emirates
* serdar.yildiz@sma.ac.ae
2 Maritime Transportation and Management Engineering Department, Fatsa Faculty of Marine Sciences, Ordu University, Ordu, Türkiye
3 Piri Reis University, 34940 Istanbul, Türkiye

Abstract. Marine casualty investigation requires the disciplined conversion of heterogeneous evidence into factual reconstruction, causal analysis, safety issues, and recommendations while preserving provenance and uncertainty. Recent maritime artificial intelligence research has focused primarily on corpus-scale knowledge graphs, extraction, and question answering. This study addresses the case level investigation workflow by presenting Marine Casualty Investigation Report Assistant (MCIRA), implemented evidence centred, human-in-the-loop multi-agent investigation-support system grounded in the International Maritime Organization (IMO) Casualty Investigation Code and MSC-MEPC.3/Circ.4/Rev.1. The framework constrains extraction and classification through the 26 chapters of the Code and the five appendices and 30 field-value tables of the harmonized reporting procedure, while retaining source-linked evidence, unresolved fields, and investigator review. A development corpus comprised 65 machine-readable investigation reports containing 2,906 pages; six additional image-only reports were identified as requiring optical character recognition. Regression testing on two reports increased page-linked retrievable evidence passages available to downstream analysis from a previous 12-passage software ceiling to 163 and 180, respectively. These results demonstrate removal of an earlier retrieval bottleneck and broader evidence availability, but not semantic accuracy. MCIRA therefore functions as a working, standards-based support system: it helps organise evidence and prepare the investigation report, while professional judgement and accountability remain with the marine safety investigator.

Keywords: Marine Casualty Investigation, Maritime Safety, Multi-agent Systems, Large Language Models, Human-in-the-loop AI.

1 Introduction

Marine casualty investigation is an evidence-intensive safety function. In practice, the task is not simply to describe an occurrence. Investigators must piece together the sequence of events, identify causal and contributing factors, look beyond the immediate event to wider organisational and system conditions, and consider what safety action may prevent a recurrence. The International Maritime Organization (IMO) Casualty Investigation Code (CIC) sets out this methodical approach and requires the investigation to look beyond the immediate evidence across the wider chain of responsibility [1]. The IMO guidance also calls for structured reporting, a clear separation between factual information and analysis, and explicit treatment of assumptions and uncertainty [2]. Marine casualty and incident data are reported through the Global Integrated Shipping Information System (GISIS) using the revised harmonized reporting procedure in MSC-MEPC.3/Circ.4/Rev.1 [3].

This regulatory framework creates a complex information processing challenge. Evidence relevant to an investigation may be spread across witness statements, statutory and classification records, voyage data recorder material, bridge and engine-room logs, photographs, communications, charts, weather information, technical examinations, emergency-response records, and documents from external authorities. The same fact may appear several times in different forms, while important inconsistencies can be located far apart in the available material. Investigators therefore need to maintain clear links to the original sources while gradually building the chronology, occurrence model, causal analysis, findings, and recommendations.

Artificial intelligence (AI) can help organise and process this material, but it introduces a different concern. A generated sentence may read convincingly even when the underlying information is inferred, uncertain, contradictory, or unsupported. For an investigation, that distinction matters. The aim is therefore not simply to produce text, but to move from evidence to analysis in a controlled way. An investigator must be able to follow a drafted statement back through the case record to the source on which it is based.

Recent maritime research has shown that knowledge graphs, large language models (LLMs), ontology-based constraints, and retrieval-augmented generation can support the extraction and analysis of accident-related knowledge across large collections of reports [4-8]. Most of these studies, however, focus on the accident-report corpus as the main unit of analysis. This study addresses a different but complementary need: supporting an investigator working through the evidence of an individual casualty. The focus is therefore on how AI can improve evidence coverage and reporting consistency while preserving professional judgement and making uncertainty visible.

The study presents the Marine Casualty Investigation Report Assistant (MCIRA) as a design-science artefact developed around this case-level workflow. Its main contribution is an evidence-first, standards-constrained, and human-supervised architecture that connects source material with structured investigation outputs. The paper describes the design and implementation of MCIRA, its demonstration on real casualty cases, the engineering evaluation carried out, and the evidentiary controls incorporated into the system.

2 Regulatory and Research Background

2.1 IMO Casualty Investigation and Harmonized Reporting

The Casualty Investigation Code (CIC) was adopted by IMO through Resolution MSC.255(84). Parts I and II are mandatory under SOLAS regulation XI-1/6, whereas Part III provides recommended practices and explanatory guidance [1]. Across its 26 chapters, the Code covers matters including purpose and definitions, investigation authority, notification and cooperation, evidence collection, reporting, investigation principles, confidentiality, witness protection, and reopening an investigation. Two chapters are particularly relevant to the way MCIRA handles analysis. Chapter 16 requires causal-factor identification to move beyond immediate evidence and consider underlying conditions, while Chapter 22 deals with the collection of evidence [1].

Resolution A.1075(28) provides further guidance for investigators on how the Code should be applied in practice. It encourages deeper questioning during safety analysis and the identification of safety issues that extend beyond the immediate casualty site. It also stresses that final reports should clearly separate factual information from analysis. Where information remains unresolved, assumptions or extrapolations may sometimes be necessary, but these should be identified together with the level of certainty attached to them [2]. For MCIRA, these principles translate into clear design requirements for source provenance, separation of fact from analysis, and visible treatment of uncertainty.

Table 1. Investigation requirements derived from the IMO framework and design implications.

Requirement	IMO basis	Design implication
Evidence traceability	Access to and systematic use of recorded and documentary evidence [1]	Every extracted item retains a source and page reference where available.
Facts separated from analysis	Final reports should distinguish facts from analysis [2]	Descriptive evidence and analytical factors are represented separately.
Visible uncertainty	Assumptions and their certainty should be identified [2]	Unsupported or unresolved fields remain open rather than being silently completed.
Harmonized classification	Five appendices and 30 field-value tables structure GISIS reporting [3]	Events, consequences, factors, and reporting fields are mapped to controlled IMO categories.
Human responsibility	The Code assigns investigation responsibility to States and investigators [1]	AI outputs are suggestions requiring investigator review; automation does not establish findings.

MSC-MEPC.3/Circ.4/Rev.1 provides the GISIS reporting structure used alongside the Code. Its five appendices cover generic information, factual information for each ship, casualty analysis data, supplementary information, and field-value options. Appendix 5 contains Tables 1–30, covering categories such as casualty events, severity,

environmental conditions, accident events, human error, operational and management factors, equipment failure, environmental effects, and external agencies [3]. In this study, these categories are used not simply as a final reporting format, but as structured constraints that guide evidence processing, classification, and analysis throughout the MCIRA workflow. Table 1 summarises the main requirements derived from these IMO instruments.

2.2 AI-Supported Maritime Accident Analysis

Knowledge-graph methods have become an important part of maritime accident analysis. Liu and Cheng [4] constructed MAKG from 581 Chinese maritime accident reports using a maritime-specific named-entity recognition model; the graph contained 16,099 entities and 20,809 relations and supported graph traversal and pattern analysis. Huang et al. [5] subsequently proposed an LLM-based pipeline for automated knowledge extraction from marine accident reports and developed a quality-assessment framework that explicitly addressed structural and semantic quality.

Research has since moved toward richer retrieval and causal analysis. Zou et al. [6] introduced MaRAG, a multi-agent knowledge-graph and retrieval-augmented framework for maritime accident question answering and causal reasoning. Xi and Yin [7] analysed 1,165 IMO GISIS reports using a zero-shot LLM pipeline combined with HFACS-based factor mapping, achieving an F1 score of 0.738 in expert validation. Zhang et al. [8] applied ontology-constrained prompt engineering to 318 ship-collision reports and reported an F1 score of 94.7% for entity recognition, with most relation types exceeding 96%. Taken together, these studies show that domain-specific constraints and structured representations can improve the extraction of knowledge from maritime accident reports.

The problem addressed in the present study, however, is different. Population-level knowledge graphs and question-answering systems are mainly designed to combine and analyse information across multiple cases. An individual marine safety investigation requires a more case-specific approach, including traceability to individual sources, preservation of unresolved evidence, reconstruction of the chronology, application of formal casualty classifications, and controlled development of findings and recommendations. For this reason, the main unit of analysis in MCIRA is the individual investigation case rather than the wider report corpus.

2.3 Research Gap and Questions

The main research gap is the lack of an integrated, case-level system that can bring together evidence from different sources, apply IMO-based classifications, support case-specific event and factor analysis, maintain source traceability, manage uncertainty, and assist with investigation-report drafting while keeping the investigator in control. The study therefore addresses three questions:

1. How can IMO casualty investigation and reporting requirements be incorporated into an AI-assisted investigation workflow as practical system constraints?

2. How effectively can the implemented system turn casualty evidence into structured outputs that remain traceable to their sources?
3. How does the standards-based multi-agent architecture help improve evidence coverage, maintain source traceability, manage uncertainty, and keep the investigator in control of the analysis?

3 Methodology

3.1 Research Design

The study follows a design-science research approach, in which knowledge is generated through the construction and evaluation of an artefact intended to address a defined practical problem [9, 10]. MCIRA is treated as investigation report supporting tool rather than a general-purpose AI tool. The design cycle comprised five linked activities: deriving requirements from IMO instruments and investigation-report practice; constructing the standards and evidence-processing layers; developing specialist analytical roles; demonstrating the artefact on published investigation reports; and evaluating engineering coverage while defining a separate expert-validation stage. Methodology flow diagram is presented in Figure 1.

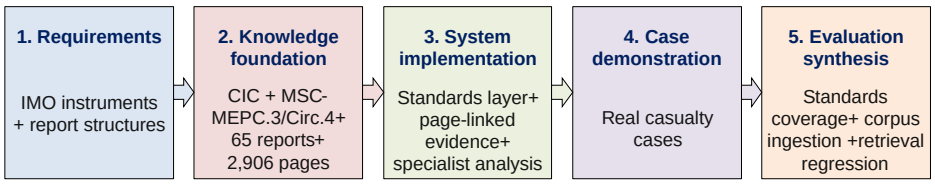


Fig. 1. Methodological sequence used to design, implement and evaluate MCIRA

The engineering evaluation establishes standards coverage, corpus ingestion, removal of a known retrieval bottleneck, and source-linked case modelling. Semantic accuracy is reserved for a separate expert-labelled validation stage.

3.2 Normative Knowledge Base and Development Corpus

The normative layer was constructed from the CIC [1], the investigator guidelines [2], and the revised harmonized reporting procedure [3]. The implemented standards representation covers the Code's 26 chapters and the harmonized procedure's five appendices and Tables 1-30. These instruments play two main roles in the system. They identify the information that needs to be considered during an investigation and provide a consistent basis for classifying events, consequences, environmental conditions, contributing factors, safety actions, and recommendations.

A collection of published marine safety investigation reports was compiled to help the system reflect professional reporting practices, including report structure, terminology, the level of evidentiary detail, and analytical writing style. Machine-

readable 65 reports including 2,906 pages were incorporated into the development process. Six additional image-only reports were identified as requiring optical character recognition (OCR) and were not silently treated as successfully parsed documents. The corpus includes different casualty types and reporting styles. It was used for design knowledge, retrieval behaviour, and regression testing; it is not presented as a statistically representative sample of global casualties. The reports were not used to fine-tune the underlying foundation model. Instead, they helped shape the domain schemas, terminology, standards mapping, retrieval approach, and reporting structure used in MCIRA.

3.3 Evidence-Centred Multi-Agent Architecture

MCIRA is built around an evidence-first approach. Source documents are divided into page-linked evidence units, specialist analytical roles work with the relevant material, IMO standards guide the required fields and classifications, and the resulting outputs are brought together in a structured case model for investigator review (Figure 2).

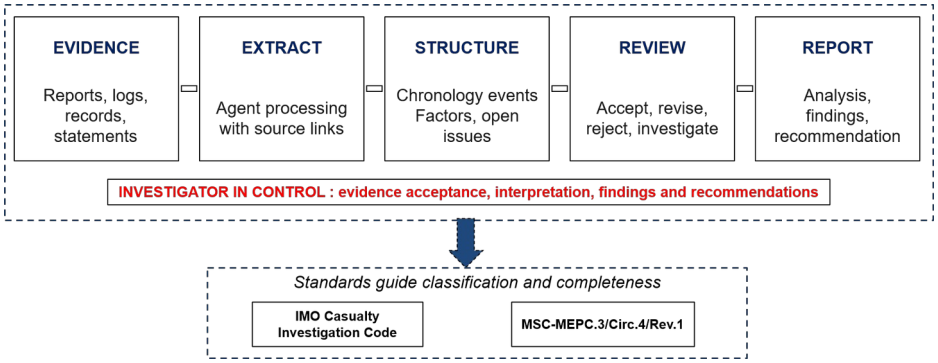


Fig. 2. Conceptual evidence-centred, IMO-constrained investigation-support architecture.

In MCIRA, “multi-agent” describes the way different specialist roles are assigned to different parts of the investigation workflow. It does not mean that the agents are statistically independent. Agreement between agents is therefore not taken as confirmation that an answer is correct. The main safeguards remain traceability to the source and verification by the investigator. The case model also keeps factual evidence separate from interpretation, retains values entered by the investigator, and leaves unresolved or conflicting information open for review.

3.4 Engineering Evaluation Strategy

At this stage, the engineering evaluation does not provide expert-validated measures of semantic accuracy, and the development corpus is not intended to represent the full range of marine casualties worldwide. The retrieval figures indicate how much source material becomes available to later stages of the system; they do not show whether

every passage has been interpreted correctly. Image-only material also remains dependent on OCR quality. In addition, specialist agents that use the same underlying model or shared context may reproduce similar errors. The regression test used the published ATINA and WUZHOU 8 investigation reports, which had exposed a hard limit in an earlier build: only 12 evidence passages were made available downstream. After redesign, the same processing workflow generated 163 and 180 page-linked retrievable passages, respectively. A retrievable passage is a source-linked text unit available for subsequent specialist analysis. The earlier value of 12 was a software ceiling, not a competing model baseline.

4 Results

4.1 Standards and Corpus Coverage

The implemented MCIRA system represented all 26 chapters of the CIC and all five appendices and 30 field-value tables of MSC-MEPC.3/Circ.4/Rev.1. This permits evidence to be organised against formal investigation and reporting concepts before narrative drafting. The development corpus contained 65 machine-readable reports and 2,906 pages. Additional 6 reports were identified as image-only and were explicitly routed to OCR requirement.

The resulting workflow supports the principal information families expected in a professional investigation: occurrence particulars, vessel and voyage information, consequences, environmental conditions, chronology, emergency response, accident events, contributing factors, findings, safety actions, recommendations, and unresolved issues. The report layer turns the reviewed case information into structured investigation text, but the generated wording is not treated as evidence on its own.

Table 2. Results from the engineering evaluation and their interpreted.

Evaluation dimension	Observed result	Interpretation
Machine-readable development corpus	65 reports; 2,906 pages	Broad design corpus for structure, terminology; not accuracy benchmark.
Image-only documents	6 reports flagged for OCR	Image-only reports are flagged for OCR instead of being treated as if no evidence were available.
IMO standards representation	26 Code chapters; 5 harmonized-procedure appendices.	Standards are built into workflow rather than added at drafting stage.
ATINA regression test	163 retrievable passages	Earlier 12-passage software ceiling removed in one regression case.
WUZHOU 8 regression test	180 retrievable passages	Removal of the retrieval ceiling replicated in a second case.
Semantic accuracy	Not claimed in this study	Requires an expert-labelled held-out validation set.

4.2 Extraction Breadth and Evidence Availability

The regression results substantially increased the amount of source material available for later analysis. Marine casualty reports often spread important information across chronology, technical examinations, witness evidence, navigation data, safety-management documentation, emergency-response records, and appendices. A fixed retrieval limit can therefore create a significant problem, as later analytical sections may rely on only a small part of the available evidence even when the full report has been provided. The redesigned workflow removed this limitation in both regression cases and made considerably more source material available for specialist analysis.

4.3 Case-Specific Analysis and Report Development

The analytical layer is designed to work at the level of the individual casualty. Evidence can be linked to the accident-event structure and to the relevant classifications in Appendix 3 and the field-value tables of the harmonized reporting procedure [3]. This allows the case representation to retain the sequence of events, possible contributing factors, supporting source references, and unresolved issues while remaining aligned with the IMO reporting taxonomy. The drafting layer then separates the factual narrative from the evidence methodology, analysis, findings, conclusions, safety issues, lessons, and recommendations. Missing information and uncertainty are kept visible so that the draft does not give a false impression of completeness. Natural-language generation therefore takes place only after the case information has been structured and reviewed. The system is intended to support report preparation rather than independently determine causal findings.

5 Discussion

5.1 Operationalising IMO Standards as System Constraints

The first research question examines how formal IMO requirements can be incorporated into an AI-assisted investigation workflow. MCIRA addresses this in three main ways. First, the CIC provides process requirements relating to areas such as evidence access, depth of causal analysis, reporting responsibility, and confidentiality. Second, Resolution A.1075(28) provides analytical and reporting guidance, particularly on separating facts from analysis and making assumptions explicit. Third, MSC-MEPC.3/Circ.4/Rev.1 provides a structured set of fields and classifications that can be used throughout the investigation process [1-3]. Together, these instruments can be translated into system requirements such as structured fields, controlled categories, validation checks, and review rules rather than being applied only after a report has been drafted. This is consistent with recent research showing that ontology-based constraints can improve maritime information extraction [8]. In MCIRA, however, the purpose goes beyond improving entity recognition. The standards also help define what the system can treat as factual information, an analytical factor, an unresolved issue, or a

reportable category. In this way, the regulatory framework functions both as a domain reference and as a control layer within the investigation process.

5.2 Evidence-to-Report Transformation

The second research question looks at what happens after evidence enters the system. In the two regression cases, removing the retrieval bottleneck meant that many more page-linked passages could reach the later stages of the workflow. Those passages are then used in building the chronology, mapping events and factors, recording unresolved issues, and preparing the report, while the source references remain attached. This traceability is important because a statement in the draft can be followed back through the case record to the extracted passage and then to the original document. The result shown here is therefore broader evidence availability together with a traceable route from evidence to report. It does not show that the multi-agent arrangement is more accurate than a single-agent system.

5.3 Evidence Coverage, Provenance and Investigator Control

The third research question examines how the standards-based multi-agent architecture supports evidence coverage, source traceability, uncertainty management, and investigator control. The specialist roles divide the work into narrower analytical tasks, while the standards layer provides the fields and classifications used across those tasks. Source and page links are retained, and contradictory or unresolved information stays visible instead of being automatically completed. The investigator remains in control for accepting the evidence, interpreting it, making causal judgements, and deciding the findings and recommendations. Confidence labels are only cues for review; they are not treated as proof that an output is correct. Generated text is also produced from the structured case model and remains subject to investigator review rather than becoming evidence in its own right

5.4 Relationship to Existing Maritime AI Research

MCIRA addresses a different part of the problem from most existing maritime accident knowledge-graph research. The approaches reviewed in [4-8] mainly work across large collections of reports to extract information, support reasoning between cases, and identify wider patterns. MCIRA starts instead with the evidence from one casualty and works toward a structured, reviewable case record and investigation report. Once a case has been reviewed, its structured information could also be exported to a cross-case knowledge graph. This would connect evidence-controlled investigation at case level with broader maritime safety analysis.

6 Conclusion

MCIRA was implemented in this study as an evidence-centred, IMO-standards-grounded system to support marine casualty investigation. The system brings together source-linked evidence processing, specialist analysis, IMO-aligned classification, the handling of unresolved or conflicting information, and structured report development while keeping interpretation, findings, and recommendations under investigator control.

Regression testing on two published casualty reports increased the amount of retrievable page-linked evidence from an earlier software ceiling of 12 passages to 163 passages for ATINA and 180 for WUZHOU 8. The change confirms that the earlier retrieval bottleneck was removed and that the later investigation stages could draw on a much wider body of evidence. It does not, however, show that the retrieved material was interpreted correctly.

The main contribution of this study is the development of an implemented and auditable pathway that transforms casualty evidence into structured investigation outputs. MCIRA demonstrates the technical feasibility of using artificial intelligence to support evidence collection, analysis, and report preparation while maintaining a clear distinction between source evidence and professional judgment. Formal expert benchmarking represents the next step in evaluating its investigative accuracy, reliability, and practical applicability.

Despite these contributions, the study has several important limitations. Semantic accuracy has not yet been validated by domain experts, and the development corpus was not designed to represent the full diversity of marine casualties worldwide. Furthermore, retrieval performance indicates the extent to which source material is made available to the system, but it does not guarantee that every item has been interpreted correctly. Reports derived from image-only documents remain dependent on OCR quality, and multiple agents operating on the same underlying model or contextual information may reproduce identical errors. A formal validation phase should therefore employ a held-out set of casualty reports independently assessed by experienced investigators. Such an evaluation should examine field-level precision, recall, and F1-score; source-attribution accuracy; chronology agreement; causal-factor identification; unsupported-claim rates; report completeness; inter-investigator agreement; and the level of investigator intervention required to finalize reports. In addition, a direct comparison between single-agent and multi-agent architectures would help determine whether specialist-agent decomposition provides measurable benefits beyond increased retrieval coverage.

Future research should focus on expert validation studies, the integration of heterogeneous data sources such as sensor records, AIS data, videos and photographs, vessel logs, and the development of explainable AI mechanisms that improve transparency and investigator trust. Further work should also examine the scalability of the framework across different casualty types, regulatory environments, and investigative organizations to establish its generalizability and operational value.

7 References

1. International Maritime Organization: Code of the International Standards and Recommended Practices for a Safety Investigation into a Marine Casualty or Marine Incident (Casualty Investigation Code), Resolution MSC.255(84), as amended by Resolution MSC.390(94). IMO, London (2008, as amended 2014).
2. International Maritime Organization: Guidelines to Assist Investigators in the Implementation of the Casualty Investigation Code (Resolution MSC.255(84)), Resolution A.1075(28). IMO, London (2013).
3. International Maritime Organization: Reports on Marine Casualties and Incidents: Revised Harmonized Reporting Procedures - Reports Required under SOLAS Regulations I/21 and XI-1/6, and MARPOL, Articles 8 and 12, MSC-MEPC.3/Circ.4/Rev.1. IMO, London (2014).
4. Liu, D., Cheng, L.: MAKG: A maritime accident knowledge graph for intelligent accident analysis and management. Ocean Engineering 312, 119280 (2024). https://doi.org/10.1016/j.oceaneng.2024.119280
5. Huang, Y., Yan, R., Zhang, Z.: Automated knowledge extraction from marine accident reports using large language models: Graph construction and evaluation. Ocean & Coastal Management 272, 108015 (2026). https://doi.org/10.1016/j.ocecoaman.2025.108015
6. Zou, Y., Wang, S., Li, G., Ning, B.: MaRAG: A knowledge graph and retrieval-augmented framework for maritime accident analysis. Ocean Engineering 344, 123622 (2026). https://doi.org/10.1016/j.oceaneng.2025.123622
7. Xi, Y., Yin, J.: Knowledge graph construction and causal analysis for maritime safety: An HFACS-guided zero-shot LLM approach. Ocean Engineering 366, 127496 (2026). https://doi.org/10.1016/j.oceaneng.2026.127496
8. Zhang, Y., Fang, Q., Zhen, R., Liu, J., Zheng, Y., Zhang, C.: A novel ship collision accident knowledge graph construction method based on ontology-constrained prompt engineering. Reliability Engineering & System Safety 277, 113148 (2027). Advance online publication (2026). https://doi.org/10.1016/j.ress.2026.113148
9. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Quarterly 28(1), 75-106 (2004). https://doi.org/10.2307/25148625
10. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. Journal of Management Information Systems 24(3), 45-77 (2007). https://doi.org/10.2753/MIS0742-1222240302

A Multi-Objective Spatial Risk Optimization Framework for Safe, Sustainable and Efficient Shipyard Layout Planning

Tolga Aycı[1] [0000-0002-0186-1510] and Barış Barlas[1] [0000-0002-5846-2369]

[1] Piri Reis University, Istanbul, Turkey
tayci@pirireis.edu.tr

Abstract. Shipyards are complex industrial environments where production efficiency, occupational safety, fire protection, and environmental performance are strongly interconnected. Conventional shipyard layout planning primarily focuses on material flow, transportation distances, process efficiency, and facility adjacency, while safety and environmental risks are often assessed separately after the layout has been established. This study proposes a multi-objective spatial risk optimization framework for shipyard facility layout planning that integrates production efficiency, occupational safety, fire risk, and environmental performance within a unified decision-making model. The proposed framework considers the spatial relationships between production activities, material and personnel flows, hazardous materials, hot-work operations, fire-prone areas, emergency routes, and environmentally sensitive processes. Risk assessment methods such as FMEA, Fine-Kinney, Fault Tree Analysis, and Bayesian Network approaches can be incorporated to quantify and prioritize different risk factors. These risk indicators are subsequently integrated into a multi-objective optimization model together with material handling and process-efficiency criteria. The framework aims to generate alternative Pareto-optimal layouts that provide different trade-offs between productivity, safety, fire risk, and environmental performance. The proposed approach is expected to support shipyard planners and safety professionals in identifying safer and more efficient spatial configurations and to provide a systematic basis for integrating risk considerations into early-stage shipyard layout decisions.

Keywords: Shipyard layout planning, multi-objective optimization, occupational health and safety, fire and explosion risk, environmental performance, spatial risk assessment.

1 Introduction

Shipyards are among the most complex industrial production environments due to the simultaneous execution of large-scale manufacturing, assembly, surface treatment, outfitting, lifting, transportation, and repair activities. Unlike conventional

manufacturing facilities, shipyards involve the production and movement of extremely large components, including steel plates, profiles, sub-blocks, ship blocks, machinery, and complete vessel structures. These activities require extensive open areas, specialized workshops, heavy lifting equipment, internal transportation networks, and direct access to waterfront facilities. The overall shipyard layout determines how these activities interact spatially. The relative locations of steel processing, block fabrication, welding, painting, outfitting, storage, and docking areas influence production flow and material-handling requirements. At the same time, these spatial relationships may create occupational hazards, fire and explosion scenarios, and environmental impacts. For example, locating welding and hot-work operations close to flammable-material storage may increase fire risk, while positioning pedestrian routes across heavy-vehicle transportation corridors may increase the likelihood of workplace accidents. Similarly, the location of blasting and painting facilities may affect the exposure of adjacent areas to dust, volatile organic compounds (VOCs), noise, and other environmental pollutants.

Consequently, shipyard layout planning should not be regarded solely as a production engineering problem. It is a strategic spatial planning problem in which operational efficiency, occupational safety, fire protection, and environmental performance must be considered simultaneously. Previous research has addressed several aspects of this problem. Choi et al. (2017) proposed a two-stage shipyard layout optimization approach based on a genetic algorithm and a stochastic growth algorithm, focusing on material-handling costs and department adjacency. Azzolini Junior et al. (2023) developed a sequential optimization procedure combining topological and geometrical layout optimization. More recently, Kafalı (2025) incorporated activity, flow, and risk factors into a shipyard workstation layout methodology using fuzzy logic and simulated annealing. These studies demonstrate that risk-aware shipyard layout optimization is already an established research direction, although there remains scope to extend the analysis to a broader set of spatially integrated performance dimensions.

Occupational accident studies have also highlighted the importance of workplace conditions in shipbuilding. Barlas (2012) analyzed fatal accidents in Turkish shipyards, identifying falls, electrical incidents, fire and explosions, and accidents involving objects or machinery among the major fatality mechanisms. Barlas and Izci (2018) subsequently examined individual and workplace factors associated with fatal occupational accidents among shipyard workers. These findings emphasize the importance of considering not only individual work practices but also the spatial organization of hazardous activities and movement networks. Fire and explosion risk represent another important dimension of shipyard planning. Welding, cutting, gas operations, painting, solvent handling, and the storage of combustible materials may occur within relatively close proximity. Guan et al. (2025) demonstrated the usefulness of Fault Tree Analysis (FTA) and Bayesian Networks (BN) in identifying and evaluating causal factors associated with ship fire and explosion accidents. Although shipboard fire scenarios are not identical to shipyard fire scenarios, the analytical approaches provide useful methodological foundations for examining ignition sources, combustible materials, and causal dependencies.

Environmental performance is similarly influenced by spatial decisions. Shipbuilding and ship repair operations generate blasting residues, paint-related waste, dust, VOC emissions, wastewater, noise, and hazardous waste streams. Qi et al. (2021) reviewed the generation and management of blasting waste in the ship repair industry, while Narci and Dogan (2025) examined environmental performance through life-cycle assessment and process optimization. These studies indicate that environmental considerations should be incorporated into production planning and facility design rather than being treated only as downstream waste-management issues.

In this context, the present study proposes a multi-objective spatial risk optimization framework for the overall shipyard layout. The research aims to integrate four separate performance dimensions—production efficiency, occupational safety, fire and explosion risk, and environmental performance—within a common spatial optimization model. The study seeks to answer this question through a representative hypothetical shipyard model, a spatial risk assessment framework, and a multi-objective optimization approach.

2 Literature Review

2.1 Shipyard facility layout and production efficiency

Shipyard facility layout planning is a complex facility layout problem because production activities involve different spatial requirements, material dimensions, handling equipment, process sequences, and operational relationships. The arrangement of facilities directly affects the distance and time required to transfer materials between production stages. Steel plate storage and processing, for example, should be spatially coordinated with cutting, profiling, panel production, and block fabrication. Similarly, block assembly areas should maintain efficient connections with welding, outfitting, and ship erection activities. Inappropriate spatial arrangements may increase internal transportation requirements, create congestion, and reduce the utilization of cranes, trailers, forklifts, and other handling equipment.

Choi et al. (2017) proposed a two-stage optimization methodology combining genetic algorithms and stochastic growth algorithms to determine shipyard department locations and geometries. Their study considered adjacency and alignment requirements, demonstrating the importance of both topological and geometrical decisions in shipyard layout planning. Azzolini Junior et al. (2023) developed a sequential structure of algorithms involving genetic algorithm-based topological optimization, stochastic growth-based geometrical optimization, and solution refinement. Their work focused on reducing material-handling costs while satisfying required department relationships. Kafalı (2025) extended this research direction by incorporating risk, activity, and flow factors into the preliminary layout of shipyard workstations. The proposed methodology used fuzzy logic to evaluate proximity relationships and simulated annealing to generate alternative arrangements. This study

is particularly relevant because it demonstrates that risk-related factors can be incorporated into shipyard layout optimization.

The present study builds upon these contributions by considering the overall shipyard rather than limiting the analysis to a selected group of workstations or production departments. It also aims to maintain occupational safety, fire risk, and environmental performance as separate objectives alongside production efficiency. Operational validation is another important consideration. Tamer et al. (2025) used ARENA discrete-event simulation to compare existing and optimized shipyard layouts. Their case study reported a reduction in total transfer distance from 4,899 m to 3,837 m and reductions in transfer and total process times. This provides a relevant example of how simulation can be used to assess the production-related implications of alternative shipyard layouts.

2.2 Occupational health and safety in shipyards

Shipyard operations involve several occupational hazards arising from heavy material handling, lifting operations, welding, cutting, working at height, confined spaces, electrical systems, vehicle movement, and exposure to hazardous substances. The spatial distribution of these activities influences the level of worker exposure. For instance, the intersection of pedestrian routes and internal transportation corridors may create vehicle–personnel conflict points. Similarly, locating frequently occupied workstations near crane operating areas or heavy material transfer routes may increase exposure to struck-by and caught-between hazards. Barlas (2012) investigated 115 fatalities in Turkish shipyards between 2000 and 2010 and identified major accident mechanisms, including falls, electric shocks, fire and explosions, and accidents involving objects. Barlas and Izci (2018) examined 126 fatal occupational accidents between 2004 and 2014 and highlighted the influence of individual and workplace factors. Their research also identified enclosed spaces, workshops, and fabrication areas as important accident locations. Şeker et al. (2017) applied a combined DEMATEL and Grey System Theory approach to occupational risk assessment in the Turkish shipbuilding industry. Their study provides a useful methodological basis for identifying relationships among occupational risk factors and prioritizing hazards.

For the proposed layout model, occupational safety should be evaluated through spatial indicators rather than relying exclusively on conventional workplace risk scores. Potential indicators include:

- Number and severity of pedestrian–vehicle conflict points.
- Worker exposure to internal transportation routes.
- Proximity of occupied work areas to crane and lifting operations.
- Interaction between personnel routes and hazardous production activities.
- Exposure to welding, cutting, blasting, and painting operations.
- Accessibility of emergency exits and evacuation routes.
- Separation of pedestrian circulation from heavy equipment movement.

The International Labor Organization's Safety and Health in Shipbuilding and Ship Repair code of practice provides a sector-specific reference for identifying and managing occupational hazards in shipbuilding and repair activities (ILO, 2019).

2.3 Fire and explosion risk

Fire and explosion hazards in shipyards are associated with the interaction of ignition sources, flammable materials, combustible substances, and operational conditions. Welding, gas cutting, grinding, and other hot-work activities may generate ignition sources, while paints, solvents, fuels, gases, and combustible storage materials may provide potential fuel sources. The spatial relationship between these elements is particularly important. For example, locating paint and solvent storage areas near welding or cutting facilities may increase the potential for fire initiation. Similarly, inadequate access for emergency vehicles or insufficient separation between hazardous facilities may increase the consequences of an incident. Guan et al. (2025) combined FTA and Bayesian Network approaches to analyze 271 ship fire and explosion accident investigation reports. Their research identified multiple causal factors associated with combustible materials and ignition sources, demonstrating the value of structured causal analysis for fire risk assessment.

For the proposed shipyard layout model, fire risk should be evaluated by considering:

- The location of ignition sources, including welding, cutting, grinding, and hot-work areas.
- The location and characteristics of combustible and flammable materials.
- The spatial relationship between hazardous-material storage and production facilities.
- Potential fire propagation between adjacent facilities.
- Fire compartmentation and separation measures.
- Access for fire-service vehicles and emergency response teams.
- The availability and accessibility of fire detection and suppression systems.
- The continuity and accessibility of evacuation routes.

The spatial assessment should not rely solely on the distance between facilities. Fire load, material properties, fire barriers, detection and suppression systems, and the characteristics of potential accident scenarios must also be considered. The application of standards such as NFPA 51B (Standard for Fire Prevention During Welding, Cutting, and Other Hot Work), where relevant to the project and jurisdiction, may support the identification of requirements associated with hot-work operations. However, the exact regulatory requirements applicable to the hypothetical shipyard must be defined separately during model development.

2.4 Environmental performance in shipyards

Environmental performance is another important dimension of shipyard layout planning. Shipbuilding and ship repair activities generate various environmental

impacts, including air emissions, solid waste, wastewater, noise, and hazardous substances. Blasting and surface preparation activities may generate significant quantities of solid waste, while painting facilities may produce VOC emissions and hazardous paint-related residues. Welding and cutting may contribute to particulate emissions and energy consumption. In addition, the transportation and storage of hazardous waste may create environmental exposure and operational risks. Qi et al. (2021) reviewed blasting waste generation and management in the ship repair industry, emphasizing the environmental significance of waste quantity, characteristics, and treatment. Narci and Dogan (2025) examined sustainable shipbuilding through life-cycle assessment and process optimization.

From a layout planning perspective, environmental performance can be improved by considering:

- Separation of blasting and painting facilities from sensitive areas.
- Reduction of internal transportation distances for hazardous materials and waste.
- Appropriate positioning of waste collection and temporary storage areas.
- Accessibility of wastewater treatment and environmental-control facilities.
- Separation of clean and potentially contaminated production zones.
- Proximity of emission-generating facilities to administrative, welfare, and other sensitive receptors.
- Reduction of noise exposure through spatial planning.

A distinction should be made between a site-level environmental performance model and a comprehensive life-cycle assessment. The present study proposes a site-level spatial environmental assessment. It does not attempt to evaluate the full life cycle of ships, materials, or equipment. ISO 14040 and ISO 14044 can provide methodological guidance if the environmental assessment is later extended to a formal life-cycle assessment.

2.5 Multi-objective spatial optimization

Facility layout problems typically involve several conflicting objectives. For example, locating two facilities close to each other may reduce material transportation distance but may create safety or environmental conflicts. Conversely, increasing the separation between hazardous activities may improve safety but increase material-handling distances.

Wang et al. (2021) demonstrated the application of multi-objective optimization to industrial park layout design by considering economic and safety performance as separate objectives. Their research is particularly relevant because it illustrates the value of retaining safety as an independent objective rather than converting all safety impacts into economic values. NSGA-II, introduced by Deb et al. (2002), is a widely used evolutionary multi-objective optimization algorithm based on non-dominated sorting and diversity preservation. It is suitable for problems in which multiple

competing objectives must be evaluated without reducing them to a single weighted objective. For the proposed study, NSGA-II is used to generate alternative shipyard layouts representing different trade-offs among production efficiency, occupational safety, fire risk, and environmental performance. A recent review by Bordbar et al. (2026) also highlights the increasing use of simulation, optimization, artificial intelligence, and digital twin approaches in shipyard production planning. This provides a broader context for the proposed integration of spatial optimization and production simulation.

3 Research Gap and Contribution

The reviewed literature demonstrates that shipyard layout optimization, occupational risk assessment, fire and explosion analysis, environmental assessment, and production simulation have all been investigated through different methodological approaches. Existing shipyard layout studies have primarily focused on material-handling efficiency, department adjacency, production flow, and workstation arrangement. Some recent studies, particularly Kafalı (2025), have also incorporated selected risk-related factors into shipyard layout optimization. Therefore, the research gap should not be defined as the absence of risk considerations in previous shipyard layout studies. Instead, the present study identifies an opportunity to extend existing approaches through the simultaneous spatial integration of four distinct performance dimensions:

1. Production efficiency.
2. Occupational health and safety.
3. Fire and explosion risk.
4. Environmental performance.

The proposed contribution is based on the following aspects:

- Development of a representative spatial model covering the overall shipyard layout.
- Integration of production facilities, support areas, hazardous activities, material flows, personnel movements, and environmental impact sources.
- Translation of occupational, fire, and environmental risk information into spatial indicators.
- Formulation of four separate objective functions within a multi-objective optimization model.
- Implementation of NSGA-II to generate alternative Pareto-optimal layouts.
- Use of production simulation, sensitivity analysis, and expert assessment to evaluate the resulting alternatives.

The main contribution is therefore not the introduction of risk into shipyard layout optimization for the first time. Rather, it is the proposed integration of production efficiency, occupational safety, fire risk, and environmental performance as separate and traceable objectives within an overall shipyard spatial planning framework.

4 Proposed Conceptual Framework

4.1 Framework architecture

The proposed framework is given in Table 1 and consists of six interconnected stages such as shipyard definition, spatial modelling, risk and performance assessment, mathematical modelling, optimization and validation.

Table 1. Proposed framework architecture.

Stage	Main Activities	Expected Output
Shipyard definition	Define site boundary, facilities, production processes and operating assumptions	Representative shipyard specification
Spatial modelling	Define facility footprints, routes, hazard locations and receptors	Spatial database
Risk and performance assessment	Calculate production, OHS, fire and environmental indicators	Objective-function inputs
Mathematical modelling	Define decision variables, objective functions and constraints	Optimization model
Optimization	Apply NSGA-II and evaluate candidate layouts	Pareto-optimal alternatives
Validation	Simulation, sensitivity analysis and expert assessment	Evaluated layout alternatives

The study considers the overall shipyard layout, including production facilities, support facilities, storage areas, environmental-control infrastructure, and emergency access systems.

1. Steel plate storage and processing: This group includes steel plate and profile storage, cutting, profiling, preparation, and related handling areas. These facilities are important origins of material flow toward panel production and block fabrication.
2. Block fabrication and assembly: This group includes panel lines, sub-block production, block assembly, and related fabrication areas. The spatial relationship between these facilities and steel processing areas is important for reducing material transfer distances.
3. Welding and hot-work areas: These facilities include welding, cutting, grinding, and other operations involving ignition sources. Their location must be evaluated in relation to combustible storage, hazardous materials, personnel circulation, and emergency access.
4. Painting and surface preparation facilities: This group includes blasting, surface preparation, coating, and drying facilities. These areas are associated with dust, VOC emissions, hazardous residues, and potentially flammable substances.

5. Outfitting and equipment assembly areas: These facilities include machinery installation, piping, electrical systems, equipment assembly, and other outfitting operations. Their location should support the movement of equipment and completed blocks toward the shipbuilding area.
6. Docks, slipways, and shipbuilding areas: These facilities represent the principal vessel construction and assembly areas. Their locations are generally constrained by waterfront access, site geometry, and existing infrastructure. Therefore, they may be treated as fixed or partially fixed elements in the optimization model rather than freely relocatable facilities.
7. Hazardous-material storage areas: These areas include storage for gases, fuels, solvents, paints, chemicals, and other hazardous materials. Their locations must satisfy applicable separation, containment, access, and fire-protection requirements.
8. Waste-management and environmental-control facilities: This group includes hazardous waste storage, blasting waste collection, wastewater treatment, waste segregation, and other environmental-control infrastructure.
9. Administrative and welfare buildings: These facilities include offices, personnel facilities, canteens, changing rooms, and other buildings used by administrative and operational personnel. Their location should consider personnel accessibility, noise, emissions, and separation from hazardous activities.
10. Emergency access and evacuation infrastructure: This group includes emergency vehicle routes, fire-service access, evacuation routes, assembly areas, and emergency exits. These elements should be represented as protected spatial networks and may also be incorporated into the layout decision variables where appropriate.

4.2 Representative hypothetical shipyard model

Since the initial research stage does not assume access to a complete operational dataset from a specific shipyard, the proposed methodology uses a representative hypothetical shipyard model. The model is developed using literature-based information, engineering assumptions, and logical production relationships. It represents a simplified but technically consistent shipyard environment in which facility locations, dimensions, production flows, transportation routes, and risk indicators can be evaluated.

The hypothetical model includes below areas:

- A defined site boundary and available land area.
- Facility dimensions and required spatial footprints.
- Fixed and relocatable facility classifications.
- Production sequences and inter-facility material flows.
- Internal transportation and pedestrian movement networks.
- Hazardous-material storage locations.
- Hot-work and fire-prone activity zones.

- Environmental emission and waste-generation sources.
- Administrative and welfare areas.
- Emergency access and evacuation infrastructure.

All assumed input values are documented in a data-assumption register. The register should identify each variable, its assumed value or range, unit, rationale, available literature source, and uncertainty. For example, facility dimensions and flow quantities may be defined through engineering assumptions, while risk indicators may be developed using structured expert-based scoring scales. These values are used to test the proposed methodology and must not be interpreted as actual measurements or risk levels for an operating shipyard. A two-dimensional coordinate-based model is considered sufficient for the initial optimization stage. GIS or BIM integration may be introduced if the complexity of the site geometry or the availability of spatial information justifies a more advanced representation.

4.3 Spatial data requirements

The spatial model should include both geometric and operational information. Geometric information includes facility footprints, site boundaries, road networks, pedestrian routes, emergency exits, storage areas, and environmental receptors. Operational information includes material flow quantities, production sequences, handling equipment requirements, personnel movements, operating frequencies, and hazardous activities. Risk-related information includes hazard sources, exposure zones, fire scenarios, combustible material locations, emergency response infrastructure, and environmental impact sources. The integration of these datasets is enabling the spatial model to evaluate the consequences of alternative facility arrangements across all four objectives.

5 Risk and Performance Assessment

5.1 Production efficiency assessment

Production efficiency is primarily evaluated through material-handling distance and transportation burden between facilities. The material flow matrix defines the amount of material transferred between each pair of facilities. The distance matrix represents the actual or estimated transportation distance along the internal route network. Where detailed operational data are unavailable, flow quantities are assigned using hypothetical production scenarios and documented engineering assumptions. The production assessment should consider:

- Material transfer distance.
- Material transfer frequency.
- Handling equipment requirements.
- Production sequence and process dependencies.
- Potential transportation congestion.
- Facility utilization, where simulation data are available.

The proposed framework is given in Table 1 and consists of six interconnected stages such as shipyard definition, spatial modelling, risk and performance assessment, mathematical modelling, optimization and validation.

5.2 Occupational safety assessment

Occupational risk is assessed by identifying spatial interactions between workers, vehicles, equipment, and hazardous activities. FMEA and Fine-Kinney assessment may be used to identify and prioritize hazards. DEMATEL and Grey System Theory may also be considered where expert-based causal relationships among occupational hazards need to be evaluated. However, the risk scores obtained from these methods must be translated into spatially meaningful indicators before being incorporated into the optimization model. Potential spatial indicators include:

- Pedestrian–vehicle crossing frequency.
- Personnel exposure to heavy material transportation routes.
- Proximity of workstations to crane operating areas.
- Worker exposure to hot-work zones.
- Proximity of occupied facilities to hazardous-material storage.
- Accessibility of emergency exits.

Risk assessment should consider severity, exposure frequency, and the characteristics of the hazardous activity. Distance alone should not be treated as a sufficient measure of occupational risk.

5.3 Occupational safety assessment

Fire and explosion risk is assessed by identifying potential interactions between ignition sources, fuel sources, exposed facilities, and emergency infrastructure. Potential fire scenarios may include:

- Welding and cutting operations near combustible storage.
- Painting and solvent handling areas exposed to ignition sources.
- Gas cylinder storage near hot-work operations.
- Fire propagation between adjacent production facilities.
- Restricted emergency access to fire-prone areas.
- Inadequate separation between hazardous-material storage and occupied buildings.

FTA and Bayesian Networks may be used to identify causal relationships and evaluate fire scenarios. Where reliable quantitative probabilities are unavailable, expert-based ordinal or fuzzy scoring methods may be adopted. The fire-risk assessment should also consider protective measures, including fire barriers, detection systems, suppression systems, separation distances, and emergency access.

5.4 Environmental performance assessment

Environmental performance is evaluated through selected site-level indicators associated with production activities and spatial relationships. The assessment may include:

- Dust and particulate emissions from blasting and surface preparation.
- VOC emissions from painting and coating operations.
- Noise exposure in adjacent occupied areas.
- Hazardous waste generation and transportation.
- Wastewater generation and treatment accessibility.
- Environmental exposure associated with material and waste storage.

The environmental assessment is focusing on the spatial implications of these activities. It is not representing a complete life-cycle assessment of ship production unless the research scope is explicitly expanded.

6 Multi-Objective Optimization Model

6.1 Problem definition

The shipyard layout problem is formulated as a multi-objective spatial optimization problem. Let:

- I represent the set of shipyard facilities.
- L represent the set of candidate facility locations.
- F_{ij} represent the material flow between facilities i and j.
- D_{ij} represent the transportation distance between facilities i and j.
- X_i represent the assigned location of facility i.
- A_i represent the required area of facility i.

The optimization process aims to determine feasible facility arrangements that minimize four independent objective functions. Facilities whose locations are physically constrained, such as docks and slipways, may be treated as fixed elements. The optimization variables primarily determine the locations and orientations of facilities that can reasonably be relocated or rearranged.

6.2 Objective 1: Production efficiency

The first objective is to minimize the total material-handling burden between shipyard facilities.

$$\min f_1 = \sum_{i \in I} \sum_{j \in I} F_{ij} D_{ij} C_{ij}$$

where F_{ij} material flow between facilities i and j. D_{ij} transportation distance between facilities i and j. C_{ij} handling cost or movement-weight factor. When reliable cost data are unavailable, the model may instead minimize the flow-weighted transportation

distance. The transportation distance should preferably be calculated using the internal transportation network rather than straight-line Euclidean distance. This is particularly important in shipyards where material movement is constrained by road networks, crane coverage, waterfront access, and large-scale handling equipment.

6.3 Objective 2: Occupational health and safety

The second objective is to minimize spatially related occupational risk exposure. A preliminary formulation is:

$$\min f_2 = \sum_{k \in K} W_k E_k R_k$$

where W_k severity or importance weight of occupational hazard category k. E_k exposure or interaction intensity, R_k assessed risk level. The exposure component may be derived from pedestrian–vehicle interactions, proximity to lifting operations, worker movement through hazardous zones, and the spatial relationship between occupied facilities and hazardous production activities. For a more detailed model, occupational risk can be calculated using pairwise spatial interactions between hazard sources and exposed personnel or facilities. The formulation should avoid counting the same exposure more than once.

6.4 Objective 3: Fire and explosion risk

The third objective is to minimize fire and explosion risk associated with the spatial arrangement of hazardous activities and materials. A preliminary formulation is:

$$\min f_3 = \sum_{m \in M} P_m C_m S_m$$

where P_m likelihood or estimated frequency of fire scenario m. C_m potential consequence of scenario m. S_m spatial exposure or fire propagation factor. The spatial factor may consider separation distance, facility adjacency, fire load, combustible material characteristics, protective barriers, detection and suppression systems, and emergency access. If reliable probability data are unavailable, the model may use structured ordinal or fuzzy risk scores. Such scores should be clearly identified as assumed or expert-assessed values and examined through sensitivity analysis.

6.5 Objective 4: Environmental performance

The fourth objective is to minimize the environmental impact associated with the spatial arrangement of shipyard facilities. A preliminary formulation is:

$$\min f_4 = \sum_{n \in N} V_n I_n$$

where V_n magnitude or frequency of environmental impact source n. I_n spatial impact or receptor-exposure factor. The environmental objective may incorporate normalized indicators for dust, VOC emissions, noise, wastewater, and hazardous waste transportation. Since these indicators have different measurement units, normalization

is required before they are combined into an environmental performance index. The selected normalization procedure and any weighting factors must be explicitly justified.

6.6 Constraints

The optimization model has to generate layouts that satisfy mandatory regulatory, engineering, and operational requirements. The principal constraints are expected to include:

1. Geometrical constraints
    a. Facilities must remain within the defined shipyard boundary.
    b. Facility footprints must not overlap.
    c. Required facility dimensions and areas must be respected.
    d. Fixed facilities must remain in their designated locations.
2. Operational constraints
    a. Production sequence and process relationships must be maintained.
    b. Required adjacency and accessibility relationships must be satisfied.
    c. Crane, trailer, forklift, and other handling equipment access must be preserved.
    d. Internal transportation networks must remain operationally feasible.
3. Occupational safety constraints
    a. Pedestrian routes must be separated from hazardous vehicle routes where required.
    b. Required access to emergency exits must be maintained.
    c. Mandatory safety clearances and restricted zones must be respected.
    d. Hazardous activities must comply with applicable safety requirements.
4. Fire and explosion safety constraints
    a. Mandatory fire separation requirements must be satisfied.
    b. Hazardous-material storage must comply with applicable requirements.
    c. Fire-service access routes must remain available.
    d. Emergency exits and evacuation routes must remain unobstructed.
    e. Required fire protection and compartmentation conditions must be maintained.
5. Environmental constraints
    a. Applicable environmental separation and containment requirements must be satisfied.
    b. Hazardous waste storage and handling requirements must be respected.
    c. Wastewater and environmental-control infrastructure must remain accessible.
    d. Required environmental protection measures must be maintained.

An important methodological principle is that mandatory regulatory and technical requirements must be treated as hard constraints. The optimization algorithm must not

be allowed to violate these requirements in exchange for shorter transportation distances or higher production efficiency.

6.7 NSGA-II optimization approach

The Non-dominated Sorting Genetic Algorithm II (NSGA-II) is proposed as the main optimization algorithm. The algorithm generates an initial population of candidate shipyard layouts. Each candidate layout is evaluated using the four objective functions and the defined feasibility constraints. The optimization process involves:

1. Generation of an initial population of candidate layouts.
2. Evaluation of the four objective functions.
3. Identification of non-dominated solutions.
4. Application of selection, crossover, and mutation operators.
5. Generation of new candidate layouts.
6. Evaluation of feasibility and objective-function performance.
7. Iterative improvement of the population.
8. Extraction of the final Pareto-optimal solution set.

The primary advantage of this approach is that it does not require the four objectives to be converted into a single weighted score. Instead, the algorithm provides a set of alternative layouts representing different trade-offs. The algorithm's population size, crossover and mutation probabilities, stopping criteria, and constraint-handling method is determined during implementation. Its convergence and solution diversity also needs to be evaluated. The selection of NSGA-II is a proposed methodological decision and should not be presented as a validated result before computational implementation.

6.8 Pareto-optimal solution analysis

The optimization process is expected to produce several non-dominated layout alternatives. A layout is considered Pareto-dominated if another feasible layout performs at least as well in every objective and better in at least one objective. The Pareto front provide information about the trade-offs between:

- Production efficiency and occupational safety.
- Production efficiency and fire risk.
- Production efficiency and environmental performance.
- Occupational safety and environmental performance.
- Fire risk and production efficiency.
- The combined performance of all four objectives.

The final selection should be based on transparent decision criteria and stakeholder assessment rather than automatically selecting the layout with the lowest value in only one objective.

7 Proposed Research Methodology and Validation

The research is conducted through the following seven stages:

- Stage 1 Definition of the representative shipyard: The overall shipyard boundary, facility inventory, production activities, fixed infrastructure, and operating assumptions are defined.
- Stage 2 Development of the hypothetical dataset: Facility dimensions, material flows, process relationships, workforce movements, hazard indicators, and environmental impact factors are defined using literature-based information and engineering assumptions.
- Stage 3 Spatial modelling: A two-dimensional coordinate-based model is developed to represent the shipyard facilities, transportation routes, pedestrian circulation, hazard sources, and emergency infrastructure.
- Stage 4 Risk and performance assessment: Production efficiency, occupational safety, fire risk, and environmental performance indicators are defined and calculated for the baseline layout.
- Stage 5 Mathematical model development: Decision variables, objective functions, geometrical requirements, operational relationships, and mandatory safety constraints are formulated.
- Stage 6 Multi-objective optimization: NSGA-II is implemented to generate alternative feasible shipyard layouts and identify the Pareto-optimal solution set.
- Stage 7 Validation and sensitivity analysis: Candidate layouts are evaluated through production simulation where feasible, expert assessment, and sensitivity analysis.

7.1 Production simulation using discrete-event simulation

Discrete-event simulation (DES) may be used to evaluate the production-related performance of selected layouts. The simulation model may represent the movement of materials through steel processing, panel production, block fabrication, outfitting, and ship assembly stages. The principal performance indicators may include:

- Total material transfer distance.
- Material transfer time.
- Production throughput.
- Equipment utilization.
- Waiting time.
- Transportation congestion.
- Work-in-process levels.

ARENA may be considered as a suitable simulation platform. Tamer et al. (2025) provide a relevant example of ARENA-based shipyard layout comparison and validation using production data. However, production simulation alone cannot validate

occupational or fire risk. These dimensions require separate assessment through risk analysis, fire engineering review, and expert evaluation.

7.2 Sensitivity and robustness analysis

Since the proposed model is based on hypothetical data, sensitivity analysis is particularly important. The analysis should examine how changes in assumed input values influence the optimization results. Potential parameters include:

- Material flow quantities.
- Facility dimensions.
- Transportation frequencies.
- Occupational risk scores.
- Fire scenario likelihoods.
- Fire consequence estimates.
- Environmental impact factors.
- Facility separation requirements.
- Workforce and vehicle movement patterns.

The purpose is to determine whether the preferred layouts remain acceptable under different plausible assumptions. Where possible, scenario-based analysis may be used to represent different production conditions, such as high production intensity, increased block movement, or changes in painting and blasting activity.

7.3 Expert assessment

The resulting Pareto-optimal layouts should be reviewed by experts from relevant disciplines, including:

- Shipyard production and planning.
- Occupational health and safety.
- Fire protection and fire engineering.
- Environmental management.
- Facility layout and industrial engineering.

Expert assessment helps identify impractical arrangements, omitted hazards, operational conflicts, and potentially unrealistic model assumptions.

8 Expected Contributions and Discussion

The proposed research is expected to contribute to shipyard layout planning in several ways. First, it provides a spatial decision-support framework that integrates production efficiency, occupational safety, fire risk, and environmental performance within a common model. Second, it extends conventional shipyard facility layout planning by incorporating occupational, fire, and environmental indicators directly into the optimization process rather than evaluating these dimensions only after the layout has been established. Third, it maintains four independent objective functions and use

Pareto-based analysis to make the trade-offs between competing performance dimensions transparent. Fourth, it establishes a methodological approach for developing and testing shipyard layout optimization models even when complete operational datasets from a real shipyard are unavailable.

From a practical perspective, the framework may support shipyard planners in comparing alternative facility arrangements before major infrastructure investments or production planning decisions are finalized. It may also help safety and environmental professionals identify spatial conflicts at an earlier stage of the design process. Nevertheless, the practical reliability of the framework depends on the quality of the assumed data, the validity of the risk indicators, and the adequacy of the mathematical model. The hypothetical shipyard model should therefore be regarded as an initial methodological test environment rather than a substitute for validation using real operational data.

9 Conclusions

This study proposes a multi-objective spatial risk optimization framework for safe, sustainable, and operationally efficient shipyard layout planning. The framework considers the overall shipyard environment, including production facilities, support areas, hazardous-material storage, environmental-control infrastructure, and emergency access systems. It integrates four independent performance dimensions: production efficiency, occupational safety, fire and explosion risk, and environmental performance. These dimensions are represented through separate objective functions and evaluated within a common spatial optimization model. A representative hypothetical shipyard model is developed using literature-based inputs and transparent engineering assumptions. NSGA-II is proposed as the optimization algorithm for generating Pareto-optimal layout alternatives, while discrete-event simulation, sensitivity analysis, and expert assessment are identified as complementary methods for evaluating the resulting solutions.

The central proposition of this research is that shipyard layout decisions should not be evaluated solely in terms of material flow and production efficiency, but should also account for their spatial implications for worker safety, fire protection, and environmental performance. By integrating these dimensions within a single decision-support framework, the proposed methodology provides a systematic basis for developing safer, more sustainable, and operationally efficient shipyard layouts.

The present study is conceptual and methodological in nature and does not report an implemented optimization model or empirical results from an actual shipyard. The reliance on hypothetical data represents an important limitation. Although literature-based inputs and engineering assumptions provide a basis for developing the framework, they cannot fully capture the operational complexity and site-specific conditions of a real shipyard. Furthermore, the proposed objective functions require

further operationalization. Occupational safety and fire risk indicators may involve uncertainty and subjective assessments, while environmental indicators may vary according to production technologies, material characteristics, operating conditions, and local environmental requirements.

Another important limitation is the dynamic nature of shipyard operations. Production activities, facility utilization, and transportation requirements may change throughout the different stages of shipbuilding. Steel preparation, block fabrication, outfitting, painting, and final assembly, for example, may require different spatial arrangements and material transportation patterns. Consequently, a static layout optimization approach may not adequately represent the changing operational requirements of a shipyard over time.

Future research should therefore focus on applying the proposed framework to a real shipyard case study and integrating actual production, transportation, and occupational safety data. Further development may include dynamic layout optimization across different shipbuilding phases, GIS- and BIM-based spatial data integration, more detailed fire and explosion scenario modelling, and evacuation and emergency response simulations. Environmental performance could also be evaluated through expanded life-cycle assessment (LCA) methods. In addition, integration with digital twin technologies and real-time production monitoring systems could improve the framework's ability to reflect changing operational conditions.

A particularly important direction for future research is the development and validation of a dynamic multi-objective layout optimization model that accounts for changing production requirements, spatial constraints, and risk conditions throughout the shipbuilding process. Such an approach would extend the proposed conceptual framework towards a practical decision-support tool for shipyard planning and management.

References

1. Ayci, T., Barlas B, (2016). Characteristics of Non-Fatal Occupational Accidents in Shipyards, 1st International Congress on Ship and Marine Technology, Istanbul, Turkey.
2. Azzolini Junior, W., Azzolini, F. G. P., Mundim, L. R., Porto, A. J. V., & Amani, H. J. S. (2023). Shipyard facility layout optimization through the implementation of a sequential structure of algorithms. Heliyon, 9(6), e16714. https://doi.org/10.1016/j.heliyon.2023.e16714
3. Barlas, B. (2012). Shipyard fatalities in Turkey. Safety Science, 50(5), 1247–1252. https://doi.org/10.1016/j.ssci.2011.12.037
4. Barlas, B., & Izci, F. B. (2018). Individual and workplace factors related to fatal occupational accidents among shipyard workers in Turkey. Safety Science, 101, 173–179. https://doi.org/10.1016/j.ssci.2017.09.012
5. Bordbar, A., Tadros, M., Nazemian, A., Aung, M. Z., Georgoulas, K., Louvros, P., & Boulougouris, E. (2026). Digitalisation of shipyard production planning: A review of

simulation, optimisation, AI, and digital twin methods (2010–2025). Journal of Marine Science and Engineering, 14(4), 396. https://doi.org/10.3390/jmse14040396
6. Choi, M., Kim, S. H., & Chung, H. (2017). Optimal shipyard facility layout planning based on a genetic algorithm and stochastic growth algorithm. Ships and Offshore Structures, 12(4), 486–494. https://doi.org/10.1080/17445302.2016.1176294
7. Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multi-objective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), 182–197. https://doi.org/10.1109/4235.996017
8. Fine, W. T. (1971). Mathematical evaluations for controlling hazards. Journal of Safety Research, 3(4), 157–166.
9. Guan, W., Zhang, C., Dong, C., & Xia, Y. (2025). Ship fire and explosion accident statistical analysis based on fault tree and Bayesian network. Fire Safety Journal, 153, 104358. https://doi.org/10.1016/j.firesaf.2025.104358
10. International Electrotechnical Commission. (2018). IEC 60812:2018: Failure modes and effects analysis (FMEA and FMECA). IEC.
11. International Labour Organization. (2019). Safety and health in shipbuilding and ship repair (Revised ed.). International Labour Office.
12. International Organization for Standardization. (2006). ISO 14040:2006: Environmental management—Life cycle assessment—Principles and framework. ISO.
13. International Organization for Standardization. (2006). ISO 14044:2006: Environmental management—Life cycle assessment—Requirements and guidelines. ISO.
14. International Organization for Standardization. (2018). ISO 31000:2018: Risk management—Guidelines. ISO.
15. Kafalı, M. (2025). A fuzzy logic and optimisation based approach for shipyard facility layout problem with consideration of risk, activity, and flow factors. Brodogradnja, 76(1), 76103. https://doi.org/10.21278/brod76103
16. Kinney, G. F., & Wiruth, A. D. (1976). Practical risk analysis for safety management (NWC Technical Publication 5865). Naval Weapons Center.
17. Narci, A. O., & Dogan, E. C. (2025). Sustainable shipbuilding: Enhancing environmental performance through life cycle assessment and process optimization. Process Safety and Environmental Protection, 202, Article 107712. https://doi.org/10.1016/j.psep.2025.107712
18. National Fire Protection Association. (2024). NFPA 51B: Standard for fire prevention during welding, cutting, and other hot work. NFPA.
19. Qi, C., Weinell, C. E., Dam-Johansen, K., & Wu, H. (2021). A review of blasting waste generation and management in the ship repair industry. Journal of Environmental Management, 300, 113714. https://doi.org/10.1016/j.jenvman.2021.113714
20. Şeker, Ş., Recal, F., & Başlıgil, H. (2017). A combined DEMATEL and grey system theory approach for analyzing occupational risks: A case study in Turkish shipbuilding industry. Human and Ecological Risk Assessment: An International Journal, 23(6), 1340–1372. https://doi.org/10.1080/10807039.2017.1308815
21. Tamer, S., Barlas, B., Gunbeyaz, S. A., & Eren, S. (2025). Shipyard layout simulation using ARENA: A case study. Ships and Offshore Structures. Advance online publication. https://doi.org/10.1080/17445302.2025.2569431
22. Wang, R., Wang, Y., Gundersen, T., Wu, Y., Feng, X., & Liu, M. (2021). A multi-objective optimization method for industrial park layout design: The trade-off between economy and safety. Chemical Engineering Science, 235, 116471. https://doi.org/10.1016/j.ces.2021.116471

23. A comprehensive review of static and dynamic facility layout problems. (2024). Annual Reviews in Control, 58, 100970. https://doi.org/10.1016/j.arcontrol.2024.100970

Multi-Objective Optimization of Marine Current Turbine Blade Geometry for Enhanced Hydrodynamic Efficiency and Structural Strength

Tunahan Şık[1[0009-0008-7752-2583]], Engin Kanun*[1,2[0000-0002-2369-5322]], İsmail Kahraman[1[0000-0003-2858-1567]], Selahattin Özsayan[1[0000-0003-2727-2076]], Mehmet Delikan[1[0009-0009-5244-4731]], Feyza Ertarhanacı[1[0009-0009-8644-1950]], Cihad Çelik[1[0000-0002-2335-0455]], Burcu Erbaş Öz[1[0000-0002-3559-1264]], Uğur Oral Ünal[1[0000-0001-5435-1138]]

[1] Faculty of Naval Architecture and Ocean Engineering, Istanbul Technical University, Istanbul, Türkiye
[2] Maritime Faculty, Mersin University, Mersin, Türkiye
kanun@itu.edu.tr

ABSTRACT

Marine current energy is a predictable and dense renewable resource. The Istanbul Bosphorus is a promising site for its use, with persistent currents driven by the level and density difference between the Black Sea and the Sea of Marmara. Within the scope of the project “Electricity Generation from Bosphorus Currents Using Turbines: Toward the Electrification of Maritime Transport in Istanbul.”, the study presents the hydrodynamic and structural design of a 6 m diameter, three-bladed, fixed-pitch horizontal-axis marine current turbine. The project investigates the feasibility of harnessing the kinetic energy of Bosphorus currents to support the electrification of Istanbul’s maritime transport. Unlike wind turbines, marine current turbines face high loads, cavitation-limited tip speeds and, without pitch control, severe runaway loading. The design chain is integrated in ITUKATBLADE, an in-house software combining a blade element momentum solver, neural-network airfoil polars, cavitation and stall assessment, structural sectional and beam analysis, and NURBS blade generation for CFD and manufacturing. Airfoil sections were first optimized with multi-objective evolutionary algorithms, yielding a blade with high efficiency and wide off-design performance. However, structural analysis showed that this hydrodynamically optimal blade could not carry the runaway loads in solid stainless steel. Airfoil shape, thickness and design lift distributions were therefore co-optimized with the blade under structural constraints. The resulting design satisfies the structural criteria and is 9.3% more efficient than a structurally feasible version of the previous design. The study shows that blade geometry and material must be determined together.

Keywords: Marine current turbine, Blade–airfoil co-optimization, Structural design.

1. INTRODUCTION

Marine currents are a predictable source of renewable energy. Since sea water is about 840 times denser than air, a current of 1 m/s has approximately the power density of a 10 m/s wind. Current turbine rotors are therefore small in comparison with wind turbines of the same rating, but their blades are heavily loaded. In fixed-pitch rotors the blades cannot be feathered. If the generator load is lost, the rotor accelerates to more than twice its design speed, and this runaway condition usually determines the strength of the blade.

The hydrodynamic design of current turbines is based on the momentum and blade element theories developed for wind turbines [1–5], with cavitation as an additional limit on the tip speed [6]. Cavitation tunnel and towing tank tests have shown that blade element momentum (BEM) analysis predicts rotor performance with good accuracy when reliable section data are available [7–9]. Section data are generally obtained from viscous panel codes such as XFOIL [10]. NeuralFoil [11], a neural-network model trained on XFOIL results, gives a complete polar in a few milliseconds, which makes it practical to treat the section shape as a design variable of the blade.

This paper presents the design of a 6 m diameter rotor developed within the scope of the project “Electricity Generation from Bosphorus Currents Using Turbines: Toward the Electrification of Maritime Transport in Istanbul.” The project investigates the feasibility of harnessing the kinetic energy of Bosphorus currents to support the electrification of Istanbul’s maritime transport.

In most blade design studies, the sections are selected or optimized first and the structure is checked afterwards. The blade obtained from hydrodynamic considerations could not carry the runaway loads in solid stainless steel, and strengthening it with the same section caused a considerable loss of efficiency. The section shape, thickness ratio and design lift coefficient were therefore optimized together with the blade under structural constraints, and the procedure was repeated for three blade materials. All results given here were obtained with the design models of the in-house program ITUKATBLADE. Verification of the final designs with CFD and with a commercial finite element code is left to a subsequent study.

2. DESIGN CONDITIONS

The current in the Bosphorus is driven by the level and density difference between the Black Sea and the Sea of Marmara. According to the measurements made within the project, the velocity is nearly uniform over the rotor disc at the planned depth. The rotor was designed for the annual mean speed, and its strength was checked at the highest measured speed (Table 1). Three load cases were considered at the maximum speed: normal operation, runaway at zero shaft torque, and parked rotor. In addition, the blade should not cavitate at its top position and should operate with an adequate margin to stall, since a fixed-pitch rotor is kept near its design point only by speed control.

Table 1. Design conditions.

Quantity	Value
Rotor diameter / hub diameter / number of blades	6 m / 0.6 m / 3 (fixed pitch)
Design current speed (annual mean) / maximum speed	1.1 m/s / 1.5 m/s
Depth of rotor axis / tip submergence at top position	10 m / 7 m
Sea water density / vapour pressure	1025 kg/m³ / 1.7 kPa
Reynolds number of blade sections at design point	$\approx 10^6$

3. METHOD

The design steps are combined in ITUKATBLADE (Fig. 1). Each analysis in the program takes a few seconds at most, so that the structural check can be used as a constraint during the optimization.

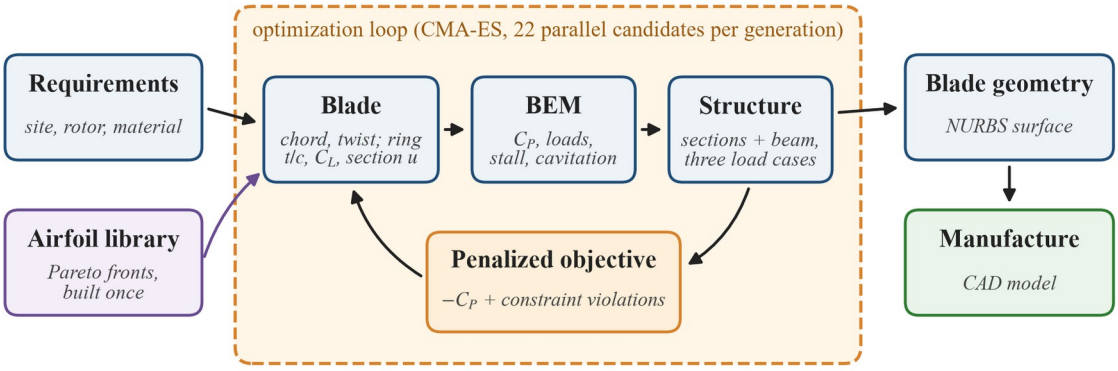


Fig. 1. Design procedure implemented in ITUKATBLADE.

3.1 Hydrodynamic model

Rotor performance is calculated with BEM theory, including the tip and hub loss factors of Prandtl [2] and the high-induction correction of Buhl [12]. The formulation of Ning [13] is used, in which the equations reduce to a single residual in the inflow angle φ,

$$R(\varphi) = \frac{\sin\varphi}{1-a(\varphi)} - \frac{\cos\varphi}{\lambda_r\,[1+a'(\varphi)]} = 0 \tag{1}$$

which can be solved reliably with a bracketing method. Lift, drag, moment and minimum pressure coefficients are taken from NeuralFoil tables as functions of angle of attack, Reynolds number and thickness ratio. The initial chord and twist distributions are obtained from the Schmitz optimum, which accounts for wake rotation [3, 4]:

$$\varphi = \tfrac{2}{3}\arctan\tfrac{1}{\lambda_r}, \qquad c = \frac{16\pi r}{B C_{L,d}}\sin^2\tfrac{\varphi}{2}, \qquad \theta = \varphi - \alpha_d \tag{2}$$

For a given design tip-speed ratio λ_d, the chord is thus inversely proportional to the design lift coefficient $C_{L,d}$. Cavitation is assessed with the quasi-steady criterion used in [8]. The cavitation margin M_{cav} is the minimum of $\sigma + C_{p,min}$ along the span, where σ is the local cavitation number, and it should be positive. The stall margin is defined as the smallest difference between the maximum and the operating lift coefficient along the blade. The blade sections are finally interpolated with periodic cubic B-splines to form a closed NURBS surface [14] for CAD and manufacturing.

3.2 Structural model

The section properties (stiffnesses, shear and torsion centres, mass) are computed from the section contour, either as a multi-cell thin-walled box or as a solid section discretized with two-dimensional finite elements. The blade is modelled as a cantilever beam with shear deformation and restrained warping [15–17]. Hydrodynamic loads from BEM, centrifugal force, weight and buoyancy are applied, and a dynamic factor of 1.2 is used. For metals the von Mises stress is limited to 25% of the yield strength, whereas composites are checked with partial safety factors. The result is expressed as the utilization $\sigma_{eq}/\sigma_{allow}$, which should not exceed unity at any section.

3.3 Optimization

The optimization problem is stated as

$$\max_{\mathbf{x}} \; C_P(\mathbf{x}) \qquad \text{subject to} \qquad \max_{r,\ \text{load cases}} \frac{\sigma_{eq}}{\sigma_{allow}} \le 1, \qquad \min_r \left(C_{L,max} - C_L\right) \ge 0.3 \tag{3}$$

where **x** contains the design tip-speed ratio, the design lift coefficient, the thickness ratio at each BEM annulus and the section shape. Section shapes are described with the CST parameterization of Kulfan [18]. When the CST coefficients of five stations were used directly as variables, the problem had 92 variables, and very few feasible candidates were obtained in trial runs. For this reason the section problem is solved beforehand. On a grid of about 1000 combinations of thickness ratio, Reynolds number and target lift coefficient, a two-objective problem is solved with SMS-EMOA [19]. The objectives are L/D at the target lift and the flapwise bending stiffness, and the NACA 63_3-618 section from the UIUC database [20] is used as the starting geometry (Fig. 2). The Pareto fronts are stored as a section library. During blade optimization, the section of each annulus is taken from this library by means of a single parameter u, which varies from 0 (highest L/D) to 1 (highest stiffness). In this way the number of variables is reduced to 12.

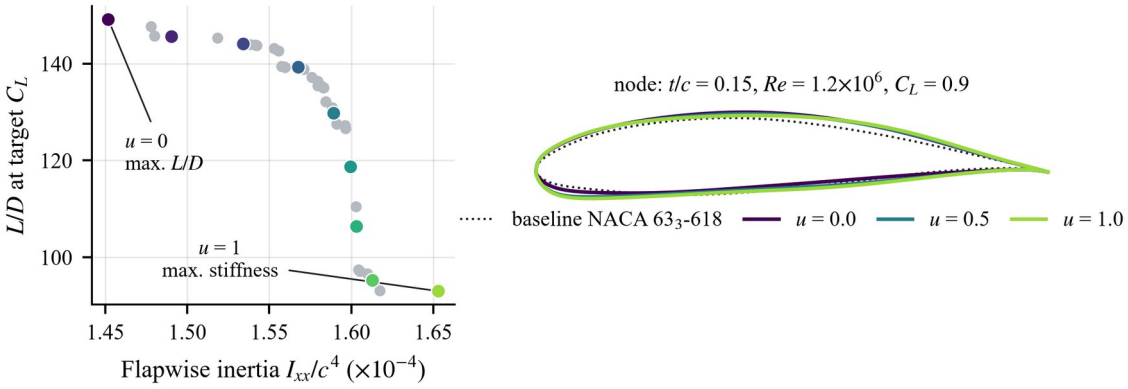


Fig. 2. Example node of the section library: Pareto front between L/D and bending stiffness, and three sections on the front.

Since a single annulus has little influence on rotor power, the annulus values of the thickness ratio tend to oscillate along the span during the optimization. The spanwise distributions are therefore smoothed with a Whittaker smoother [21] before each evaluation. Constraints are handled with the feasibility rules of Deb [22], and CMA-ES [23] is used as the optimizer. One optimization campaign required about 10^5 blade evaluations, which corresponds to approximately six hours on 22 processor cores.

4. RESULTS

4.1 First design

The first blade was designed in the usual order. Among the sections frequently used for turbine blades (NACA 44xx, 63- and 64-series, DU 06-W-200, FX 63-137), the Wortmann FX 63-137 has the highest maximum L/D at the design Reynolds number. This section was thickened to 18% and re-optimized with SMS-EMOA in order to recover its L/D range and to limit its pitching moment. The resulting section is referred to as FX-O18. The Schmitz blade with this section and $\lambda_d = 5.5$ has $C_P = 0.482$, which is 81% of the Betz limit [1], and C_P remains above 90% of its maximum between $\lambda \approx 4.3$ and 8. The cavitation margin is 7.3. A parametric study showed that increasing λ_d from 5.5 to 7.5 raises C_P by only about 1%, while the cavitation margin is more than halved and the chord is reduced by about one quarter. The value $\lambda_d = 5.5$ was therefore retained.

The structural analysis of this blade is summarized in Table 2. Runaway is the governing load case for all constructions. The highest utilization occurs in the middle and outer part of the span, where the chord and thickness are small, rather than at the root. Only the solid carbon/epoxy construction satisfies the criterion. In solid stainless steel 316 the allowable stress is exceeded by a factor of 4.6.

Table 2. First design at the maximum current speed: largest utilization over the span and load cases.

Construction	Utilization	Blade mass (kg)	Result
Box beam, unidirectional carbon caps	1.1	23	not feasible
Box beam, unidirectional glass caps	2.1	29	not feasible
Solid carbon/epoxy	0.7	86	feasible
Solid glass/epoxy	1.3	105	not feasible
Solid stainless steel 316	4.6	448	not feasible

According to Eq. (2), the blade can be strengthened by reducing the design lift coefficient, which increases the chord, and by increasing the thickness ratio. This was done systematically for the solid steel blade (Fig. 3). Each gray curve corresponds to one thickness ratio with varying design lift. The blade becomes stronger as the design lift is reduced, but the efficiency falls, because the highly cambered FX-O18 then operates away from its best L/D. The best feasible blade obtained in this way has $C_P = 0.447$, which is 7% lower than the initial design, and a mass of 766 kg.

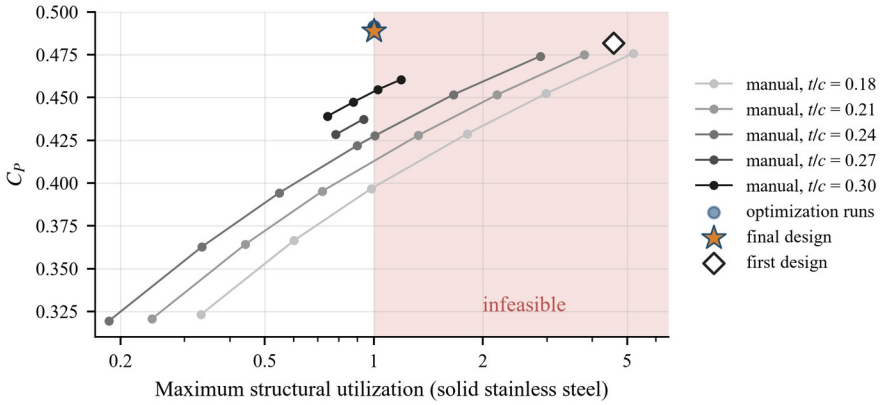


Fig. 3. Solid steel blade: manual modification of the first design (gray curves) and the co-optimized design (star).

4.2 Optimized blade in stainless steel

The blade optimized for solid stainless steel satisfies the structural criterion, and its power coefficient is slightly higher than that of the first design (Table 3). Compared with the first design, the design tip-speed ratio is higher, the design lift coefficient is about half, and the maximum chord is 21% larger. The FX-O18 is replaced by library sections that have their best L/D at the lower lift coefficient. The sections are relatively thin near the root, where the chord is large, and the thickness ratio is highest around mid-span (Fig. 5).

Table 3. First and optimized designs in solid stainless steel 316.

Quantity	First design	Optimized	Change
Design tip-speed ratio / rotor speed (rpm)	5.5 / 19.3	6.2 / 21.7	
Power coefficient C_P	0.482	0.489	+1.5%
Thrust (kN) / torque (kN·m)	14.5 / 4.6	14.5 / 4.1	0% / −10%
Maximum chord (m)	0.72	0.87	+21%
Stall margin / cavitation margin	0.51 / 7.3	0.92 / 6.3	
Largest utilization	4.6	1.0	
Blade mass (kg)	448	489	+9%
Tip deflection in runaway (mm)	73	22	
First flapwise frequency in water (Hz)	11.2	11.0	

The utilization along the span is shown in Fig. 4. In the runaway case the utilization of the optimized blade is close to unity between $0.3R$ and $0.8R$, although a uniform stress level was not imposed. This indicates that the thickness was increased only where it was required by strength. The stall margin is nearly doubled, and the first flapwise natural frequency in water is about 7 times the blade-passing frequency. The cavitation margin decreases from 7.3 to 6.3 because of the higher tip speed, but it is still large. Compared with the best manually modified blade of Section 4.1, the optimized blade is 9.3% more efficient and 36% lighter.

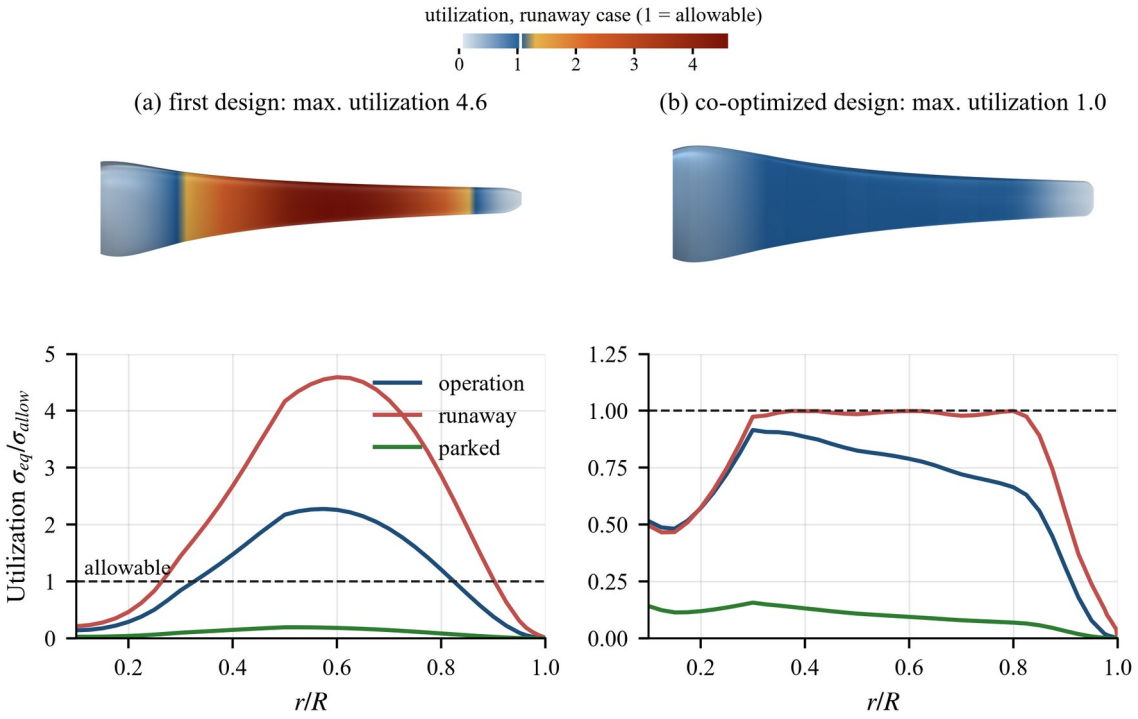


Fig. 4. Utilization of the solid steel blade at the maximum current speed: (a) first design, (b) optimized design. Top: runaway case; bottom: spanwise distribution for the three load cases.

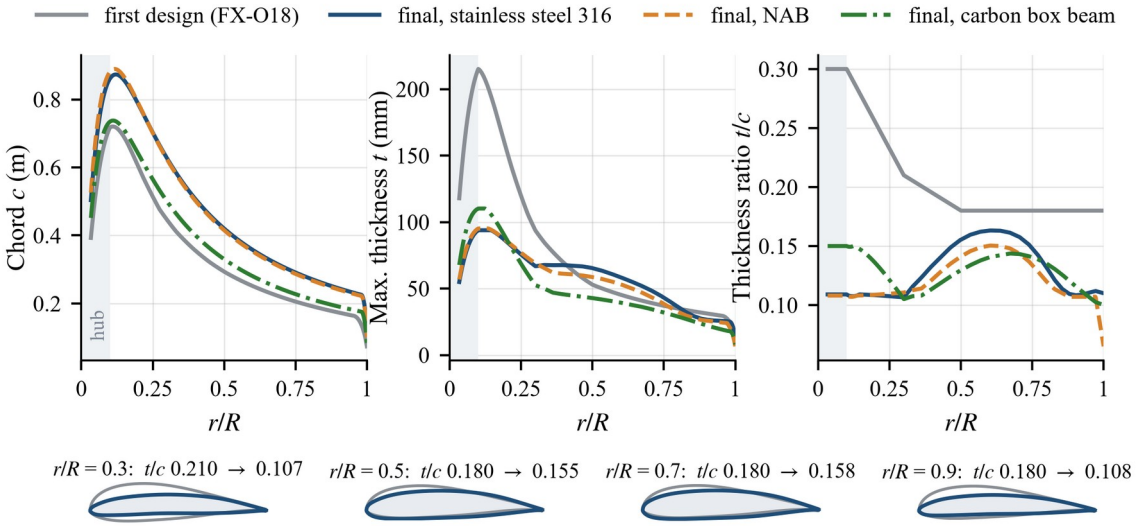


Fig. 5. Chord, maximum thickness and thickness ratio of the first design and of the blades optimized for three materials; bottom: sections of the first (gray) and the optimized steel (blue) blade.

4.3 Effect of blade material

The optimization was repeated for nickel-aluminium bronze (NAB) and for a box beam with unidirectional carbon caps (Fig. 6, Table 4). NAB has a higher allowable stress than stainless steel, and the blade optimized for it has thinner sections and a higher design tip-speed ratio. For the carbon box beam the strength constraint is less restrictive. The blade is more slender, and its mass is about one twentieth of that of the metal blades. However, its tip deflection in runaway reaches 14% of the radius, and its first flapwise frequency is only 2.9 times the blade-passing frequency. Stiffness was not constrained in the present formulation, and for such light constructions a deflection or frequency constraint would be necessary. The power coefficients of the three blades are close to each other ($C_P \approx 0.49$). The choice of material therefore depends on mass, stiffness, manufacturing and corrosion resistance rather than on efficiency.

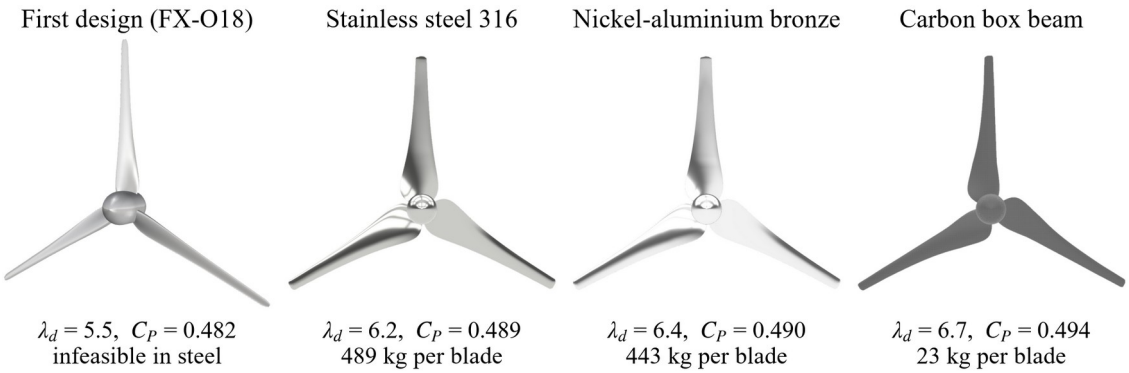


Fig. 6. First design and the rotors optimized for three blade materials.

Table 4. Blades optimized for three materials.

Quantity	Stainless steel 316	NAB	Carbon box beam
Construction / allowable stress (MPa)	solid / 51	solid / 61	box beam / strain-based
Design tip-speed ratio / design lift coefficient	6.20 / 0.64	6.42 / 0.61	6.74 / 0.70
Maximum chord (m)	0.87	0.89	0.74
Power coefficient C_P	0.489	0.490	0.494
Stall margin / cavitation margin	0.92 / 6.3	0.87 / 5.9	0.77 / 5.3
Blade mass (kg)	489	443	23
Tip deflection in runaway (mm)	22	42	430
First flapwise frequency in water (Hz)	11.0	8.4	4.7

5. CONCLUSIONS AND FUTURE WORK

The hydrodynamic and structural design of a 6 m diameter, fixed-pitch marine current turbine for the Istanbul Bosphorus was presented. The main findings are as follows:

1. The blade designed from hydrodynamic considerations alone has C_P = 0.482, but in solid stainless steel the allowable stress is exceeded by a factor of 4.6. Runaway is the governing load case, and the critical region is the mid-span rather than the root.
2. When this blade is strengthened by changing the design lift and the thickness with the same section, C_P decreases by 7%.
3. When the section shape, thickness and design lift are optimized together with the blade under the structural constraint, a feasible steel blade is obtained whose C_P is 1.5% higher than that of the first design. It is 9.3% more efficient than the manually modified blade. The pre-computed section library reduces the number of variables from 92 to 12 and makes this optimization practical.
4. The optimum blade depends on the material. The stainless steel, NAB and carbon box beam blades have different geometries but nearly the same efficiency. Blade geometry and material should therefore be determined together.

The results of this paper are based on BEM theory with surrogate section polars and on a beam model of the blade. In the next stage of the work, the final designs will be verified with CFD analyses and with three-dimensional finite element analyses in commercial software. Fatigue and buckling analyses, an annual energy estimate based on the site measurements, and model tests in the ITUKAT Cavitation Tunnel are also planned.

ACKNOWLEDGEMENTS

This study was supported by the Scientific Research Projects Coordination Unit of Istanbul Technical University (ID: MGA-2026-47415).

REFERENCES

[1] Betz, A. (1920). Das Maximum der theoretisch möglichen Ausnützung des Windes durch Windmotoren. Zeitschrift für das gesamte Turbinenwesen, 26, 307–309.
[2] Glauert, H. (1935). Airplane propellers. In: Durand, W.F. (Ed.), Aerodynamic Theory, Vol. IV, Division L, 169–360. Springer, Berlin.
[3] Schmitz, G. (1955/56). Theorie und Entwurf von Windrädern optimaler Leistung. Wissenschaftliche Zeitschrift der Universität Rostock, 5, 379–391.
[4] Manwell, J.F., McGowan, J.G., Rogers, A.L. (2009). Wind Energy Explained: Theory, Design and Application, 2nd ed. Wiley, Chichester.

[5] Burton, T., Jenkins, N., Sharpe, D., Bossanyi, E. (2011). Wind Energy Handbook, 2nd ed. Wiley, Chichester.
[6] Bahaj, A.S., Batten, W.M.J., McCann, G. (2007). Experimental verifications of numerical predictions for the hydrodynamic performance of horizontal axis marine current turbines. Renewable Energy, 32(15), 2479–2490.
[7] Bahaj, A.S., Molland, A.F., Chaplin, J.R., Batten, W.M.J. (2007). Power and thrust measurements of marine current turbines under various hydrodynamic flow conditions in a cavitation tunnel and a towing tank. Renewable Energy, 32(3), 407–426.
[8] Molland, A.F., Bahaj, A.S., Chaplin, J.R., Batten, W.M.J. (2004). Measurements and predictions of forces, pressures and cavitation on 2-D sections suitable for marine current turbines. Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment, 218(2), 127–138.
[9] Batten, W.M.J., Bahaj, A.S., Molland, A.F., Chaplin, J.R. (2008). The prediction of the hydrodynamic performance of marine current turbines. Renewable Energy, 33(5), 1085–1096.
[10] Drela, M. (1989). XFOIL: An analysis and design system for low Reynolds number airfoils. In: Mueller, T.J. (Ed.), Low Reynolds Number Aerodynamics, Lecture Notes in Engineering 54, 1–12. Springer, Berlin.
[11] Sharpe, P., Hansman, R.J. (2025). NeuralFoil: An airfoil aerodynamics analysis tool using physics-informed machine learning. arXiv:2503.16323.
[12] Buhl, M.L. (2005). A new empirical relationship between thrust coefficient and induction factor for the turbulent windmill state. NREL/TP-500-36834, National Renewable Energy Laboratory, Golden, CO.
[13] Ning, S.A. (2014). A simple solution method for the blade element momentum equations with guaranteed convergence. Wind Energy, 17(9), 1327–1345.
[14] Piegl, L., Tiller, W. (1997). The NURBS Book, 2nd ed. Springer, Berlin.
[15] Timoshenko, S.P. (1921). On the correction for shear of the differential equation for transverse vibrations of prismatic bars. Philosophical Magazine, 41(245), 744–746.
[16] Vlasov, V.Z. (1961). Thin-Walled Elastic Beams, 2nd ed. Israel Program for Scientific Translations, Jerusalem.
[17] Benscoter, S.U. (1954). A theory of torsion bending of multicell beams. Journal of Applied Mechanics, 21(1), 25–34.
[18] Kulfan, B.M. (2008). Universal parametric geometry representation method. Journal of Aircraft, 45(1), 142–158.
[19] Beume, N., Naujoks, B., Emmerich, M. (2007). SMS-EMOA: Multiobjective selection based on dominated hypervolume. European Journal of Operational Research, 181(3), 1653–1669.
[20] Selig, M.S. UIUC Airfoil Coordinates Database. University of Illinois at Urbana-Champaign. https://m-selig.ae.illinois.edu/ads/coord_database.html
[21] Eilers, P.H.C. (2003). A perfect smoother. Analytical Chemistry, 75(14), 3631–3636.
[22] Deb, K. (2000). An efficient constraint handling method for genetic algorithms. Computer Methods in Applied Mechanics and Engineering, 186(2–4), 311–338.
[23] Hansen, N., Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. Evolutionary Computation, 9(2), 159–195.

Orbital Welding of Stainless-Steel Pipes in the Nuclear Industry: History, Technological Development, Applications, and Future Perspectives

Gursel Yildiz [1,2] [0009-0002-8430-7255] and Barış Barlas[3] [0000-0002-5846-2369]
[1] Istanbul Technical University, Turkey
[2] BMC Marine, Turkey
[3] Piri Reis University, Istanbul, Turkey
yildizgur@itu.edu.tr

Abstract. Orbital welding has become an important joining technique for stainless steel piping and tubing in industries where weld quality, consistency, and reliability are essential. Initially developed to meet the demanding requirements of aerospace applications during the 1960s, the technology has improved considerably through advances in welding equipment, electronic control systems, and automation. These developments have also created opportunities for its application in the nuclear industry, where the integrity of welded joints is critical to the safe and reliable operation of nuclear facilities. This paper examines the historical development of orbital welding and its application to stainless steel piping in nuclear power plants. Attention is given to orbital gas tungsten arc welding (GTAW), including its operating principles, metallurgical considerations, and advantages over conventional manual welding. The paper also discusses practical challenges associated with joint preparation, heat input, internal shielding, weld defects, and the corrosion resistance of stainless steel. In addition, it considers the importance of welding procedure criterion, inspection, material traceability, and compliance with applicable nuclear construction codes. Recent developments in digital process monitoring and automated quality documentation are reviewed. The discussion highlights that, although orbital welding can improve consistency and productivity, its successful application in nuclear facilities depends on welding procedure, welding preparation, and compliance with the requirements of the specific piping system.

Keywords: Orbital welding, nuclear industry, stainless steel piping, nuclear quality welding.

1 Introduction

The temperature of the water entering the nuclear reactor pressure vessel is approximately 285°C, while the exit temperature is around 320°C. The pressure inside these tubes is approximately 160 bars. Safety is a critical issue in nuclear power plants, and it becomes even more prominent in the context of nuclear-powered vessels. Welding technology plays a fundamental role in the construction and maintenance of nuclear power plants. Extensive piping networks transport water, steam, gases, chemicals, and other process fluids through systems that may perform essential safety

functions or support normal plant operation. Depending on the system, these pipes must withstand combinations of internal pressure, temperature, vibration, thermal cycling, corrosion, and mechanical loading. AISI 321 grade stainless steel is widely used in nuclear facilities because of its corrosion resistance, ductility, toughness, and suitability for numerous demanding service environments.

The quality of a welded joint depends on several interacting factors, including base material composition, filler metal selection, joint geometry, welding parameters, purging and shielding gas, cleanliness, and operator competence. Manual welding can produce excellent results when performed by qualified welders using appropriate procedures. However, variation in manual torch movement, travel speed, electrical parameter variations and heat application can make consistency more difficult to maintain over large production volumes. Orbital welding addresses this challenge by mechanizing the movement of the welding electrode around a stationary tube or pipe by preprogrammed parameters. The process provides repeatable control over the welding cycle and creates an electronic record of selected process parameters for further quality inspections. Historical studies of automated nuclear piping fabrication demonstrate that mechanized GTAW attracted increasing industrial interest during the 1970s, while aerospace research had already established the potential benefits of automatic tube welding [1].

The objective of this paper is to explain how orbital welding evolved, why it is relevant to nuclear stainless-steel piping, and which technical and quality requirements determine whether it can be used successfully.

2. Historical Development of Orbital Welding

The origins of modern orbital welding are closely associated with aerospace manufacturing. The aircraft and space-systems needed dependable welded connections for hydraulic lines, fuel systems, and other small-diameter fluid circuits. These applications demanded leak-tight joints, controlled penetration, low component weight, and consistent manufacturing quality. Traditional manual welding could satisfy some of these requirements, but the production of numerous similar joints introduced variability. The development of equipment that rotated a welding arc around a fixed tube joint offered a practical alternative. Early systems combined a tungsten electrode, a mechanized welding head, an electrical power source, and basic control equipment. The electrode travelled around the joint while the tube remained stationary. The approach reduced the need for the welder to manipulate a torch continuously around the pipe circumference. NASA’s technical literature describes the emergence of automatic orbital GTAW in the aerospace sector during the early 1960s and its subsequent commercial development. Early equipment was comparatively large and technically demanding, requiring substantial engineering and electronic support [2]. Although the first applications were not primarily associated with nuclear power plant construction, they established the basic technological principles that later supported industrial and nuclear applications.

Orbital welding equipment became increasingly practical for specialized industrial production. More capable welding heads, improved power supplies, and better control of current and electrode movement helped extend its use beyond aerospace manufacturing. Nuclear facilities represented a promising application because instrumentation systems and small-bore stainless-steel tubing often required clean, reliable, repeatable joints. Orbital welding could reduce dependence on manual torch manipulation and improve the consistency of welds when joint preparation and welding conditions were controlled. Research and engineering work also explored mechanized welding for larger nuclear piping systems. Oak Ridge National Laboratory published a report on automated orbital pipe-welding systems in 1973, while a 1979 technical paper documented the increasing role of mechanized and automatic welding in nuclear piping fabrication [3]. These developments are important because they show that orbital welding evolved from a specialized tube-joining method into a technology with broader relevance to nuclear plant fabrication and repair. More compact power supplies and integrated control units made it easier to transport equipment to construction sites and perform welds where moving the component to a workshop was impractical. Electronic controls improved the ability to program welding current, rotational speed, and other process variables. Welding schedules could be developed for specific tube dimensions and joint configurations, helping reduce unnecessary variation between production welds. A 1985 technical paper on automatic orbital GTAW highlighted the benefits of repeatability, controlled heat application, and the suitability of the process for stainless steel and other demanding alloys [2][4]. These advances supported wider use in industrial piping, including applications where quality requirements and the cost of defective joints made repeatability particularly valuable.

Orbital welding systems benefited from more sophisticated electronic controls, improved welding heads, programmable power supplies, and better process documentation. The technology became more accessible to fabrication contractors and maintenance organizations. Equipment could store multiple welding schedules, allowing qualified procedures to be selected for different tube sizes and material combinations. In nuclear-related applications, this period reinforced the distinction between simply automating a weld and establishing a qualified, repeatable manufacturing process. Welding automation could reduce variability, but only when supported by controlled fit-up, suitable materials, correct parameter selection, and systematic inspection.

Last two decades, welding equipment increasingly incorporated digital control systems, electronic data storage, programmable welding sequences, and more comprehensive documentation. Depending on the system, recorded information may include welding current, voltage, rotation speed, programmed segments, and alarms or interruptions. Some systems also support data export and integration with manufacturing records. This development is particularly valuable for nuclear fabrication, where each weld may need to be linked to material certificates, a welding procedure, personnel credentials, inspection results, and component documentation. During the same period, nuclear construction and refurbishment projects continued to emphasize welding quality, material traceability, inspection, and management of degradation mechanisms affecting

stainless steel piping. It is important to distinguish a digitally controlled welding system from a fully autonomous one. A programmable power source can improve repeatability, but it does not independently verify material identity, joint cleanliness, root geometry, or the acceptability of the completed weld.

Modern orbital welding development increasingly focuses on data integration, enhanced monitoring, improved equipment portability, and the ability to document manufacturing activities consistently. Current systems may offer more advanced parameter monitoring, computerized procedure selection, automated reporting, and connections to broader production-management systems. The availability and functionality of these features vary between equipment manufacturers and models. In parallel, nuclear industry developments (including new reactor designs, small modular reactors, plant life extension, and refurbishment projects) continue to create demand for reliable welding and fabrication methods. The central direction of progress is therefore not simply faster welding. It is the integration of repeatable welding operations with traceable materials, controlled procedures, documented inspection, and lifecycle quality records.

3. Basic Principles of Orbital Welding

Orbital GTAW is a mechanized arc-welding process in which a non-consumable tungsten electrode travels around the circumference of a tube or pipe joint. An electric arc forms between the electrode and the workpiece, generating the heat required to melt the base material. An inert shielding gas, commonly argon or a qualified argon-based mixture, protects the molten weld pool and hot metal from atmospheric contamination. For stainless steel joints where, internal root oxidation must be controlled, the pipe interior may also be purged with a suitable shielding gas. Orbital welding machine examples is given in Fig.1. The principal components of an orbital welding system generally include welding power source with programmable controls, orbital welding head and electrode assembly, clamping or alignment arrangement, control unit for the welding schedule, shielding-gas delivery equipment, suitable purge-gas arrangement and monitoring, data recording, and inspection equipment as specified by the project.



Fig. 1. Orbital welding machines

The welding head travels around the stationary joint. In automatic systems, the programmed sequence controls the motion and relevant welding parameters. Orbital GTAW can be performed with or without filler metal, depending on the joint design, material thickness, and qualified welding procedure. Autogenous welding uses only the base material to form the joint. It is commonly associated with thin-wall tubing and suitable joint configurations. Filler-metal welding introduces additional material into the weld pool. It may be required for thicker walls, larger-diameter piping, specified joint geometries, or applications where the qualified procedure calls for filler metal. For nuclear service, the choice between these methods must be established through engineering assessment and procedure. Autogenous welding should not be assumed suitable for every stainless-steel pipe joint.

Based on industry experience, two broad categories of orbital welding heads are commonly used. Enclosed welding heads surround the joint in a controlled chamber. They are particularly suitable for small-diameter, thin-wall tubing and applications where close control of the local shielding environment is beneficial. Open welding heads allow greater access to the joint and are used for a wider range of tube and pipe sizes, including applications that require filler-wire feeding. The appropriate equipment depends on pipe diameter, wall thickness, joint design, site access, production volume, and the qualified welding procedure. An orbital welding schedule typically divides the circumference into programmed segments. Welding current and travel speed may vary according to position around the pipe, because heat dissipation and weld-pool behavior change as the electrode moves. Important variables include welding current and current waveform, electrode rotational or travel speed, arc voltage and arc length, where controlled, pulse frequency and pulse-current settings, shielding-gas flow and its purity, internal purge conditions, electrode geometry and its condition, joint alignment, root gap, and surface preparation, and filler-wire feed rate. All these variables interact with each other. For example, excessive heat input can increase distortion and alter the microstructure, while insufficient energy may cause incomplete fusion or penetration.

The correct settings must be established and validated for the actual material, dimensions, joint design, and applicable requirements.

4. Stainless Steel Materials Used in Nuclear Piping

Stainless steel is used for many nuclear piping applications because it combines corrosion resistance with useful mechanical properties. However, stainless steel is not a single material category with identical welding behavior. Common austenitic grades include 304L and 316L, together with other grades specified for particular nuclear systems, like 321 grades. The low-carbon grades 304L and 316L can help reduce the risk of sensitization associated with chromium-carbide precipitation during exposure to certain thermal cycles. Their use does not eliminate all corrosion risks, and material selection must consider the actual operating environment. Grade 316L contains molybdenum, which generally improves resistance to certain forms of localized corrosion compared with 304L. The selection of one grade over another depends on the system requirements rather than on welding convenience alone. Some nuclear applications specify specially controlled materials, such as nuclear-grade 321 and 316NG, where enhanced resistance to particular degradation mechanisms is required. The material specification, chemical composition, product form, heat treatment, and procurement requirements must be verified against the applicable design documentation. A standard commercial stainless-steel grade should not automatically be treated as interchangeable with a nuclear-grade material.

The welding process locally melts and re-solidifies the material, producing a weld metal microstructure different from that of the original pipe. The surrounding heat-affected zone may also undergo microstructural changes. Depending on material chemistry and thermal history, potential concerns include sensitization and susceptibility to intergranular corrosion, hot cracking under unfavorable metallurgical or restraint conditions, excessive oxidation of the internal root surface., unfavorable ferrite content in the weld metal, distortion and residual stresses, and localized corrosion or stress corrosion cracking under susceptible service conditions. Austenitic stainless-steel piping in nuclear facilities has a history of service-related degradation, including intergranular stress corrosion cracking in certain boiling-water reactor piping systems. Material improvements, fabrication controls, inspection, and corrective actions have been used to manage these risks [6]. Orbital welding can help control the thermal cycle and improve repeatability, but it cannot independently eliminate material-related degradation mechanisms.

5. Quality Assurance and Regulatory Requirements

Quality assurance is one of the most significant distinctions between general industrial orbital welding and its application to nuclear components. Nuclear fabrication is subject to stringent regulatory, code, and quality requirements intended to ensure component

integrity, safety, traceability, and long-term reliability. Where the project adopts the ASME Boiler and Pressure Vessel Code, relevant provisions may include ASME Section III: Rules for Construction of Nuclear Facility Components, including applicable piping requirements; ASME Section IX, covering welding, brazing, and fusing; and ASME Section V, which establishes nondestructive examination (NDE) methods and requirements where invoked by the governing code [5]. The applicable edition, division, subsection, addenda, and project-specific requirements must be identified for each project. ASME identifies Section III as the principal code framework for the construction of nuclear facility components and maintains a nuclear certification program for organizations manufacturing components under its requirements [7]. Other jurisdictions may apply different national regulations, nuclear codes, standards, or combinations of requirements. Compliance with one code should not be assumed to satisfy all applicable regulatory obligations.

The welding procedure shall be qualified in accordance with the governing code and project specification. This typically requires a documented Welding Procedure Specification (WPS) supported by the records. This procedure demonstrates that the selected combination of base material, joint design, welding process, filler metal (where applicable), and essential process variables can consistently produce acceptable welds [5]. Personnel involved in setup, programming, operation, inspection, and quality control must possess the qualifications and competencies required by the governing code and project specifications [5]. Responsibilities should be clearly assigned for approving welding schedules, verifying equipment readiness, authorizing production welding, and addressing process deviations to ensure effective quality control throughout fabrication.

Material identification and traceability are fundamental requirements in nuclear fabrication. Weld records should be linked, as applicable, to base-material identification, material certificates, heat or lot numbers, filler-metal records, weld locations, approved procedures, responsible personnel, and inspection results. Such documentation provides objective evidence that the installed component complies with approved design, procurement, and quality requirements [5]. Inspection and nondestructive examination requirements depend on the governing code, component classification, service conditions, joint configuration, and project specification. Verification activities may include visual examination of external welds, inspection of internal root surfaces where required, dimensional verification, radiographic examination, ultrasonic examination using qualified techniques, liquid penetrant examination, and other tests specified in the approved inspection plan. Not all examination methods are required for every weld, and certain techniques may have boundaries for small-diameter tubing or specific joint geometries. Acceptance criteria shall be established by the governing code and approved project documentation. Automatic welding parameter records may support the quality dossier and process verification but do not replace required inspections, examinations, or acceptance decisions [5].

6. Technical Tasks and Applications

Orbital welding is sensitive to inconsistent joint preparation. Variations in squareness, root gap, alignment, wall thickness, and surface condition can affect penetration and weld-pool stability. For small-bore tubing, even minor geometric inconsistencies can be significant relative to the wall thickness. An orbital cutting machine photo is given in Fig.2.



Fig. 2. Orbital cutting machine photo

Effective production therefore requires controlled cutting, end preparation, cleaning, alignment, clamping, and inspection before welding. Applications of orbital beveling and fit up photos are shown in Fig.3 and Fig.4 respectively.

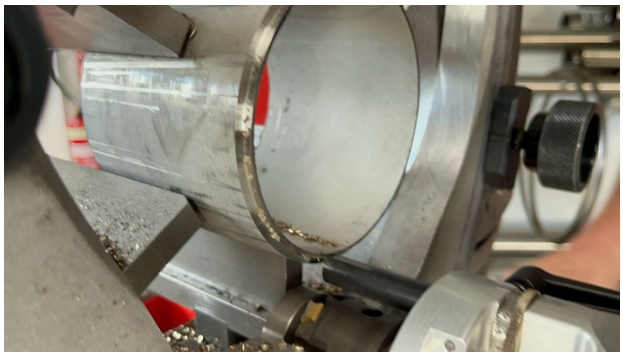


Fig. 3. Orbital beveling

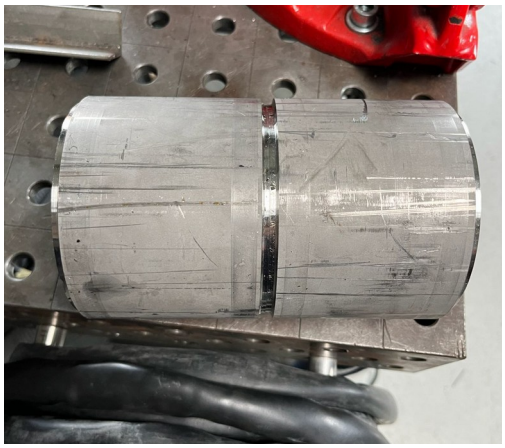


Fig. 4. Fit Up

Stainless steel roots exposed to oxygen at elevated temperatures can develop oxide scale and heat tint. Depending on severity and service conditions, this can impair corrosion resistance and create an undesirable internal surface. An appropriate internal purge may be required to reduce oxygen exposure during welding and cooling. Purge-gas purity, flow, pressure, and termination criteria should be specified in the qualified procedure or approved work instructions. A purge monitor is shown in Fig.5. Excessive purge pressure can also affect the weld pool, particularly in thin-wall tubing. Purging must therefore be controlled rather than simply maximized.



Fig. 5. Purge monitor

Heat input affects penetration, bead geometry, distortion, and microstructure. In orbital welding, the changing orientation of the weld around the circumference can influence molten-metal behavior. Programmed changes in current and travel speed may be needed to maintain consistent results. Nevertheless, a welding schedule developed for one pipe size or material combination should not be transferred to a different application without the necessary technical review. Fig.6 gives the application of open head orbital welding machine.

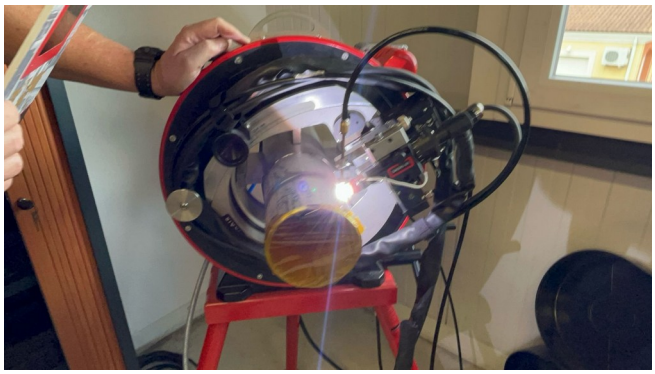


Fig. 6. Open head orbital welding machine

Field installations may involve restricted space, nearby equipment, awkward pipe orientations, and limited access for the welding head. The selected system must fit around the joint and permit the required clamping, shielding, purging, inspection, and maintenance activities. Orbital welding can reduce the physical demands of torch manipulation. Orbital welding of two stainless-steel pipes is given in Fig.7. Reliable results depend on the condition of the power source, welding head, drive mechanism, electrode assembly, gas delivery system, and associated monitoring equipment. Maintenance, calibration or verification, functional checks, and software or program control should be managed under the applicable quality system. Unauthorized parameter changes or use of an unapproved programcan undermine process repeatability.

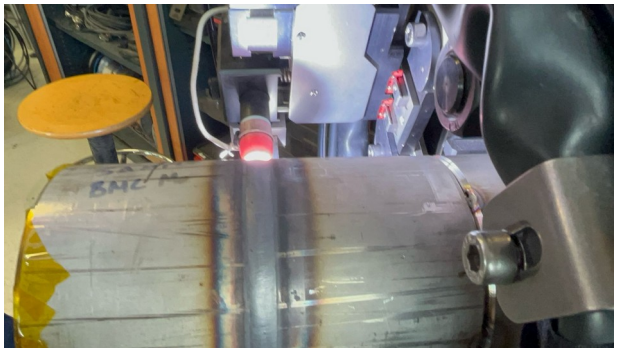


Fig. 7. Orbital welding

7. Comparison of Orbital and Manual GTAW

Orbital GTAW and manual GTAW share the same fundamental arc-welding principle, but differ in the way the welding motion and parameters are controlled. Neither method is inherently acceptable or unacceptable for a nuclear application. Selection should be based on the component requirements, approved procedure, geometry, productivity, access, inspection strategy, and demonstrated quality. Orbital GTAW differs from

manual GTAW primarily in the level of automation, repeatability, and process control. In orbital GTAW, the electrode movement is mechanized around the joint, whereas in manual GTAW the movement is controlled directly by the welder. As a result, orbital welding generally offers higher repeatability when the procedure is properly qualified and the equipment setup is consistent, while weld quality in manual GTAW is more dependent on operator skill and working conditions. Orbital systems also permit welding schedules and process parameters to be programmed, stored, and reused, whereas manual GTAW requires the welder to continuously control welding parameters and torch movement during operation.

Both processes require proper joint preparation; however, orbital GTAW typically demands more tightly controlled joint geometry and alignment to ensure consistent machine operation. In terms of accessibility, orbital welding requires sufficient space for the weld head and clamping system, while manual GTAW can be more adaptable in confined or difficult-to-access locations. For repetitive production welding, orbital GTAW is often advantageous because it can reduce operator-to-operator variation and provide more consistent results. Manual GTAW can also be effective for repetitive work but may exhibit greater variability between welders. Documentation and quality control approaches also differ. Orbital welding systems may support automatic recording of selected welding parameters, contributing to process traceability and documentation. By contrast, manual GTAW generally relies more heavily on manual recordkeeping unless supplementary monitoring systems are employed. Orbital GTAW typically involves higher initial equipment costs, whereas manual GTAW generally requires lower capital investment. Nevertheless, both processes require appropriate procedure qualification, process control, inspection, and quality assurance measures to achieve acceptable weld quality and compliance with applicable codes and standards.

Table 1. Welding motion and parameters comparison of Orbital and Manual GTAW

Criterion	Orbital GTAW	Manual GTAW
Electrode movement	Mechanized around the joint	Controlled by the welder
Repeatability	High when the process is qualified and setup is consistent	Depends more directly on individual technique and conditions
Process programming	Welding schedules can be stored and reused	Parameters and movement are controlled during the operation
Joint preparation	Requires controlled geometry and alignment	Also requires good preparation, with some ability to adapt during welding
Access	Requires suitable space for the welding head and clamping	Can be more adaptable in restricted locations
Production of repetitive joints	Often advantageous	Can be effective but may involve greater operator-to-operator variation

Documentation	May support automatic recording of selected parameters	Typically relies more on manual records, unless additional monitoring is used
Initial equipment cost	Generally higher	Generally lower
Quality assurance	Requires qualification, process control, and inspection	Requires qualification, process control, and inspection

The main differences between orbital GTAW and manual GTAW relate to electrode movement, repeatability, process control, joint preparation requirements, accessibility, suitability for repetitive production, documentation capabilities, and equipment costs, while both processes require appropriate qualification, inspection, and quality assurance measures (Table 1).

8. Discussions and Conclusions

Orbital welding can provide several operational advantages in suitable applications. First, it can improve consistency across repeated joints by reducing variation in electrode travel and programmed welding parameters. Second, it can reduce the amount of manual torch manipulation required, which may improve productivity for suitable geometries. Third, programmable control and electronic records can help standardize production and support documentation. The economic case must nevertheless consider more than the welding cycle. Equipment purchase, setup, fixture design, training, maintenance, inspection, repair rates, and record management all contribute to the total cost. The greatest benefit is often achieved when the same qualified configuration is used repeatedly, and the production environment supports consistent joint preparation. For one-off welds, unusual geometries, or extremely restricted access, manual welding may remain more practical. A project-specific comparison should consider the complete fabrication and inspection process rather than comparing arc-on time alone.

A practical limitation observed in nuclear applications are expected to focus on enhanced process monitoring, digital traceability, advanced data analysis, robotic integration, and support for emerging reactor technologies. Continued improvements in process monitoring systems may enable more detailed collection and interpretation of welding data, helping identify abnormal conditions, interruptions, or deviations from approved welding schedules. The effectiveness of such systems, however, depends on validated measurement techniques, reliable data acquisition, appropriate alarm thresholds, data integrity, and clearly defined corrective actions when deviations occur. Digital weld traceability is also expected to become increasingly important. Integrated digital records can link weld identification to approved procedures, equipment configurations, material certifications, welding personnel, process parameters, inspection reports, and examination results. This approach can improve the completeness, accessibility, and management of manufacturing records while supporting quality assurance and regulatory compliance. For nuclear projects, any

digital traceability system must maintain record integrity, controlled access, and compliance with project-specific documentation requirements. Advances in data analytics may allow manufacturers to better understand relationships between welding parameters and weld quality, supporting optimization of procedures, production planning, and process consistency. Artificial intelligence (AI) technologies may further contribute through anomaly detection, trend analysis, and decision-support functions. Nevertheless, AI-generated assessments are not equivalent to code acceptance criteria, and any system used for quality-critical decisions must be properly validated for its intended application and incorporated within an approved quality management framework.

The integration of robotics and mechanized handling systems is also expected to expand, particularly in prefabricated pipe-module fabrication and repetitive manufacturing environments. Such technologies can improve component positioning, reduce manual handling, and enhance production consistency. In nuclear construction, however, these potential benefits must be evaluated alongside practical considerations such as accessibility, component tolerances, system classification, inspection requirements, approved procedures, and personnel credentials. Advanced reactor concepts, small modular reactors (SMRs), and the refurbishment of existing nuclear facilities may create additional opportunities for automated welding and digitally managed fabrication processes. The greatest long-term benefit is likely to arise from combining proven orbital welding technologies with disciplined engineering practices, rigorous materials control, effective inspection programs, and comprehensive traceable documentation. In this context, orbital welding should be viewed not as a standalone solution, but as one element of an integrated system designed to achieve reliable, high-quality nuclear construction and lifecycle management.

Orbital welding has developed from a specialized aerospace joining method into a mature mechanized welding technology used across several industries, including nuclear power generation. Its historical progress has been driven by improvements in welding heads, electronic power sources, programmable controls, portability, and process documentation. For stainless steel piping, orbital GTAW offers precise control of electrode movement and welding parameters. When the equipment, joint preparation, shielding, and qualified procedure are appropriately selected, it can improve repeatability and support efficient fabrication of numerous similar joints. Its use in the nuclear industry requires particular attention to material selection, metallurgical behavior, internal oxidation, heat input, residual stresses, joint preparation, personnel competency, traceability, and nondestructive examination [8]. The governing code and project requirements must determine whether orbital welding is suitable for a specific component. In projects using ASME requirements, the relevant provisions of Section III, Section IX, and Section V should be considered alongside the applicable quality assurance program and other invoked requirements. Looking forward, digital monitoring, improved data traceability, and advanced process analytics offer opportunities to strengthen the consistency and documentation of nuclear welding operations. These developments will deliver the greatest value when they support,

rather than replace, established engineering controls and qualified inspection practices [9].

The authors consider, orbital welding should be understood not merely as an automated alternative to manual welding, but as a manufacturing process whose value depends on the integration of equipment capability, qualified procedures, materials engineering, inspection, and rigorous quality assurance.

Acknowledgements

This study forms part of the first author's PhD research at Istanbul Technical University.

References

1. Hood, D. W. (1979). *Automated Welding of Nuclear Piping Systems.* UNT Digital Library, U.S. government technical literature. https://digital.library.unt.edu/ark:/67531/metadc1098054/
2. Henon, B. K. (1985). *Automatic Orbital GTAW Welding: Highest Quality Welds for Tomorrow's High-Performance Systems.* NASA Technical Reports Server. https://ntrs.nasa.gov/citations/19860001762
3. Holz, P. P. (1973). *ORNL Automated Orbital Pipe Welding Systems.* Oak Ridge National Laboratory. https://digital.library.unt.edu/ark:/67531/metadc1031467/
4. Clautice, W. E. (1978). *Automatic Welding of Stainless-Steel Tubing.* NASA Technical Reports Server, NASA-TM-79759. https://ntrs.nasa.gov/citations/19780022260
5. American Society of Mechanical Engineers (ASME). *Nuclear Component Certification.*https://www.asme.org/certification-accreditation/nuclear-component-certification
6. International Atomic Energy Agency (IAEA). *Nuclear Energy Series: Plant ageing, degradation mechanisms, and related nuclear piping experience.* https://www-pub.iaea.org/MTCD/publications/PDF/Pub1337_web.pdf
7. American Society of Mechanical Engineers (ASME). *Required Code Books for ASME Certification.*https://www.asme.org/certification-accreditation/asme-certification-process/required-code-books
8. Engelhard, G., Habip, L. M., Pellkofer, D., Schmidt, J., and Weber, J. (2000). *Optimization of residual welding stresses in austenitic steel piping: proof testing and numerical simulation of welding and post welding processes.* Nuclear Engineering and Design, 198(1–2), 141–151. https://doi.org/10.1016/S0029-5493(99)00286-1
9. International Atomic Energy Agency (IAEA). *Dissimilar Metal Welds: Inspection and related issues in nuclear power plants.* IAEA-TECDOC-1852. https://www-pub.iaea.org/MTCD/Publications/PDF/TE-1852web.pdf

Securing Acoustic Communication for Underwater Drones Based on the Primitive Vector Cipher

Gülçin Çivi Bilir [1] [0000-0002-8375-980X] Volkan Ateş [2] [0009-0005-6056-4013] and Barış Barlas [3] [0000-0002-5846-2369]
[1] Istanbul Technical University, Istanbul, Turkey
[2] EPC Energy Systems, Istanbul, Turkey
[3] Piri Reis University, Istanbul, Turkey
civi@itu.edu.tr

Abstract. Reliable and secure data transmission in underwater environments is a critical challenge for underwater drones. Standard wireless communication methods like radio and optical waves are practically impossible or highly restricted due to extreme signal attenuation. While acoustic waves offer a viable alternative, the physical layer is severely constrained by three distinct characteristics of the underwater acoustic channel: severe multipath with delay spreads of tens to hundreds of milliseconds, strong time-varying Doppler caused by platform motion at speeds that are a meaningful fraction of the speed of sound in water, and very limited usable bandwidth because acoustic attenuation grows quickly with frequency. Due to the high latency, low bandwidth, and severe attenuation characteristic of underwater acoustic channels, traditional cryptographic frameworks often introduce prohibitive computational overhead and communication delays. To address these drawbacks, this paper introduces a practical and lightweight encryption framework based on a Primitive Vector Cipher (PVC) approach specifically optimized for secure underwater acoustic communication between underwater drones. The proposed PVC method leverages low-overhead vector transformations to achieve robust data confidentiality without compromising the tight operational constraints of underwater drone telemetry. The results demonstrate that the lightweight encryption framework based on a PVC approach optimized to secure underwater acoustic communication achieves a superior balance between encryption strength and computational efficiency, significantly reducing processing latency and power consumption compared to conventional standards. Consequently, this approach provides a scalable and resilient security solution for real-time underwater drone deployments, environmental monitoring, and marine defense operations.

Keywords: Underwater Drones, Acoustic Communication, Primitive Vector Cipher, Vector Diffie-Hellman.

Introduction

Reliable and secure communication is one of the biggest challenges for underwater drones. Traditional wireless technologies such as radio frequency and optical communication do not work effectively underwater because signals experience severe

attenuation. As a result, acoustic communication becomes the most practical option. However, acoustic channels introduce several challenges. They suffer from multipath propagation, where signals arrive through different paths and create delays. They are also affected by Doppler shifts caused by vehicle movement, and they offer only limited bandwidth. Because of these constraints, any security mechanism used underwater must be computationally efficient and introduce minimal communication overhead.

Underwater drones, also known as Autonomous Underwater Vehicles (AUVs) and Remotely Operated Vehicles (ROVs), increasingly rely on acoustic communication systems to exchange navigation, control, and mission data. Unlike terrestrial wireless networks, underwater environments do not support conventional radio-frequency communications effectively, making acoustic channels the primary means of long-range communication. However, acoustic communication systems face significant security challenges, including eavesdropping, spoofing, jamming, message manipulation, and denial-of-service attacks. This paper examines the security requirements of underwater acoustic networks, analyzes potential cyber threats, and discusses methods for protecting the confidentiality, integrity, and availability of communications between underwater drones and control stations.

The use of underwater drones has expanded in scientific research, environmental monitoring, offshore energy operations, military applications, and maritime surveillance. Reliable communication is essential for transmitting sensor data, mission updates, and control commands. Acoustic communication remains the dominant technology due to the severe attenuation of radio waves in seawater. Despite its operational advantages, acoustic communication introduces unique vulnerabilities. The open nature of acoustic propagation allows adversaries to intercept signals, while underwater environmental conditions complicate the implementation of conventional cybersecurity techniques. This paper focuses on the challenge of securing underwater drone communications and introduces a lightweight cryptographic solution designed specifically for underwater acoustic channels.

2. Underwater Communication Systems

Underwater communication systems play a critical role in a wide range of marine applications, including environmental monitoring, offshore exploration, military operations, disaster prevention, and scientific research. Unlike terrestrial wireless networks that primarily rely on radio frequency waves, underwater communication predominantly utilizes acoustic waves because radio signals experience severe attenuation in water and cannot propagate effectively over long distances. Acoustic communication systems transmit information through sound waves and are therefore the most practical and widely adopted solution for underwater environments.

A typical underwater acoustic communication system consists of several interconnected components. Acoustic modems are used to encode, transmit, and receive data between underwater devices. Underwater sensors collect environmental and operational information such as temperature, pressure, salinity, and the presence of specific marine phenomena. Autonomous Underwater Vehicles (AUVs) are often deployed to perform exploration, inspection, and data-gathering tasks while communicating with other underwater nodes. Surface gateways serve as intermediate communication points, linking underwater acoustic networks with terrestrial or satellite communication systems. Finally, shore-based control stations provide centralized monitoring, data processing, network management, and mission control functions. The four common underwater communication scenarios are depicted in Fig1.

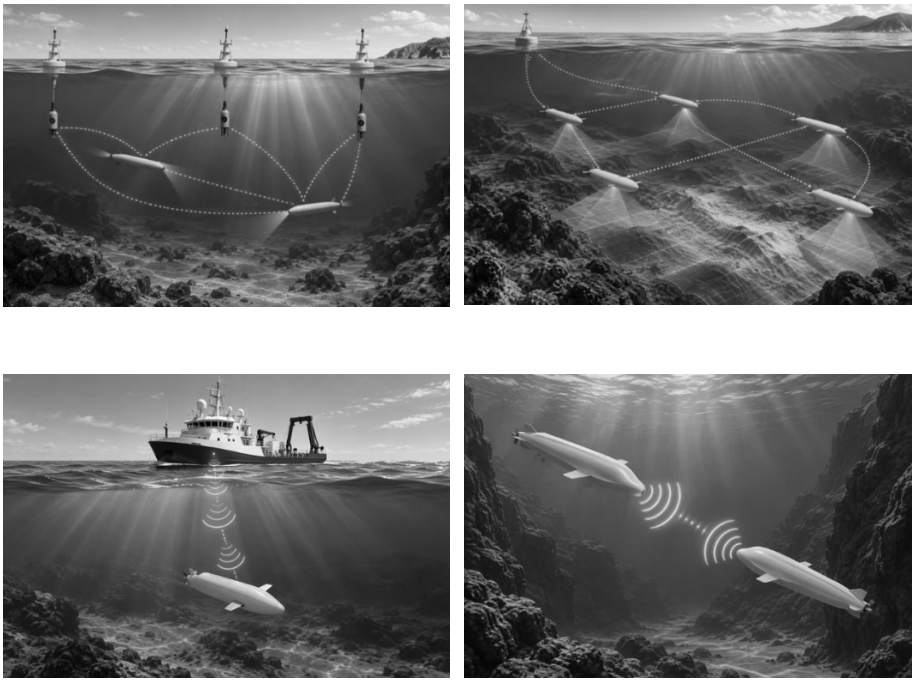


Fig.1: Representative underwater acoustic communication scenarios: (a) acoustic communication network, (b) fleet operation, (c) surface-to-underwater communication, and (d) platform-to-platform communication.

Despite their advantages, underwater acoustic communication systems face several challenges that significantly distinguish them from conventional terrestrial networks. One of the primary limitations is the availability of bandwidth, which is considerably

lower than that of radio-frequency communication systems. Additionally, the relatively slow propagation speed of sound in water introduces high communication latency, making real-time data exchange more difficult. Multipath propagation is another major challenge, as acoustic signals are often reflected by the sea surface, seabed, and other obstacles, causing multiple delayed copies of the same signal to arrive at the receiver. Furthermore, the movement of underwater objects, water currents, and waves can introduce Doppler effects, resulting in frequency shifts and synchronization issues. Signal quality may also vary significantly due to environmental factors such as water depth, temperature gradients, salinity levels, ambient noise, and weather conditions.

These unique characteristics create a highly dynamic and error-prone communication environment. Consequently, ensuring reliable and secure data transmission in underwater acoustic networks is considerably more challenging than in terrestrial communication systems. The combination of limited bandwidth, high latency, fluctuating channel conditions, and increased vulnerability to interception or disruption necessitates the development of efficient communication protocols and robust cryptographic mechanisms. Therefore, advanced encryption, authentication, and key management techniques are essential for protecting sensitive information and maintaining the integrity, confidentiality, and availability of underwater communications.

3. Modulation and Coding Techniques for Underwater Acoustic Communication

The design of underwater acoustic communication (UAC) systems is strongly influenced by the unique characteristics of the underwater acoustic channel. Unlike terrestrial wireless channels, underwater environments are characterized by severe multipath propagation, significant Doppler effects, limited available bandwidth, and high propagation delays (Ha et al, 2017). Multipath propagation occurs due to reflections from the sea surface, seabed, and various underwater objects, resulting in delay spreads that can range from tens to hundreds of milliseconds. Furthermore, the motion of underwater platforms, water currents, and surface waves introduces time-varying Doppler shifts that can distort transmitted signals and complicate synchronization. Another major limitation is the restricted bandwidth available for acoustic communication, typically ranging from a few kilohertz to several tens of kilohertz, due to the rapid increase in acoustic attenuation at higher frequencies.

High latency is an additional challenge in underwater communication systems. Since sound travels in water at approximately 1500 m/s, signal propagation is significantly slower than electromagnetic wave propagation in terrestrial networks. For example, a communication link spanning 1 km can experience a round-trip delay of approximately 1.3 seconds. As a result, retransmission-based error recovery mechanisms become inefficient, making robust forward error correction techniques essential for maintaining reliable communication performance.

One of the most widely adopted modulation schemes for underwater acoustic communication is Orthogonal Frequency Division Multiplexing (OFDM). OFDM divides the available bandwidth into numerous orthogonal narrowband subcarriers, allowing data to be transmitted in parallel across multiple frequency channels. By reducing the symbol rate on each subcarrier, OFDM effectively mitigates the adverse effects of multipath propagation and intersymbol interference (ISI). Intersymbol Interference (ISI) is a phenomenon that occurs when one transmitted symbol spreads into the time interval of the next symbol, causing the symbols to overlap at the receiver (Stojanovic & Preisig, 2009). As a result, the receiver may have difficulty determining where one symbol ends and the next begins, leading to decoding errors. In underwater acoustic communication, ISI is primarily caused by multipath propagation. When an acoustic signal is transmitted, it does not travel along a single direct path. Instead, portions of the signal are reflected from the sea surface, seabed, rocks, underwater structures, and other objects. Consequently, multiple copies of the same signal arrive at the receiver at different times (Stojanovic, 2008).

In addition, a cyclic prefix is inserted between consecutive OFDM symbols to absorb delayed signal echoes and preserve orthogonality among subcarriers. To address the time-varying characteristics of underwater channels, pilot and null subcarriers are commonly incorporated for channel estimation, synchronization, and Doppler compensation. These features make OFDM particularly well suited for underwater environments and explain its widespread use in modern acoustic communication systems.

Quadrature Phase Shift Keying (QPSK) is another important technique frequently employed in underwater acoustic communications. QPSK encodes information using four distinct carrier phase states, enabling the transmission of two bits per symbol. Compared with higher-order modulation schemes, such as 16-QAM or 64-QAM, QPSK offers superior robustness against noise, fading, and channel impairments because its constellation points are separated by larger phase distances. Although higher-order modulations can achieve greater spectral efficiency, the challenging characteristics of underwater acoustic channels often make QPSK a more reliable choice. Consequently, QPSK provides a practical balance between data rate, power efficiency, and communication reliability in dynamic underwater environments.

To further enhance communication reliability, Low-Density Parity-Check (LDPC) coding is commonly employed as a forward error correction (FEC) technique. LDPC codes introduce structured redundancy into transmitted data through a sparse parity-check matrix, enabling the receiver to detect and correct transmission errors without requiring retransmission. Decoding is performed iteratively using belief-propagation or message-passing algorithms that exchange reliability information between variable nodes and check nodes within a graphical representation known as a Tanner graph. LDPC codes are capable of achieving performance close to the Shannon capacity limit while maintaining efficient implementation in parallel hardware architectures.

Although LDPC is a channel coding technique rather than a modulation method, it constitutes a fundamental component of modern underwater communication systems. The code effectively corrects residual errors caused by noise, fading, interference, and channel distortions that remain after demodulation. Furthermore, the use of interleaving across time intervals and OFDM subcarriers helps distribute burst errors over an entire codeword, thereby improving the decoder's ability to recover the original information. Because the long propagation delays of underwater acoustic channels make Automatic Repeat reQuest (ARQ) retransmissions costly and inefficient, LDPC-based forward error correction significantly improves throughput, reliability, and overall communication efficiency. Iterative message passing between variable nodes and parity-check nodes, originally proposed by Gallager (1962), enables the correction of residual errors and enhances the robustness of the communication system under challenging underwater channel conditions.

Together, OFDM modulation, QPSK signaling, and LDPC forward error correction form a powerful physical-layer framework for underwater acoustic communications. OFDM combats multipath propagation, QPSK provides a robust modulation strategy under varying channel conditions, and LDPC coding ensures reliable error correction without excessive retransmissions. The combination of these techniques enables efficient and secure data transmission in underwater environments characterized by limited bandwidth, high latency, severe multipath effects, and dynamic channel variations.

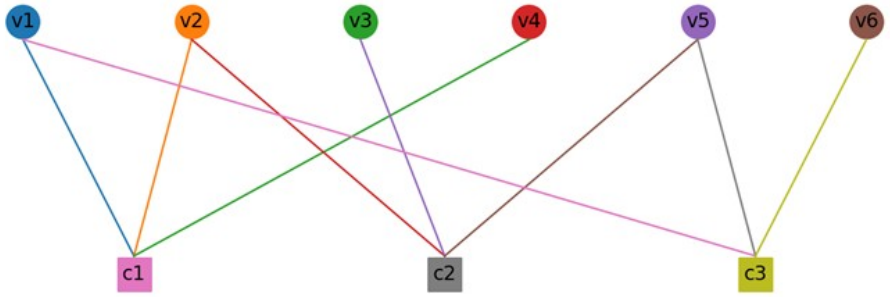


Fig. 2: Iterative belief propagation messages pass between variable and check nodes

The upper diagram in Fig.3 shows the transmitter processing chain. The process starts with the information bits. These bits are first protected using LDPC encoding, which adds error-correction capability. Next, interleaving is applied to spread burst errors across the data stream. The encoded data is then modulated using QPSK, converting bits into symbols. Pilot and null subcarriers are added to support synchronization and channel estimation. An IFFT operation generates the OFDM signal, and a cyclic prefix is inserted to reduce the effects of multipath propagation. Finally, a preamble is added for synchronization, and the signal is transmitted into the water through the acoustic transducer. The lower diagram in Fig.3 shows the receiver processing chain. The

receiver first captures the acoustic signal using a hydrophone. It then performs synchronization and Doppler estimation because motion of the underwater platform can distort the received signal. After Doppler compensation and cyclic-prefix removal, the receiver performs FFT processing to recover the OFDM subcarriers.

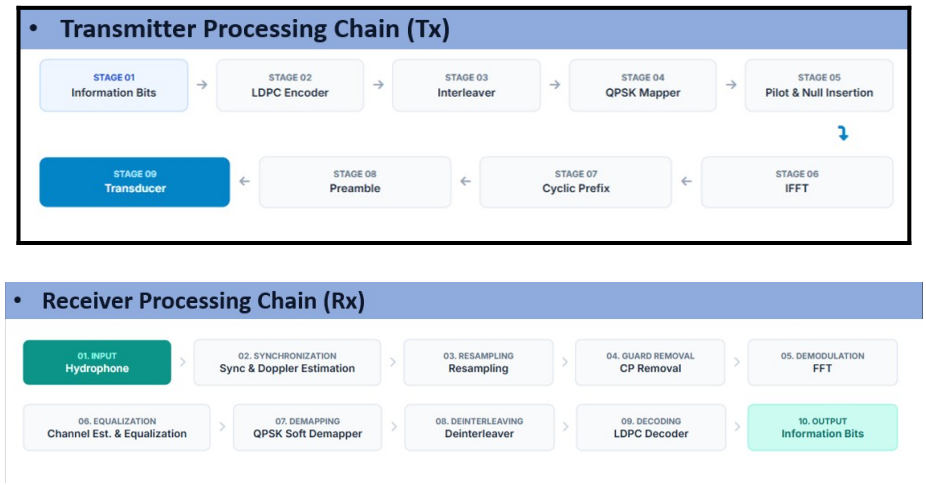


Fig. 3: Architecture of the underwater acoustic transmitter and receiver processing chains.

Channel estimation and equalization are then applied to compensate for underwater channel effects. The receiver converts the QPSK symbols back into bit probabilities, performs deinterleaving, and finally uses LDPC decoding to recover the original information bits. Together, these transmitter and receiver chains provide reliable communication despite the challenging underwater environment. On the transmitter side, information bits pass through LDPC coding, interleaving, QPSK modulation, pilot insertion, OFDM processing, and acoustic transmission. The signal then propagates through the underwater acoustic channel, which introduces several impairments such as multipath propagation, ambient noise, frequency-selective fading, and time-varying Doppler effects. The receiver must therefore perform several corrective operations before recovering the data. These include synchronization, Doppler estimation and resampling, FFT processing, channel estimation, and equalization. Only after these channel impairments have been compensated can the receiver perform QPSK soft demapping and LDPC decoding to reconstruct the original information bits.

This architecture represents a practical and widely used solution for underwater acoustic communications because it combines the robustness of OFDM, the efficiency of QPSK, and the strong error-correction capability of LDPC coding. The PVC security mechanism proposed in this work is designed to operate on top of this communication framework without introducing significant computational or communication overhead. Now that we have seen how reliable underwater communication is established at the physical and data transmission levels, the next question becomes how to secure this

communication efficiently. This leads us to the Primitive Vector Cipher, or PVC, which is the main contribution of this work.

4. Primitive Vector Cipher

Although acoustic communication is the most practical solution underwater, the channel limitations create challenges for conventional cryptographic systems. Traditional encryption algorithms may introduce significant computational overhead, communication overhead, and delays. To address this issue, the paper proposes a lightweight encryption framework based on the Primitive Vector Cipher (PVC) defined by Çivi Bilir, (2026), where the goal is to provide strong security while respecting the strict resource limitations of underwater drones.

The proposed PVC framework consists of three main phases (Fig. 4). First, authenticated key establishment is performed using an STS-based Diffie-Hellman exchange to generate a shared primitive vector (Diffie et al, 1992). Second, during encryption, plaintext is transformed into matrix form, protected with a pseudorandom mask, encrypted through matrix operations, and modified using column-wise offsets. Finally, during decryption, these operations are reversed to reconstruct the original message. The main design principle is combining matrix-based transformations with key-derived pseudorandom masking.

Phase	Main operations
Authenticated key establishment	STS-(Station to Station) based Diffie–Hellman exchange produces a shared primitive vector
Encryption	Map plaintext to a matrix, apply a pseudorandom mask, encrypt selected 3 × 3 submatrices, and add per-column offsets
Decryption	Remove offsets, reconstruct the encrypted submatrices, reverse the matrix transform, and remove the mask

Fig. 4: Phases and main operations of proposed PVC framework.

The proposed encryption scheme consists of seven integrated stages designed to ensure secure key establishment, robust authentication, and confidentiality of transmitted data. In the first stage, a Primitive Vector and Vector Computational Diffie-Hellman (V-CDH) mechanism is employed to establish a shared secret vector between communicating parties through a secure Diffie-Hellman key exchange process. The second stage incorporates Station-to-Station (STS) authentication, which strengthens the protocol by protecting the key exchange against man-in-the-middle attacks and ensuring the authenticity of the communicating entities.

Following successful authentication, the third stage focuses on key generation. In this phase, the secret matrices V and U, along with the required pseudorandom keys, are

derived using the HKDF (HMAC-based Extract-and-Expand Key Derivation Function) (Dworkin, 2001; Krawczyk and Eronen, 2010). This process ensures that cryptographic keys are generated securely and possess high entropy. The fourth stage performs plaintext masking, where the original message is first converted into an ASCII matrix representation and then masked using values generated by a Cryptographically Secure Pseudorandom Number Generator (CSPRNG) (Barker and Kelsey, 2015). This masking step obscures the plaintext structure and enhances resistance against cryptanalytic attacks.

In the fifth stage, known as Shifting SubMatrices (SSM), the masked matrix is divided into individual 3×3 submatrices to facilitate subsequent transformation operations. The sixth stage applies a two-layer encryption process. Initially, matrix-based transformations are performed on each submatrix, after which additional security is provided through column-wise keystream offsets. This layered approach significantly increases the complexity of unauthorized decryption attempts. Finally, in the seventh stage, the decryption process is carried out by applying the inverse of each transformation and masking operation in reverse order, thereby reconstructing the original plaintext message accurately and securely. Together, these seven stages create a comprehensive encryption framework that combines key exchange, authentication, pseudorandom masking, matrix manipulation, and layered cryptographic protection to achieve a high level of data security.

These stages include shared secret establishment using Vector Computational Diffie-Hellman, authentication with the STS protocol, key generation through HKDF, plaintext masking with a cryptographically secure pseudorandom generator, extraction of three-by-three submatrices, two-layer encryption, and finally decryption. Together, these stages provide both authentication and confidentiality while maintaining low computational requirements.

4.1 STS Protocol Flow and Security

The Station-to-Station, or STS, protocol combines Diffie-Hellman key exchange with authentication to protect against man-in-the-middle attacks. After exchanging public values, both drones independently compute the same shared secret, which becomes the basis for subsequent encryption operations.

The key exchange between Drone-A and Drone-B proceeds as an authenticated exchange (Fig. 5). We first define the *primitive vector* $\boldsymbol{g} = (g_1, g_2, g_3) \in (F^*)^3$ such that that each component is a distinct primitive root modulo a large prime p. The three-pass authenticated key exchange between Drone-A and Drone-B proceeds as follows:

1. $Drone\ A \rightarrow Drone\ B : g^a, S_A$
2. $Drone\ B \rightarrow Drone\ A : g^b, S_B, MAC_B(g^q, g^b, S_A, S_B)$
3. $Drone\ A \rightarrow Drone\ B : MAC_A(g^q, g^b, S_A, S_B),$

Where after the exchange, Drone-A and Drone-B obtain ephemeral public vectors and compute the shared secret key (*KMAC*) derived from $G = (g_1^{ab}, g_2^{ab}, g_2^{ab}) \in (F^*)^3$. This use of authentication tags over the transcript ensures that any adversary modifying the public exchange values cannot compute the correct MAC, effectively mitigating the MITM threat.

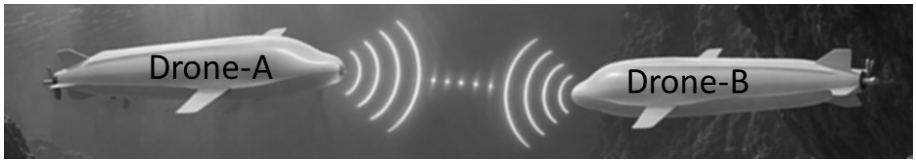


Fig. 5: Platform-to-Platform communication between underwater drones Drone-A and Drone-B

Key Establishment

After key establishment, the plaintext message is embedded into a master matrix. Each character is converted into its ASCII representation, creating a structured matrix form of the message. This matrix serves as the input for the encryption process that follows:

Using the authenticated vector $G = (g_1^{ab}, g_2^{ab}, g_2^{ab}) \in (F^*)^{33}$, we define the secret key matrices:

$$\mathbf{V} = \begin{bmatrix} 0 & k_1 & k_1 \\ k_2 & 0 & k_2 \\ k_3 & k_3 & 0 \end{bmatrix} \text{ and } U = \begin{bmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & k_3 \end{bmatrix},$$

where $k_i = g_i^{ab} \bmod p$.

Master Matrix and ASCII Mapping

The plaintext is embedded into an m × n master matrix N, which is then converted into the corresponding ASCII matrix. A pseudorandom mask matrix $R \in Fpmxn$ is generated deterministically by a CSPRNG seeded from the authenticated shared secret vector **G** via HKDF. Concretely, serialize the vector **G** as

$$\text{encode}(\mathbf{G}) = \text{I2OSP}(k_1, \ell_1) \,||\, \text{I2OSP}(k_2, \ell_2) \,||\, \text{I2OSP}(k_3, \ell_3),$$

$$\text{PRK: } = \text{HKDF_Extract}(\text{salt}, \text{encode}(G)),$$

$$K_{mask} \text{ : } = \text{HKDF_Expand (PRK, "PVC/mask", } L_{mask}),$$

$$L_{mask} = \text{m . n. } [log_2 p / 8].$$

This guarantees that the expanded key material suffices to derive the pseudorandom mask for the entire matrix. The masked matrix is

$$M' = (M + R) \, (mod \; p),$$

which both parties can deterministically reproduce.

Extraction of Shifting Submatrices (SSM)

From the masked matrix M', we extract a family of 3 × 3 submatrices, referred to as *Shifting Submatrices* (SSM). Let the sets of top–left row and column indices be

$$I \subseteq \{1, \ldots, m-2\}, J \subseteq \{1, \ldots, n-2\}.$$

For a fixed ordering π on $I \times J$, we enumerate $(i_k, j_k) = \pi(k), \; k = 1, \ldots, |I||J|$.

Each submatrix is defined by

$$S_{i_k, \; j_k} \begin{bmatrix} p_{ik} & p_{i(k+1)} & p_{i(k+2)} \\ p_{(i+1)k} & p_{(i+1)(k+1)} & p_{(i+1)(k+2)} \\ p_{(i+2)(k+1)} & p_{(i+2)(k+1)} & p_{(i+2)(k+2)} \end{bmatrix},$$

$$i = 1, 2, , \ldots, m-2, k = 1, 2, \ldots, n-2.$$

where $p_{a,b}$ are entries of M'. The total number of transmitted blocks is $B = |I||J|$.

Each selected SSM $S_{ij} \in F_p^{3x3}$ is encrypted as

$$C_{ij} = S_{ij} \, V + \Delta_{ij} U \quad (\, mod \; p \,),$$

where

$$\Delta_{ij} = \begin{cases} I_{3x3} \;, & If \;\; i = j \\ 0_{3x3} \;, & If \;\; i \neq j \end{cases}$$

The blocks $\{(i_k, j_k)\}$ and their columns $c_r \in (F_p^3)$ are enumerated by the index $l = 3(k - 1) + r, r \in \{1, 2, 3\}$

Using a separate key $K_{CTR-cols}$ derived from HKDF, per–column offsets r_l are generated. The final ciphertext columns are

$$\tilde{c_l} \equiv c_l + r_l \; (mod \; p), \;\; l = 1, \ldots, 3|I||J|,$$

which are transmitted to the receiver in Drone-B.

Decryption from Drone-B, given the header

$$hdr = (p, G, \mathbf{g}, \mathbf{g^a}, salt, nonce, m \times n, indexing)$$

and the column stream $\tilde{c}_1$, $\ell = \{1, 2, \cdots 3|I||J|\}$, the receiver computes the shared vector

$$\mathbf{G} = \mathbf{g^{ab}} \in (Z_p)^3 \text{ and derives keys via HKDF :}$$

$$\text{PRK} = \text{HKDF_Extract(salt, encode(}\mathbf{G}\text{))},$$
$$K\text{mask} = \text{HKDF_Expand(PRK, “PVC/M_mask_R/}m\times n\text{”)},$$
$$K\text{cols} = \text{HKDF_Expand(PRK, “PVC/CTR-cols/}m\times n\text{”)}.$$

The submatrix $\boldsymbol{S_{ij}}$ is reconstructed via the inverse transformation:

$$S_{ij} = (C_{ij} - \Delta_{ij} \ U)V^{-1} \ (mod \ p).$$

Finally, the master matrix $\boldsymbol{M'}$ (the masked plaintext matrix) is reconstructed by reassembling the decoded submatrices:

$$\mathbf{M}' = \text{Reassemble } (\{S_{ij}\}).$$

The original plaintext matrix **M** is then recovered by removing the CSPRNG mask **R**:

$$\mathbf{M} = \mathbf{M}' - \mathbf{R} \ (\text{mod } p),$$

8. Conclusions

This study reports linear operation-count scaling and preliminary entropy and keystream results. Establishing the claimed suitability for underwater drones will require a reconciled protocol specification and measurements on the target hardware and acoustic link. The encryption mechanism is a two-layer approach: plaintext masking followed by matrix-based encryption and column-wise keystream offsetting. The computational performance reported the peak memory requirement is only 1.1 KB for a 12 × 23 example. The operation counts also demonstrate linear scaling with the number of selected submatrices.

The proposed PVC method leverages low-overhead vector transformations to achieve robust data confidentiality without compromising the tight operational constraints of underwater drone telemetry. The results demonstrate that the lightweight encryption framework based on a PVC approach optimized to secure underwater acoustic communication achieves a superior balance between encryption strength and computational efficiency, significantly reducing processing latency and power consumption compared to conventional standards.

The approach combines plaintext masking, matrix-based encryption, and keystream offsets while maintaining low memory and computational requirements. The authors consider, PVC aims to achieve a balance between security and computational efficiency, making it a promising approach for underwater drones, environmental monitoring systems, and marine defense applications.

References

[1] E. Barker and J. Kelsey, Recommendation for Random Number Generation Using Deterministic Random Bit Generators, NIST Special Publication 800-90A Revision 1, National Institute of Standards and Technology (NIST), 2015.

[2] G. Çivi Bilir, “Primitive Vector Cipher (PVC): A Hybrid Encryption Scheme Based on the Vector Computational Diffie–Hellman (V-CDH) Problem,” Journal of Discrete Mathematical Sciences and Cryptography, in press, 2026.

[3] W. Diffie, P. C. van Oorschot, and M. J. Wiener, “Authentication and authenticated key exchanges,” Designs, Codes and Cryptography, vol. 2, no. 2, pp. 107–125, 1992.

[4] M. Dworkin, Recommendation for Block Cipher Modes of Operation: Methods and Techniques, NIST Special Publication 800-38A, National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA, 2001.

[5] R. G. Gallager, “Low-density parity-check codes,” IRE Transactions on Information Theory, vol. 8, no. 1, pp. 21–28, 1962. doi:10.1109/TIT.1962.1057683.

[6] D. V. Ha, V. D. Nguyen, and Q. K. Nguyen, “Modeling of Doppler power spectrum for underwater acoustic channels,” Journal of Communications and Networks, vol. 19, no. 3, pp. 270–281, 2017.

[7] H. Krawczyk and P. Eronen, HMAC-based Extract-and-Expand Key Derivation Function (HKDF), RFC 5869, Internet Engineering Task Force (IETF), May 2010. doi:10.17487/RFC5869.

[8] M. Stojanovic, “OFDM for underwater acoustic communications: Adaptive synchronization and sparse channel estimation,” in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), Las Vegas, NV, USA, 2008, pp. 5288–5291.

[9] M. Stojanovic and J. Preisig, “Underwater acoustic communication channels: Propagation models and statistical characterization,” IEEE Communications Magazine, vol. 47, no. 1, pp. 84–89, 2009.

ABSTRACT Track 3 · Emerging Trends and Technologies in Naval Architecture, Marine Engineering and Maritime Sciences

Full-Wave Design and Broadband Characterization of a Center-Fed Planar Elliptical Dipole Antenna for 3.1–10.6 GHz Ultra-Wideband Communication

Zakaria Mesmoudi[1]

[1] *LEPCI Laboratory, Department of Electronics, Faculty of Technology, Sétif 1 University – Ferhat Abbas, Algeria*
zakariamesmoudi19@gmail.com

This paper presents the full-wave design and physical characterization of a center-fed planar elliptical dipole antenna, which is intended for ultra-wideband (UWB) wireless communication within the FCC-allocated frequency range of 3–11 GHz. The antenna features two co-planar elliptical arms, each incorporating an inner elliptical slot, and is constructed on a Rogers RT/duroid 5880LZ substrate (ε_r = 2.2). It is fed between the two arms via a 50 Ω discrete port. Time-domain full-wave simulations conducted in CST Microwave Studio 2021 indicate a reflection coefficient below -10 dB across the frequency range of 3.2 to 10.55 GHz, yielding a fractional bandwidth of 106.9% that effectively encompasses the entire UWB allocation. The corresponding voltage standing-wave ratio (VSWR) remains close to or below 2 throughout the band, achieving a minimum of 1.6 at the 3.5 GHz matching frequency. The input impedance oscillates between 65 and 85 Ω in its real part, while the reactance stays within approximately ±30 Ω above 2.5 GHz, reflecting the non-resonant, gradually tapered current path typical of elliptical UWB dipoles. The peak realized gain, derived from simulated far-field patterns at six frequencies across a decade (1.1–11 GHz), increases from 1.8 dBi to 7.1 dBi as frequency rises. Additionally, the radiation pattern transitions from a simple twin-lobed dipole shape to a multi lobed structure as the aperture becomes electrically large; this trend is explained in terms of the increasing electrical size of the fixed physical aperture and the emergence of higher-order current distributions on the elliptical arms. The design, along with its governing eight parameter geometry, is linked to the invasive-weed-optimization (IWO) methodology outlined in previous studies. Furthermore, the antenna is compared against five recently published (2022–2025) UWB and UWB-MIMO antennas, several of which are based on the same elliptical radiating-element concept, situating the current single-element design within the contemporary state of the art.

Keywords: *Ultra-Wideband Antenna, Elliptical Dipole Antenna, Planar Antenna, CST Microwave Studio, UWB-MIMO.*

ABSTRACT Track 3 · Emerging Trends and Technologies in Naval Architecture, Marine Engineering and Maritime Sciences

Design, Simulation and Physical Performance Analysis of an Ultra-Wideband Cavity-Backed Log-Periodic Folded-Slot Dipole Array Antenna for UHF/HF and EMC Applications

Zakaria Mesmoudi[1]

[1] *LEPCI Laboratory, Department of Electronics, Faculty of Technology, Sétif 1 University – Ferhat Abbas, Algeria*
zakariamesmoudi19@gmail.com

In this paper, an ultra-wideband log-periodic folded-slot dipole array (LPFSA) antenna backed by a three-layer lossy dielectric absorber cavity for UHF/HF terrestrial broadcasting and electromagnetic compatibility (EMC) measurement applications is presented. The eight-element array conforms to classical log-periodic scaling theory, with scale factor $\tau = 0.890$ and spacing factor $\sigma = 0.0756$. These two parameters are shown to reproduce the 20° tooth angle implemented in the full-wave model via the standard log-periodic apex-angle relation, confirming internal design consistency. Simulated in CST Microwave Studio (finite integration technique) about a design frequency $f_0 = 8$ GHz, the antenna occupies a compact volume of 44.54 mm × 38.10 mm × 12.38 mm. Over the simulated 4.64–11.64 GHz window (≈88% fractional bandwidth referred to f_0), the input reflection coefficient remains below -8 dB and the voltage standing wave ratio (VSWR) mostly below 2.5:1, with a deep resonance of about -27.8 dB near 6.3 GHz. The realized gain varies between 3 dBi and 8.9 dBi, and end-fire-like, forward-tilted main beams are obtained at the lower, center, and upper edges of the band. The physical origins of the periodic impedance ripple, the gain notch near f_0, the upper-band mismatch, and the influence of the absorptive cavity on back-radiation suppression are analyzed and discussed. The design is further benchmarked, quantitatively and qualitatively, against four recently reported log-periodic dipole array antennas (2022–2024).

Keywords: *Log-periodic Dipole Array (LPDA), Folded-slot Antenna, Ultra-Wideband (UWB) Antenna, Cavity-backed Antenna, Absorber Loading, CST Microwave Studio.*

TRACK 4

Data-Driven Circular Built Environment

SPECIAL TRACK ORGANIZERS

Burcu Kısmet Conk Özyeğin University, Türkiye



KEYNOTE SPEAKER

Prof. Hexin (Johnson) Zhang

Edinburgh Napier University, United Kingdom

Engineering the Future with Bio-Based Materials (Structural Timber & Bamboo Focus)

Prof. Hexin (Johnson) Zhang is Professor of Structural Engineering and Associate Dean International at the School of Computing, Engineering and the Built Environment, Edinburgh Napier University. He holds a PhD in Civil Engineering from the Hong Kong University of Science and Technology and has extensive academic and professional experience in structural engineering. His research focuses on innovative structural materials, sustainable construction, engineered bamboo and timber, bamboo–timber composites, gridshell structures, laminated glass, prestressed concrete, and complex structural systems.

A major strand of his research explores the development and real-scale application of bio-based structural systems as low-carbon alternatives for the built environment. Prof. Zhang has pioneered research on bamboo–timber composite gridshell structures, leading the development of the world's first real-scale bamboo–timber composite gridshell, followed by Europe's first bamboo–timber composite gridshell at Edinburgh Napier University. His work bridges material and structural experimentation, computational and engineering analysis, full-scale construction, structural monitoring, and industry-oriented implementation, and has been supported through a Royal Academy of Engineering Industrial Fellowship.

Mapping Artificial Intelligence Applications for a Circular Economy in the Built Environment: A Bibliometric Analysis

Kudzai Mbowa[1]* and Clinton Aigbavboa[1]

[1] Sustainable Human Settlement and Construction Research Centre, Faculty of Engineering and the Built Environment, University of Johannesburg, Johannesburg, South Africa
* Corresponding author: Kudzai Mbowa; sammykmbowa@gmail.com

Abstract.

Artificial intelligence (AI) offers opportunities to improve resource decisions in the built environment, but research on these opportunities extends across several technological and disciplinary domains. This study mapped publication trends, international collaboration and keyword connections in a Scopus export concerning AI, circularity and the built environment. A bibliometric review was conducted using 861 records dated 2015–2026. Descriptive indicators were combined with country co-authorship analysis and an author-keyword co-occurrence network. Keyword variants were consolidated, and terms appearing in at least 15 documents were analysed using association-strength normalisation and Louvain community detection. Publication output increased from 34 records in 2022 to 274 in 2025; the incomplete 2026 year contained 336 records. The keyword network comprised 21 terms and 131 links, organised into four communities concerning computational methods and waste circularity, connected digital infrastructure, sustainability and organisational transformation, and computing resource efficiency. These communities exposed an important scope distinction: resource-efficient computing and general digital sustainability occur alongside research directly concerned with construction materials. The study therefore interprets the export as a retrieved research landscape rather than a fully screened set of circular construction interventions. It contributes a transparent account of the connections between AI and circularity and proposes research priorities linking predictive performance with verified reuse, material recovery and lifecycle outcomes.

Keywords: artificial intelligence; circular economy; built environment; bibliometric review; keyword co-occurrence; construction waste

1 Introduction

Circularity in the built environment depends on decisions made throughout the life of buildings and infrastructure. Material selection, maintenance, adaptation, disassembly and recovery influence whether resources retain value or become waste. Digital technologies can support these decisions by making information easier to collect and use. Çetin et al. (2021) connected digital technologies with circular strategies in a built-environment framework, while Çetin et al. (2023) examined the information required to construct material passports for existing buildings. These contributions establish a useful starting point: circular decisions require accessible information about physical assets and their potential future uses.

AI adds capabilities for recognising material characteristics, predicting waste quantities and comparing resource alternatives. Akanbi et al. (2020), for example, developed a deep-learning approach to demolition-waste prediction. However, prediction is only one part of the circular process. A technically accurate model contributes to resource retention only when its outputs inform feasible collection, recovery or reuse decisions. This distinction motivates a review of how AI and circularity are connected in the research landscape.

Existing reviews approach this landscape from different directions. Ametepey et al. (2024) mapped construction digitalisation broadly, whereas Wu et al. (2025) examined intelligent technologies in construction and demolition waste management. Their scopes show why digitalisation, AI and circularity require careful differentiation. A building information model can organise asset records without using AI, and an AI system can improve energy or computing efficiency without addressing construction material loops. A bibliometric review can reveal these conceptual connections, provided that retrieved records are not automatically treated as evidence of successful implementation.

The objective of this study was to map the development and knowledge connections of literature retrieved at the intersection of AI, circularity and the built environment, and to identify priorities for research on circular applications. Three questions guided the analysis: how did publication output develop; which countries and collaboration links were visible; and which keyword communities structured the retrieved landscape? The contribution is an empirical map accompanied by an interpretation of the boundary between material circularity and broader digital resource efficiency.

2 Research methodology

2.1 Review design and data source

The study employed bibliometric performance analysis and science mapping, following the distinction described by Donthu et al. (2021). Tambwe and Aigbavboa (2026) provided a structural precedent for presenting a Scopus-based review through publication trends, country collaboration and keyword communities. The present study applied that sequence to a different research problem and dataset. Analysis was implemented in Python using pandas, NetworkX 3.7 and Matplotlib; the figures were calculated from the supplied metadata rather than generated in VOSviewer.

The source was a Scopus CSV export retrieved on 23 October 2026, containing 861 records dated 2015–2026. The documented search scope combined AI expressions, circularity or resource-efficiency expressions, and built-environment expressions. The export contained titles, abstracts, authors, affiliations, sources, years, keywords, DOIs and citation information. It did not contain the executed Boolean query or a complete search-history record. Consequently, the retrieval stage cannot be independently reconstructed from the CSV alone, and the analysis is explicitly bounded to this supplied export. No unverified initial retrieval or exclusion counts were introduced.

2.2 Record preparation and analytical scope

All 861 records had abstracts. Forty-two lacked DOIs, 45 lacked author keywords and one lacked affiliation information. Checks of case-insensitive exact titles and non-empty DOIs identified no duplicates, so all records were retained for descriptive analysis. Records without author keywords contributed to publication statistics but not keyword networks. The corpus comprised 569 journal articles and 292 conference papers. No full-text eligibility or quality appraisal was conducted. The resulting map describes a retrieval landscape, including adjacent topics, rather than a systematically screened body of intervention evidence.

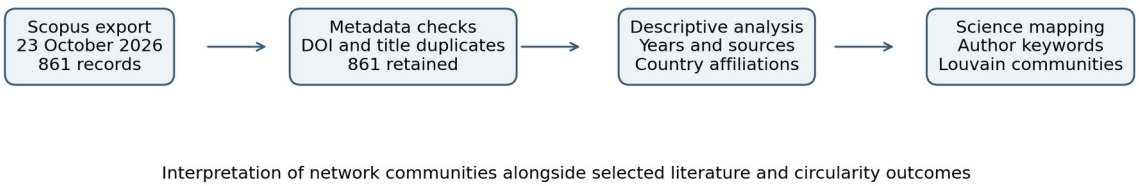


Figure 1. Analysis workflow. Counts refer to the supplied export; no undocumented screening stage is implied.

2.3 Network construction and interpretation

Author keywords were separated at semicolons, converted to lower case and counted once per document. Explicit equivalents were merged, including AI with artificial intelligence; ML with machine learning; BIM and US/British spelling variants with building information modelling; digital twins with digital twin; IoT variants with internet of things; and digitalization with digitalisation. Selected variants of construction and demolition waste were also consolidated. This produced 2,887 distinct normalised terms. The threshold of 15 documents retained 21 terms for a readable map; it was a display and analysis choice, not an eligibility criterion for papers.

For each retained pair, co-occurrence equalled the number of documents containing both terms. Association strength was calculated as $c_{ij}/(o_i o_j)$, where c_{ij} is pair co-occurrence and o_i and o_j are document frequencies. Louvain community detection used these normalised edge weights, resolution 1.0 and random seed 42 (Blondel et al., 2008). Node size represented document frequency; edge width represented raw co-occurrence. A force-directed layout used the normalised weights. Total link strength in Table 1 is the sum of raw co-occurrences with the other 20 retained terms. Community names were assigned interpretively after inspecting member terms; they are not validated research categories.

HTML entities were decoded and non-geographic fragments removed. Country and territory names were extracted from the final comma-separated component of each affiliation and deduplicated within documents. Each represented country received one whole publication count, so totals across countries can exceed 861. A country pair received one collaboration count when both countries occurred in a document. The displayed collaboration map retained countries with at least 40 documents and links supported by at least three joint documents. Such affiliation-based measures indicate research participation, not the location of fieldwork or technology adoption.

3 Research results

3.1 Publication growth and outlet distribution

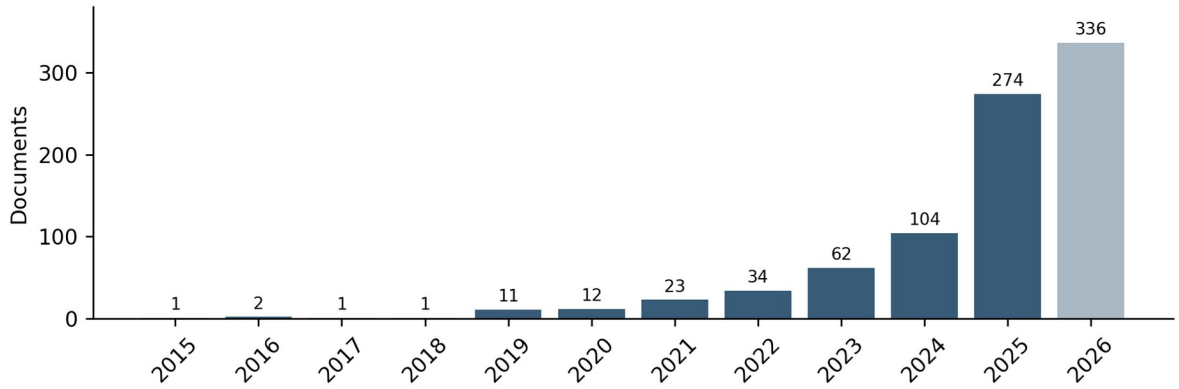


Figure 2. Annual publication counts in the 861-record export. Grey identifies the incomplete 2026 year.

Output remained modest from 2015 to 2021, with annual counts of 1, 2, 1, 1, 11, 12 and 23. It then rose to 34 in 2022, 62 in 2023, 104 in 2024 and 274 in 2025. The 2023–2025 period accounted for 440 records (51.1% of the export). A further 336 records were dated 2026, but this incomplete year should not be used to calculate a directly comparable annual growth rate. The pattern demonstrates increased publication attention within the retrieved literature; it does not establish the extent of practical AI adoption.

Journal articles represented 66.1% of records and conference papers 33.9%. Sustainability was the most frequent source, with 36 documents, followed by Lecture Notes in Civil Engineering (17), Scientific Reports (16) and Lecture Notes in Networks and Systems (16). IEEE Access contributed 10 records, while Buildings and Resources, Conservation and Recycling each contributed nine. The outlet distribution connects sustainability, construction and computing, supporting the need to distinguish direct material applications from adjacent technological work.

3.2 Country participation and international collaboration

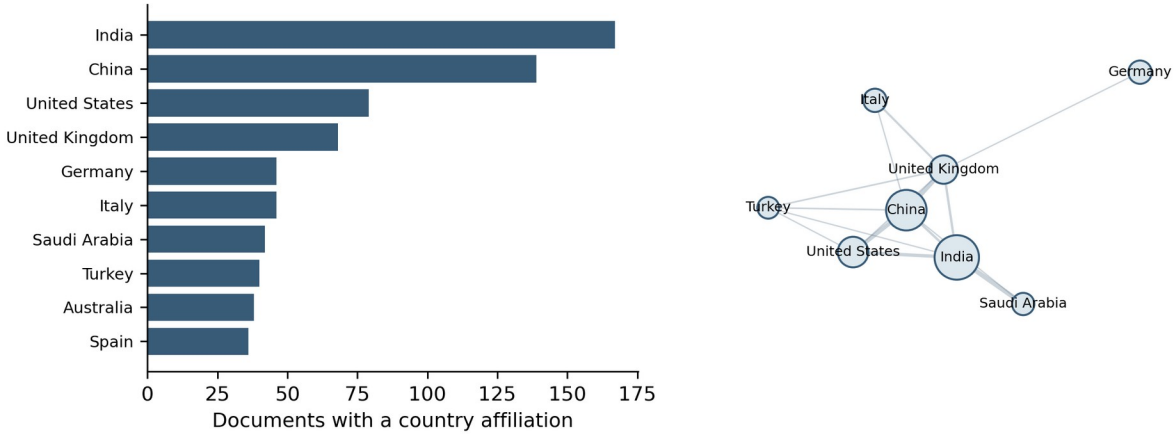


Figure 3. Leading affiliation countries and selected international collaboration links. Whole counting is used; the network shows countries with at least 40 documents and links with at least three joint documents.

The affiliation extraction identified 95 country and territory labels. India appeared in 167 records and China in 139, followed by the United States (79) and United Kingdom (68). Germany and Italy each appeared in 46 records; Saudi Arabia appeared in 42 and Turkey in 40. Australia and Spain contributed 38 and 36 records respectively. These values describe participation in the full retrieval landscape. They do not establish which countries lead specifically in circular construction implementation, because the corpus also contains broader digital and sustainability research.

A total of 277 records (32.2%) contained affiliations from more than one country. The largest extracted collaboration pairs were China–United Kingdom and India–Saudi Arabia, each with 16 shared documents, followed by China–United States (14) and India–United States (11). Collaboration can provide routes for sharing data and expertise, although a co-authored publication does not reveal the depth or continuity of a partnership. No inference about national funding, language barriers or research quality was drawn from these counts alone.

3.3 Keyword prominence and network structure

Author keywords were available in 816 records (94.8%). After normalisation, 21 of the 2,887 distinct terms met the minimum frequency of 15 documents. The retained network contained 131 undirected links and four Louvain communities. Circular economy occurred in 228 documents, artificial intelligence in 191 and machine learning in 167. These frequencies are calculated from author keywords, not from all abstract mentions, which explains why they are lower than the total number of retrieved records.

Keyword	Documents	Link strength
Circular economy	228	258
Artificial intelligence	191	238
Machine learning	167	144
Sustainability	71	110
Deep learning	50	47
Internet of things	48	87
Resource efficiency	29	33
Waste management	28	44

Table 1. Leading normalised author keywords. Link strength is restricted to the 21-term network.

Circular economy and artificial intelligence had the highest raw link strengths, indicating connections with several frequently used terms. These values should be read as network prominence within the selected vocabulary. They do not rank the effectiveness of technologies, and the frequency threshold means that specialised or emerging applications may be absent from the map.

4 Keyword communities and research focus

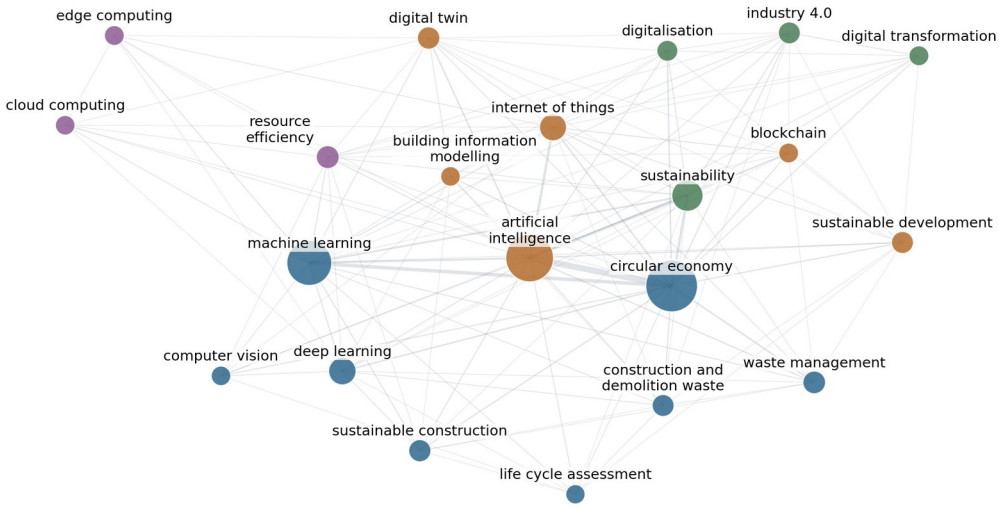


Figure 4. Author-keyword co-occurrence map. Blue: computational methods and waste circularity; orange: connected digital infrastructure; green: sustainability and organisational transformation; purple: computing resource efficiency. Colours reflect the calculated community partition.

4.1 Computational methods and waste circularity

The first community joined circular economy, machine learning, deep learning, waste management, sustainable construction, construction and demolition waste, computer vision and life cycle assessment. It provides the clearest direct connection between computational methods and construction material problems. Akanbi et al. (2020) illustrated the predictive direction through demolition-waste modelling. Wu et al. (2025) subsequently examined intelligent technologies across waste-management activities. Together, these examples help interpret the community as a research space linking the recognition and prediction of waste streams with planning and environmental assessment. Membership itself does not demonstrate that every paper combined all these techniques or measured a circular benefit.

4.2 Connected digital infrastructure

The second community contained artificial intelligence, internet of things, digital twin, sustainable development, blockchain and building information modelling. These terms point towards the infrastructure through which information is captured, organised and exchanged. Their association is relevant to circularity because reuse decisions depend on the identity, condition and history of materials. Nevertheless, a digital twin or blockchain record should not automatically be classified as an AI application. Çetin et al. (2023) showed why the availability of material information is a substantive issue in its own right. The community therefore indicates connections between enabling systems, rather than evidence that any particular technological combination guarantees reuse.

4.3 Sustainability and organisational transformation

The third community comprised sustainability, Industry 4.0, digitalisation and digital transformation. Its vocabulary concerns the wider organisational setting in which digital technologies are introduced. It complements the technical emphasis of the first two communities by locating AI within changes to workflows, skills and management. Ametepey et al. (2024) similarly mapped construction digitalisation as a broad field. Circularity research can build on that work by specifying which organisational changes enable recovery and continued material use, rather than treating digital uptake as a circular outcome.

4.4 Computing resource efficiency

The fourth community grouped resource efficiency, edge computing and cloud computing. This is an analytically important result because the term resource efficiency can refer to computing capacity and digital infrastructure as well as physical construction resources. Its appearance exposes the breadth of the retrieval vocabulary. The community cannot be assumed to represent material circularity without further record-level assessment. Reporting it preserves the distinction between what the network actually contains and what the review initially sought to investigate. A focused follow-up review should explicitly separate computing efficiency, building operational efficiency and material circularity during eligibility assessment.

5 Discussion

5.1 What the map adds to existing reviews

The results connect two observations: publication attention has grown substantially, and the retrieved knowledge base remains conceptually heterogeneous. The first community anchors the map in waste and construction material questions, while the other communities extend towards data infrastructure, organisational transformation and computing efficiency. This pattern helps explain why a broad AI–circularity search can retrieve a larger literature than the set of studies directly evaluating circular construction practices. The map adds value by making those boundaries visible rather than removing them through undocumented exclusions.

Compared with the broad digitalisation review of Ametepey et al. (2024), the present analysis gives circular economy an explicit organising role. Compared with Wu et al. (2025), whose review concentrated on construction and demolition waste technologies, it includes the wider digital and sustainability connections visible in the supplied export. Çetin et al. (2021) established a conceptual relationship between digital tools and circular strategies; the current network shows how related vocabulary co-occurs in a later retrieval landscape. These are complementary contributions. The bibliometric map identifies connections and directions for closer investigation, while application studies are needed to establish what technologies accomplish. Related scholarship also examines AI-enabled systemic circularity in construction, the relationship between AI and energy/material circularity, and AI applications across architecture, engineering and construction (Adopting artificial intelligence, 2023; Awuzie et al., 2024; Applications of artificial intelligence, 2024).

5.2 Linking technical performance with circular outcomes

The most useful interpretation of the computational community is that AI can support decisions about material flows. Predicting waste volumes may improve recovery planning, and recognising materials may assist sorting. However, a forecast or classification becomes circularly useful only when it changes a feasible decision. A sorting model may achieve high accuracy while recovered material remains unusable because of contamination, certification requirements or insufficient demand. Future evaluation should therefore report the sequence from model output to operational action and then to retained material value. This application-oriented interpretation is consistent with reviews of AI for construction circularity and intelligent construction-and-demolition waste management (Adopting artificial intelligence, 2023; Wu et al., 2025).

This distinction also applies to digital information systems. A material passport can describe a component without ensuring that it will be recovered, stored or sold. Çetin et al. (2023) examined the data requirements and availability underlying such passports. Langley et al. (2023) addressed the coordination of digital product passports in a smart circular economy. These studies support examining information and organisational arrangements together. For the built environment, records about composition and condition must connect with practical decisions concerning disassembly, responsibility and subsequent use.

5.3 Interpreting research gaps responsibly

Neither keyword counts nor collaboration links establish how many studies used field trials, tested models outside their training settings or quantified avoided virgin material demand. The present analysis therefore does not claim a measured shortage of such evidence across all 861 records. Instead, it identifies questions that the mapped communities make especially important. These include whether predictive models transfer across projects, whether interoperable records improve actual reuse, and whether computing demands offset part of the environmental benefit. Answering them requires targeted full-text synthesis and comparative empirical studies.

6 Implications and future research

For researchers, the first priority is to define the circular process being evaluated. Waste prevention, component reuse, recycling and operational efficiency require different indicators and baselines. Studies of machine-learning prediction should link forecast error with recovery planning and realised material outcomes. Studies of material matching should measure the proportion of components actually reused, the effort required to verify suitability and the change in virgin material demand. Environmental assessment should then consider lifecycle burdens, including data acquisition and computing, rather than reporting technical accuracy as a substitute for circular performance.

A second priority concerns transferable data and evaluation. Models trained on a narrow set of buildings or waste images may perform differently with other materials, construction practices and recording systems. Cross-project validation, clear material classifications and transparent uncertainty reporting would improve comparability. Where data cannot be openly shared, researchers can still report collection procedures, variable definitions, sample composition and validation protocols. These practices would help determine whether differences between models reflect algorithm choice or differences in available information.

For practitioners, pilots should begin with a defined resource decision and an observable baseline. A demolition contractor could compare conventional recovery planning with an AI-supported process, recording recovery yield, labour input, cost and rejected material. An asset owner could test whether improved component records lead to more reuse during refurbishment. Such designs make the practical contribution reviewable. They also allow simpler information-management improvements to be compared with more demanding AI systems, which is relevant when implementation resources are limited.

For policymakers and industry bodies, interoperability and reliable secondary-material information are priorities for investigation. Shared classifications, material certification and clear responsibilities for updating records may determine whether digital recommendations can be acted upon. In Southern African settings, these priorities should be examined against local construction practices, digital capacity and recovery markets. The present country map does not establish regional implementation barriers or performance; it provides a global context for designing those local studies.

7 Limitations

The findings are bounded by a single Scopus export, the retrieval vocabulary and the metadata available on 23 October 2026. The exact executed search expression was not supplied with the CSV, limiting retrieval reproducibility. Full-text eligibility screening was not undertaken; accordingly, broad digital sustainability and computing topics remain in the corpus. Country extraction reflects affiliation strings rather than verified study locations. Keyword communities depend on the thesaurus, frequency threshold, normalisation and clustering settings, and the chosen partition was not subjected to a stability analysis. Missing author keywords affect network coverage, while the incomplete 2026 year constrains annual comparisons. These conditions support an exploratory mapping contribution rather than claims of exhaustive coverage or intervention effectiveness. Bibliometric software and mapping choices can also affect network representation and interpretation (Aria and Cuccurullo, 2017; van Eck and Waltman, 2010; van Eck et al., 2010).

8 Conclusion

This bibliometric review mapped 861 Scopus records through publication trends, country collaboration and author-keyword connections. The network identified four communities connecting computational methods and waste circularity, digital infrastructure, organisational sustainability and computing resource efficiency. Publication growth demonstrates increasing research attention, while the community structure shows why AI-enabled circular construction should be distinguished from wider digital sustainability. The central implication is that future work should connect information and prediction with observable decisions and verified material outcomes. The map offers a basis for a more tightly screened review and for empirical studies that assess whether AI improves recovery, reuse and lifecycle resource performance under practical operating conditions.

References

Aria, M. and Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics, 11(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

van Eck, N.J. and Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics, 84, 523–538. https://doi.org/10.1007/s11192-009-0146-3

van Eck, N.J., Waltman, L., Dekker, R. and van den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. Journal of the American Society for Information Science and Technology, 61, 2405–2416. https://doi.org/10.1002/asi.21421

Awuzie, B., Ngowi, A. and Aghimien, D. (2024). Towards built environment decarbonisation: A review of the role of artificial intelligence in improving energy and materials' circularity performance. Energy and Buildings. https://doi.org/10.1016/j.enbuild.2024.114491

Adopting artificial intelligence for enhancing the implementation of systemic circularity in the construction industry: A critical review. (2023). Sustainable Production and Consumption, 35, 509–524. https://doi.org/10.1016/j.spc.2022.12.002

Applications of artificial intelligence in the AEC industry: A review and future outlook. (2024). https://doi.org/10.1080/13467581.2024.2343800

Akanbi, L.A., Oyedele, A.O., Oyedele, L.O. and Salami, R.O. (2020). Deep learning model for demolition waste prediction in a circular economy. Journal of Cleaner Production, 274, 122843. https://doi.org/10.1016/j.jclepro.2020.122843

Ametepey, S.O., Aigbavboa, C., Addy, H. and Thwala, W.D. (2024). A bibliometric review of the trends of construction digitalization research in the past ten years. Buildings, 14(9), 2729. https://doi.org/10.3390/buildings14092729

Blondel, V.D., Guillaume, J.-L., Lambiotte, R. and Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Çetin, S., De Wolf, C. and Bocken, N. (2021). Circular digital built environment: An emerging framework. Sustainability, 13(11), 6348. https://doi.org/10.3390/su13116348

Çetin, S., Raghu, D., Honic, M., Straub, A. and Gruis, V. (2023). Data requirements and availabilities for material passports: A digitally enabled framework for improving the circularity of existing buildings. Sustainable Production and Consumption, 40, 422–437. https://doi.org/10.1016/j.spc.2023.07.011

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N. and Lim, W.M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Langley, D.J., Rosco, E., Angelopoulos, M., Kamminga, O. and Hooijer, C. (2023). Orchestrating a smart circular economy: Guiding principles for digital product passports. Journal of Business Research, 169, 114259. https://doi.org/10.1016/j.jbusres.2023.114259

Tambwe, O. and Aigbavboa, C. (2026). Mapping research trends on risk in construction data management in the 5th Industrial Revolution: A bibliometric review. Proceedings of the 15th SACQSP International Research Conference 2025, paper SACQSP2025-118. Cape Town: University of Cape Town. ISBN 978-1-83492-219-5. Conference held 11–13 November 2025; proceedings printed February 2026.

Wu, Z., Pei, T., Bao, Z., Ng, S.T., Lu, G. and Chen, K. (2025). Utilizing intelligent technologies in construction and demolition waste management: From a systematic review to an implementation framework. Frontiers of Engineering Management, 12, 1–23. https://doi.org/10.1007/s42524-024-0144-4

ABSTRACT Track 4 · Data-Driven Circular Built Environment

Bamboo–Timber Composite Gridshell: Design, Construction and Full-Scale Experimental Investigation

Yanghao Pei and Hexin Zhang

School of Engineering and the Built Environment, Edinburgh Napier University, Edinburgh, UK

Actively bent bamboo–timber composite gridshells offer a lightweight structural form that can be constructed from initially straight members. Assessing their structural performance requires an understanding of their deformation behaviour at full scale. This study presents the design, construction and experimental investigation of a full-scale bamboo–timber composite gridshell measuring 12.4 m × 12.4 m. A parametric design-to-construction workflow was developed to guide prototype construction. Laminated bamboo and softwood timber were combined in a composite grid, with active bending used to generate the curved geometry without prefabricated curved members. The prototype was tested using a total applied mass of 3,000 kg distributed over a local area. Analysis of the load–displacement response revealed creep and residual deformation and provided information on the in-plane stiffness of the gridshell. Three-dimensional point-cloud models were generated to visualise the global deformation response. The construction and testing demonstrate the feasibility of the composite gridshell concept and provide experimental evidence of its full-scale structural behaviour. These findings inform further development of analysis and design methods for actively bent bamboo–timber composite gridshells.

Keywords: *bamboo–timber composite, active bending, parametric design, full-scale load testing, load–displacement response, point-cloud modelling.*

ABSTRACT Track 4 · Data-Driven Circular Built Environment

Bauhaus Earth – Towards a Systemic Transformation of the Built Environment

Patricia Jeglitsch

Bauhaus Earth, Berlin, Germany

The construction sector contributes substantially to global carbon emissions, resource consumption and waste generation. Addressing these impacts requires a regenerative approach that reduces environmantal damage and activly supports the restoration of ecosystems.

Bauhaus Earth is an international, interdisciplinary non-profit think-and-make tank that works at the interface between research and practical application, covering systems research and data analytics, material innovation, demonstration, education and policy advice. Its Research Workshop translates research into built demonstrators at different scales. Material properties, environmental impacts and technical requirements are considered together with regional resource availability and local production capacities. Material investigations, design, full-scale prototyping and in-use performance monitoring form an iterative process, accompanied by life-cycle assessment, circularity analysis and evaluation of regionality. Pilot projects allow the resulting systems to be tested under real building conditions.

The Holz-Lehm-Hybrid pilot project in Berlin-Weißensee applied this methodology to five transferable building components using regional timber, unstabilised earth, reused concrete elements and bio-based fibres. Developed within existing regional manufacturing infrastructure, the components offer scalable solutions that comply with applicable regulations. Together, they achieved a 77% reduction CO_2e emissions compared with functionally equivalent conventional components across the assessed life-cycle stages.

ABSTRACT Track 4 · Data-Driven Circular Built Environment

Circularity Across Scales and Disciplines. A New Education Programme for the Built Environment in the Netherlands

Bianca Andaloro (presenting), Poorya Eghtesadi, Cecilia Chiappini, Nurhan Abujidi

Smart Urban Redesign (SURD), Zuyd University of Applied Sciences, The Netherlands

Circularity in the built environment is still mostly framed as a question of resource efficiency, centred on the optimisation of resource flows (Pekdemir et al., 2025). This focus tends to treat the environmental, social and economic pillars of sustainability separately and overlooks how circular principles take shape in specific places and at specific scales. Building on the concept of Circular (Urban) Development (Williams, 2023), this contribution positions circularity as an integrated Triad of technical, spatial and social dimensions (Bosschaert, 2023), addressed through an inter-scalar and trans-disciplinary approach. Within this framing, circularity is understood as design, shaping spaces and their uses; as flow, connecting resources, people and ecosystems; and as process, involving the actors who govern and inhabit urban transformation. Translating this vision into practice remains difficult, as professionals with distinct roles and expertise struggle to coordinate circular ambitions.

In response, a new education programme is being developed at Zuyd University of Applied Sciences, building on the cross-border nature of the province of Limburg and the opportunities it offers for exchange across national, institutional and disciplinary boundaries. The programme aims to build circular implementation capacity among practitioners through practice-oriented support, equipping them to shape urban transformations that are technically circular, spatially integrated and socially embedded.

Keywords: *Circular design; inter-scalarity; trans-disciplinarity; circular education*

REFERENCES

Bosschaert, T. (2023). Circularity is not sustainability: How well-intentioned concepts distract us from our true goals, and how SiD can help navigate that challenge. In H. Lehmann, C. Hinske, V. de Margerie, & A. S. Nikolova (Eds.), The impossibilities of the circular economy: Separating aspirations from reality (pp. 72–80). Routledge. https://doi.org/10.4324/9781003244196-9

Pekdemir, S., Guaralda, M., & Limb, M. (2025). The missing link: Circularity in urban design – A systematic review of circular paradigms for resilient and self-sustaining cities. Cities, 162, Article 106007. https://doi.org/10.1016/j.cities.2025.106007

Williams, J. (2023). Circular cities: Planning for circular development in European cities. European Planning Studies, 31(1), 14–35. https://doi.org/10.1080/09654313.2022.2060707

ABSTRACT Track 4 · Data-Driven Circular Built Environment

Adaptive Reuse for Urban Food Waste Collection: Hub Location and Routing Optimization

Esma Selen Aksoy

Istanbul Technical University, Türkiye

Urban food-waste management presents a growing challenge for densely developed cities, where dispersed waste generation and collection-related transport demands complicate the transition towards circular urban systems. Given the biodegradable nature of food waste and associated risks of odour and hygiene-related problems, timely and regular collection is particularly important (European Commission, 2020). This study investigates how existing shopping malls and office buildings can be strategically selected and repurposed as food-waste collection hubs to improve the spatial efficiency of urban collection networks, using Besiktas, Istanbul, as a case study. The research integrates geospatial analysis, activity-based waste-demand estimation, and network-based location and routing optimization. A dataset of 381 food-service establishments was compiled, and daily food-waste generation was estimated using business characteristics and Google review activity as an indirect indicator of operational intensity. Forty existing buildings were identified as potential collection hubs, and road-network distances were calculated to evaluate their accessibility within the collection system. Single- and two-hub configurations were systematically assessed using a Capacitated Vehicle Routing Problem (CVRP) to identify hub locations and collection routes that minimize total travel distance under vehicle capacity constraints. These findings demonstrate how the integrated evaluation of existing building locations and collection routes can support the strategic selection of decentralized food-waste collection hubs. By reducing collection-related travel distances, the proposed approach may contribute to more efficient and reliable collection operations, supporting the timely transfer of perishable food waste while positioning adaptive reuse as a component of circular urban infrastructure.

ICASE 2026

General Session



The General Session brings together contributions from all areas of applied sciences and engineering. It was held in two parallel parts, in virtual rooms 5 and 6.

PART 1 · SESSION CHAIRS

Prof. Dr. Ümit Işıkdağ
Prof. Dr. Gebrail Bekdaş
Dr. Ayla Ocak

PART 2 · SESSION CHAIRS

Prof. Dr. Sinan Melih Nigdeli
Yaren Aydın

The Effect of Isolator Mobility in Base Isolation Systems Optimized by Accounting for Soil-Structure Interaction

Ayla Ocak[1[0000-0003-2236-7834]], Sinan Melih Nigdeli[2[0000-0002-9517-7313]] and Gebrail Bekdaş[2[0000-0002-7327-9810]]

[1] Istanbul Gelisim University, 34310 Istanbul, Turkey
ayocak@gelisim.edu.tr
[2] Istanbul University - Cerrahpaşa, 34320 Istanbul, Turkey
melihnig@iuc.edu.tr, bekdas@iuc.edu.tr

Abstract. Seismic base isolators are highly effective control systems that rapidly reduce the acceleration of seismic vibrations transmitted to a structure. Key factors to consider in the seismic isolation of a structure include the region's soil conditions and seismicity. This study focuses on investigating changes in a structure's dynamic response based on different isolator mobility levels for two soil conditions, soft and dense, in the seismic isolation of a structure located in a region exposed to distant fault earthquakes. To conclude, the Jaya algorithm, a metaheuristic algorithm, was used to optimize the system period and damping ratio characteristics of an isolator-equipped structure subjected to a FEMA far-fault earthquake for both soft and dense soil conditions, considering different isolator mobility levels. In the optimization, FEMA earthquake records known to be critical for both soft and dense soil conditions were considered for the given structural characteristics. Critical earthquake analyses were performed using the obtained optimal isolator system parameters, and the effect of isolator mobility on the damping of earthquake vibrations was investigated for both soft and dense soil types. Earthquake analyses have shown that taking the SSI effect into account in an isolator-based system can increase the system's required isolator mobility for a given damping ratio and that, when isolator mobility is increased relatively, the system's required period can be reduced.

Keywords: Seismic Base Isolation, Structural Control, Soil-Structure Interaction, Optimization.

1 Introduction

Base isolators are effective control devices that reduce a structure's dynamic responses by damping strong vibrations, such as those caused by earthquakes. They play an active role, particularly in minimizing the transmission of vibrations such as seismic vibrations, which threaten structural safety and negatively impact human life. It is well established that the fundamental operating principle of seismic isolators is to interrupt and reduce the transmission of vibrations that could be transmitted from the ground to the structure through oscillations generated at the isolator level. Their high performance

in reducing acceleration sets seismic isolators apart from similar control devices. Isolators typically achieve this effect by extending the structure's natural period. There are two key parameters for an isolator to be added to a structure. These are the control system period, which allows the structure to oscillate in order to prevent damage caused by vibrations to the superstructure, and the system damping ratio, which affects the isolator's ductility when seismic energy is transmitted. The optimal tuning of these parameters is crucial for reducing the structure's dynamic responses, such as maximum displacement and total acceleration. In the optimization of base isolators, the rubber type, elastomer layer dimensions, allowable isolator displacement, period, damping ratio, isolator stiffness, the normalized yield strength of the lead core, the yield deformation of the lead-rubber bed, the yield base shear ratio, the coefficient of friction, the radius of curvature, the number of isolators, and their arrangement have been optimized in various studies [1–7]. The type of isolator selected and its characteristic damping properties play a significant role in determining which parameters to optimize. The vast majority of studies on control devices represent research that does not account for structure-soil interaction. However, it has been observed that soil typology affects the isolator's design parameters, shear displacement, and the acceleration accumulated in the structure [8]. Furthermore, a structure on soft ground is affected by structural motion and soil stiffness during an earthquake and must accommodate significant displacement demands [8–10]. The influence of the soil on a control system to be installed in the structure has been supported by various studies [11–15]. All of the findings obtained highlight the need to incorporate the structure-soil interaction into control system designs and to integrate control systems into the structure, taking into account not only the region's seismicity but also the soil conditions.

The seismic base isolator evaluated in this study is linearized, and its two fundamental parameters, period and damping ratio, have been optimized. Meta-heuristic algorithms have gained prominence in the literature as a method frequently preferred in recent years for the optimization process. Meta-heuristic algorithms are known to be mathematical models with an infrastructure that mimics natural phenomena and instinctive animal behavior, inspired by nature. In studies on the design of optimal control systems for base isolators, meta-heuristic algorithms such as the harmony search algorithm, a derivative aimed at finding the best harmony [2, 16] the genetic algorithm inspired by natural selection and evolutionary processes [3, 7, 17], particle swarm optimization inspired by the swarm behavior of birds, fish, and similar organisms [4, 18], and the grasshopper optimization algorithm [5, 19], inspired by the swarm interactions and movements of grasshoppers, have been preferred. Within the scope of this study, the Jaya algorithm [20], a good optimization tool due to its speed in reaching the optimal solution and its simple structure, was used. Jaya is an algorithm that aims to find the best solutions by moving away from poor solutions; it features a simple and easy-to-understand structure, contains a single solution-generation equation, and does not involve any algorithmic steps.

Within the scope of this study, a 50% damped seismic isolator was placed on the base of a structure for two types of soil (soft and dense) and optimized using the Jaya algorithm for isolator displacement limits of 30 and 40 cm. The structural model used by Ocak et al. [12] was adopted for the optimization. In a study by Ocak et al. [21],

which optimized a structural model without accounting for soil-structure interaction (SSI), it was found that for a system equipped with isolators having 50% damping, the best control performance was achieved with an isolator displacement limit of 30 cm, followed by the system with a 40 cm limit. Consequently, the present study aims to investigate whether the inclusion of SSI effects makes a significant difference when the isolator displacement capacity is set to 30 cm or 40 cm. The NORTHR/MUL279 and DUZCE/BOL090 earthquakes identified by Ocak et al. [12] as critical for soft and dense soils, respectively, in a 50% damped system were applied to the structure using FEMA P-695 [22] earthquake records. The structural properties were defined in the Matlab [23] program, and the selected critical earthquakes were applied to the structure via Matlab Simulink. The damping ratio and period of the isolated system were optimized to minimize the structure's total acceleration. In the critical earthquake analyses performed using the obtained optimal parameters, the effect of isolator mobility on control performance was evaluated for a 50% damped system on soft and dense soils.

2 Methodology

2.1 Design of a Structure with Soil-Structure Interaction and Seismic Base Isolator

Seismic isolators are vibration-damping systems designed to reduce structural acceleration and displacement. In the design of a seismically isolated structure that accounts for structure-soil interaction (SSI), various parameters representing the soil's influence are included. Figure 1 shows the mathematical model of a multi-story structure that accounts for structure-soil interaction.

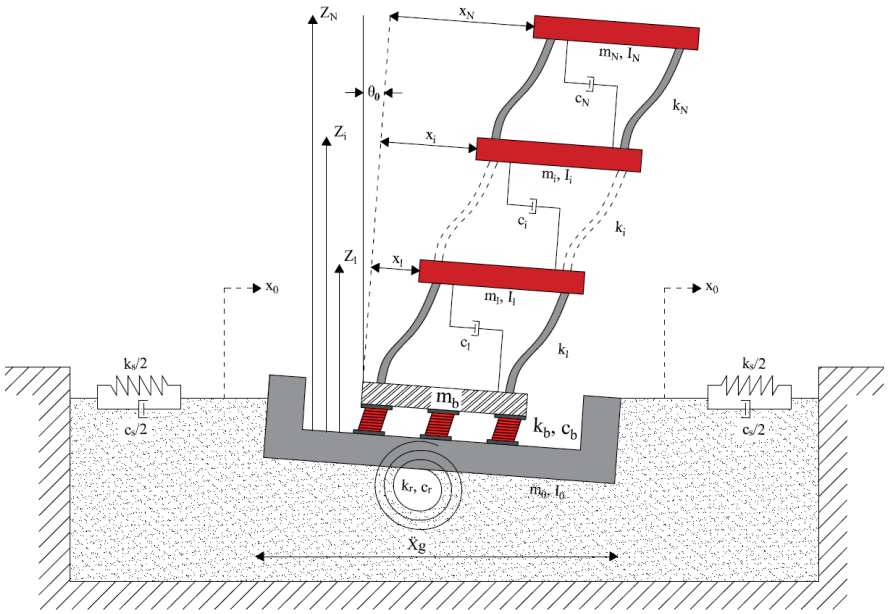


Fig. 1. A structural model of a base-isolated multi-story building that accounts for SSI effects.

In Figure 1, the index "*b*" represents the properties of the isolator, "*i*" represents the floor number, "*z*" represents the floor heights, and "*N*" represents the number of the topmost floor of the structure; the structure has floors ranging from $i = 1$ to N. "I" represents the floor inertia, "*m*" represents the mass, "*k*" represents the stiffness, and "*c*" represents the damping coefficient. The mass and stiffness of the foundation are denoted by "m_0" and "I_0," respectively. The floor's swaying is denoted by the index "s," and its rocking by the index "r." The floor's swaying stiffness and damping coefficient are denoted by k_s and c_s, respectively, while the floor's rocking stiffness and damping coefficient are denoted by k_r and c_r. The total mass of the structure (m_{total}) is calculated by summing the mass of the isolator (m_b) and the total mass of the structure ($m_{structure}$), as shown in Equation 1. The mass of the isolator story is assumed to be equal to the mass of one story of the structure.

$$m_{total} = m_b + m_{structure} \tag{1}$$

It alters the system's degrees of freedom, natural period, stiffness, and damping calculations. In a single-degree-of-freedom (SDOF) system, the isolator and structure move together, exhibiting common stiffness, natural period, and damping. For a single-degree-of-freedom (SDOF) structure with an isolator, the system's period (T_b) is given in Equation 2, its stiffness (k_b) in Equation 3, and its damping coefficient c_b) in Equation 4. In these equations, the system's damping ratio is denoted by ζ_b and its natural angular frequency by w_b.

$$T_b = \frac{2\pi}{w_b} \tag{2}$$

$$k_b = m_{total} \times w_b^2 \tag{3}$$

$$c_b = 2 \times \zeta_b \times m_{total} \times w_b \tag{4}$$

The elements representing the system's behavior under structure-soil interaction are presented in a matrix in Equation 5. $X(t)$ denotes the system's displacement, and N denotes the floor number.

$$X(t) = \begin{bmatrix} X_b \\ X_1 \\ X_2 \\ \vdots \\ X_{N-1} \\ X_N \\ X_0 \\ \theta_0 \end{bmatrix} \tag{5}$$

The matrix containing the layer masses (M_v) is shown in Equation 6, the matrix obtained by multiplying these by the layer heights (MZ) is shown in Equation 7, and the mass matrix (m^*) multiplied by the ground acceleration in the fundamental equation of motion is shown in Equation 8.

$$[M_v] = \begin{bmatrix} m_b \\ m_1 \\ m_2 \\ \vdots \\ m_{N-1} \\ m_N \end{bmatrix} \tag{6}$$

$$[MZ] = \begin{bmatrix} m_b z_1 \\ m_1 z_1 \\ m_2 z_2 \\ \vdots \\ m_{N-1} z_{N-1} \\ m_N z_N \end{bmatrix} \tag{7}$$

$$m^* = \begin{bmatrix} m_b \\ m_1 \\ m_2 \\ \vdots \\ m_{N-1} \\ m_N \\ m_0 + m_b + \left(\sum_{i=1}^{N} m_i\right) \\ \left(\sum_{i=1}^{N} m_i z_i\right) + m_b z_1 \end{bmatrix} \tag{8}$$

The mass, stiffness, and damping matrices M, K, and C of the isolated system are given in Equation 9, Equation 10, and Equation 11, respectively. In its simplest form, the motion is formulated in Equation 12.

$$[M]=\begin{bmatrix}[M_f] & [M_v] & [MZ]\\ [M_v]^T & m_0+m_b+\left(\sum_{i=1}^{N} m_i\right) & \left(\sum_{i=1}^{N} m_i z_i\right)+m_b z_1\\ [MZ]^T & \left(\sum_{i=1}^{N} m_i z_i\right)+m_b z_1 & \left(\sum_{i=1}^{N} m_i {z_i}^2\right)+m_b {z_1}^2 I_0+I_b+\sum_{i=1}^{N} I_i\end{bmatrix} \quad (9)$$

$$[K]=\begin{bmatrix}[K_f] & 0 & 0\\ 0\ldots 0 & k_s & 0\\ 0\ldots 0 & 0 & k_r\end{bmatrix} \quad (10)$$

$$[C]=\begin{bmatrix}[C_f] & 0 & 0\\ 0\ldots 0 & c_s & 0\\ 0\ldots 0 & 0 & c_r\end{bmatrix} \quad (11)$$

$$[M]\ddot{X}(t)+[C]\dot{X}(t)+[K]X(t)=-[m^*]\ddot{X}_g \quad (12)$$

In the equations, the gravitational acceleration is denoted by $\ddot{X}_g$, the displacement of the system by $X(t)$, the velocity by $\dot{X}(t)$, and the acceleration by $\ddot{X}(t)$. The values M_f, K_f, and C_f represent the mass, stiffness, and damping coefficient matrices that result when an isolator system is added to a fixed-base structure.

2.2 Optimization with Jaya Algorithm

The Jaya algorithm [20], developed by Rao, is a meta-heuristic algorithm that generates new solutions based on the best and worst results relative to the objective function during the solution search process. Since it does not involve any algorithmic factors, it quickly reaches the optimal solution. It generates solution vectors using a single solution-generation equation. Equation 13 presents the solution-generation equation for the Jaya algorithm.

$$X_{i,new} = X_{i,j} + rand(X_{i,best} - |X_{i,j}|) - rand(X_{i,worst} - |X_{i,j}|) \quad (13)$$

In the equation, each new candidate solution is denoted by $X_{i,new}$. After the initial solution matrix is constructed, the best and worst solutions in the solution matrix, based on the objective function, are denoted by $X_{i,best}$ and $X_{i,worst}$, respectively. The jth candidate solution for the ith variable is denoted by $X_{i,j}$.

The objective of the optimal design of an isolator-based system subjected to SSI is to minimize the maximum total acceleration transmitted to the superstructure. The system's displacement is constrained as shown in Equation 14 to protect the isolator from deformation caused by large displacements. The objective function of the problem is shown in Equation 15.

$$g(x) = \max(|X|) \quad (14)$$

$$f(x) = Minimize\left(\max\left(\left|\ddot{X}\right|\right)\right) \quad (15)$$

3 Numerical Example

In this study, based on the example structural model shown in Figure 1, a single-degree-of-freedom structural model was used in which the superstructure was assumed to be rigid and to move in unison with the isolator. Two cases were investigated in which the structure was located on soft and dense soil types with two degrees of freedom. Optimal design parameters were obtained for the structure with a damper limit of 50% and displacement limits of 30 and 40 cm under the critical displacement record. The critical earthquake records were selected from the FEMA earthquake records identified as critical for soft and dense soil types in the studies by Ocak et al. [12], which used the same 10-story structural model for a system with 50% damping. For soft soil, this is the NORTHR/MUL279 earthquake that occurred in 1994, and for dense soil, it is the DUZCE/BOL090 earthquake record that occurred in 1999. The structural model properties were taken from the study by Liu et al. [24]. The floor-to-floor height of the structure was set at 4 m, the floor mass at 980 metric tons, and the foundation mass at 1,960 metric tons. The isolator and floor inertia moments were set at 1.31×10^8 kg·m², and the foundation inertia moment at 1.96×10^8 kg·m². The stiffness between the first and tenth floors of the structure ranges from 2,130 to 1,870 MN/m, while the damping coefficient ranges from 42.6 to 34.7 MNs/m. The soil properties used in the study are presented in Table 1.

Table 1. Soil and foundation characteristics [24].

Soil Type	c_s (Ns/m)	k_s (N/m)	k_r (N/m)	c_r (Ns/m)
Soft Soil	$2.19x10^8$	$1.91x10^9$	$7.53x10^{11}$	$2.26x10^{10}$
Dense Soil	$1.32x10^9$	$5.75x10^{10}$	$1.91x10^{13}$	$1.15x10^{11}$

The Jaya algorithm was used to optimize the structure. During optimization, the system period with isolators was optimized within the range of 1–5 s, and the damping ratio was optimized within the range of 1–50%. The population size was set to 10, the number of iterations to 1,000, and the isolator displacement limits to 30 and 40 cm. The optimum parameters obtained from optimization are presented in Table 2. The displacement and total acceleration values of the structure, derived from the critical seismic analysis performed for soft and dense soil conditions using these optimum parameters, are shown in Table 3.

The displacement and total acceleration plots for the isolated system corresponding to a 40 cm displacement limit, which yielded the best results for both soft and dense soil, are presented in Figure 2 and Figure 3, respectively.

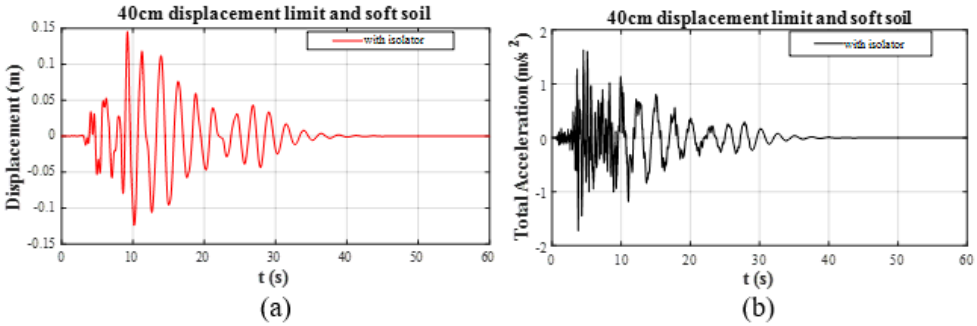


Fig. 2. Analysis plots for soft soil for a system with a 40 cm displacement limit, subjected to the NORTHR/MUL279 earthquake record: (a) Structural displacement, (b) Total structural acceleration.

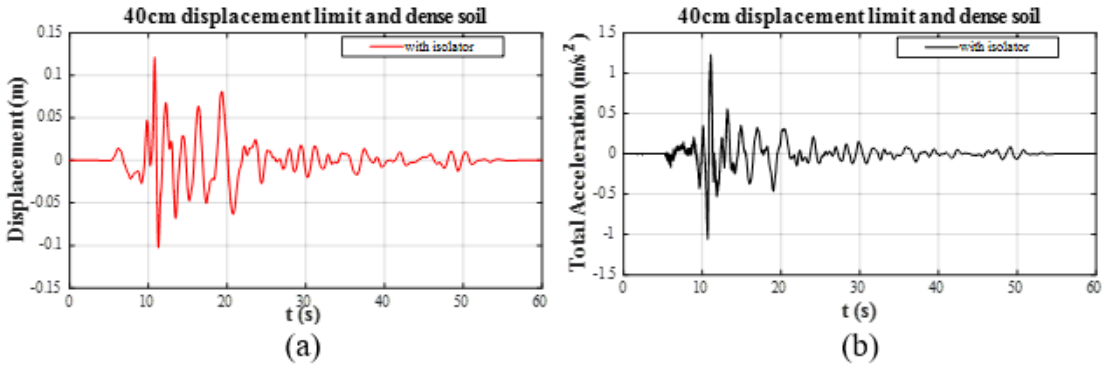


Fig. 3. Analysis plots for dense soil a system with a 40 cm displacement limit, subjected to the DUZCE/BOL090 earthquake record: (a) Structural displacement, (b) Total structural acceleration.

Table 2. Optimum parameters.

Soil Type		Isolator movement limit	
		30cm	40cm
Soft Soil	ζ_b	0.0604	0.1533
	T_b (s)	2.8927	2.1512
Dense Soil	ζ_b	0.1313	0.5000
	T_b (s)	5	3.6485

Table 3. Optimum responses.

Soil Type	Limit	Displacement (m)		Total Acceleration (m/s^2)	
		Without isolator	With isolator	Without isolator	With isolator
Soft Soil	30cm	0.2112	0.1724	6.5254	1.0799
	40cm		0.1454		1.7259
Dense Soil	30cm	0.3264	0.2020	16.7801	0.4863
	40cm		0.1211		1.2278

4 Discussion and Conclusions

In this study, an investigation was conducted into the effect of isolator mobility on control performance when the structure-soil interaction (SSI) effect is taken into account, using a single-degree-of-freedom (SDOF) structure model in which the superstructure is assumed to be rigid. In a previously tested structural model [21], for a system with 50% damping, when the SSI effect was not considered, a 30 cm isolator displacement was sufficient for the different displacement capacities of 30, 40, and 50 cm isolators; in this study, however, the contribution of the SSI effect to control performance was investigated for isolator movement limits of 30 cm and 40 cm, which had yielded the best results without the SSI effect. In this study, an investigation was conducted into the effect of isolator mobility on control performance when the structure-soil interaction (SSI) effect is taken into account, using a single-degree-of-freedom (SDOF) structure model in which the superstructure is assumed to be rigid. In

a previously tested structural model, for a system with 50% damping, when the SSI effect was not considered, a 30 cm isolator displacement was sufficient for the different displacement capacities of 30, 40, and 50 cm isolators; in this study, however, the contribution of the SSI effect to control performance was investigated for isolator movement limits of 30 cm and 40 cm, which had yielded the best results without the SSI effect. For an isolator displacement capacity of 40 cm, the total acceleration of the structure decreased by 73% for soft soil and by 92% for dense soil. At a displacement limit of 30 cm, the total acceleration of the isolated system decreased by 83% for soft soil and by 97% for dense soil. A 40 cm displacement capacity proved more effective in reducing acceleration, while a 30 cm displacement limit was more effective in reducing displacement. The findings indicate that accounting for the SSI effect in an isolator-based system can increase the system's required isolator displacement for a given damping ratio compared to a scenario where the SSI effect is not considered; conversely, the SSI effect can reduce the required period in a system with relatively greater isolator displacement limit. In the optimization performed for the isolator-based system, the required system period was longer for both soil types in the 30 cm limited system, which has relatively lower mobility. As the isolator's mobility limit increased, the required period decreased, but the demand for damping ratio increased. The findings indicate that accounting for the SSI effect in an isolator-based system can increase the system's required isolator displacement for a given damping ratio compared to a scenario where the SSI effect is not considered; conversely, the SSI effect can reduce the required period in a system with relatively greater isolator displacement.

References

1. Léger, N., Rizzian, L., & Marchi, M. (2017). Reliability-based design optimization of reinforced concrete structures with elastomeric isolators. *Procedia engineering*, *199*, 1193-1198.
2. Ocak, A., Bekdaş, G., & Nigdeli, S. M. (2024). Optimum design of hybrid base isolation and tuned liquid damper systems for structures under near-fault ground motions. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, *10*(4), 04024061.
3. Charmpis, D. C., Komodromos, P., & Phocas, M. C. (2012). Optimized earthquake response of multi-storey buildings with seismic isolation at various elevations. *Earthquake Engineering & Structural Dynamics*, *41*(15), 2289-2310.
4. Ghosh, S. K., Kashyap, I., & Das, D. (2026). Machine learning-enhanced metaheuristic optimization of lead rubber bearings for inter-story isolated buildings under seismic load. *ENGINEERING Structure and Civil Engineering*, *20*(1), 36-79.
5. Mehri, F., Mollaei, S., Farsangi, E. N., Babaei, M., & Ghahramani, F. (2023). Application of a novel optimization algorithm in design of lead rubber bearing isolation systems for seismic rehabilitation of building structures. *International Journal of Engineering, Transactions C: Aspects*, *36*(3), 594-603.
6. Tsipianitis, A., Spachis, A., & Tsompanakis, Y. (2022). Combined optimization of friction-based isolators in liquid storage tanks. *Applied Sciences*, *12*(19), 9879.
7. El Ouardani, A., & Tbatou, T. (2024). Seismic isolators layout optimization using genetic algorithm within the pymoo framework. *Civil Engineering Journal*, *10*(8), 2517-2535.

8. Radkia, S., Rahnavard, R., Tuwair, H., Gandomkar, F. A., & Napolitano, R. (2020, October). Investigating the effects of seismic isolators on steel asymmetric structures considering soil-structure interaction. In *Structures* (Vol. 27, pp. 1029-1040). Elsevier.
9. Spyrakos, C. C., Maniatakis, C. A., & Koutromanos, I. A. (2009). Soil–structure interaction effects on base-isolated buildings founded on soil stratum. *Engineering Structures*, *31*(3), 729-737.
10. Kelly, J. M. (1991). *Shake table tests of long period isolation system for nuclear facilities at soft-soil sites*. Argonne National Laboratory.
11. Fosoul, S. A., & Tait, M. J. (2021). Soil-pile-structure interaction effects on seismic demands and fragility estimates of a typical Ontario highway bridge retrofitted with fiber reinforced elastomeric isolator. *Soil Dynamics and Earthquake Engineering*, *151*, 106967.
12. Ocak, A., Nigdeli, S. M., & Bekdaş, G. (2023, October). Optimization of the base isolator systems by considering the soil-structure interaction via metaheuristic algorithms. In *Structures* (Vol. 56, p. 104886). Elsevier.
13. Jasim, F. A., Jasim, N. A., & Al-Hussein, A. A. (2024). Assessment of Soil-Structure Interaction Effects on Seismic Behavior of Isolator and Mass Damper Equipped Buildings. *Mathematical Modelling of Engineering Problems*, *11*(2), 325.
14. Öner, E., Aydoğan, A., Yurdakul, M., Ergün, M., & Artar, M. (2026). Seismic performance of prestressed girder bridges considering soil-pile-structure interaction and isolation systems. *Structural Engineering and Mechanics*, *97*(2).
15. Mellas, N., Mellas, M., Mabrouki, A., Hebbache, K., & Alemdar, Z. F. (2026). A novel study on seismic response of base-isolated wind turbine taking the soil-structure interaction impact. *Marine Georesources & Geotechnology*, *44*(5), 1178-1195.
16. Geem, Z.W., Kim, J.H., Loganathan, G.V.: A new heuristic optimization algorithm: harmony search, Simulation, 76(2), 60-68 (2001).
17. Holland, J.H.: Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor, MI (1975).
18. Kennedy, J., Eberhart, R. C.: Particle swarm optimization, Proceedings of IEEE International Conference on Neural Networks No. IV, 27 November–1 December Perth, IEEE Service Center, Piscataway, NJ, 1942–1948 (1995).
19. Saremi, S., Mirjalili, S., & Lewis, A. (2017). Grasshopper optimisation algorithm: theory and application. *Advances in engineering software*, *105*, 30-47.
20. Rao, R. V. (2016). Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems. *International Journal of Industrial Engineering Computations*, *7*, 19-34.
21. Ocak, A., Nigdeli, S. M., Bekdaş, G., Kim, S., & Geem, Z. W. (2022). Optimization of seismic base isolation system using adaptive harmony search algorithm. *Sustainability*, *14*(12), 7456.
22. FEMA P-695. Quantification of Building Seismic Performance Factors. Washington.
23. MathWorks Inc., 2018, MATLAB R2018a. Natick, MA, USA.
24. Liu, M. Y., Chiang, W. L., Hwang, J. H., & Chu, C. R. (2008). Wind-induced vibration of a high-rise building with tuned mass damper including soil-structure interaction. *Journal of Wind Engineering and Industrial Aerodynamics*, *96*(6-7), 1092-1102.

Evolution of Artificial Intelligence Paradigms and Emerging Architectures for Geohazard Prediction

Hao Xu[1]

[1] Xi'an University of Technology, Xi'an, Shaanxi 710048, China

Abstract. Intelligent geohazard prediction is evolving from static identification toward dynamic evolution modeling and physics-informed integration. Focusing on landslides, earthquakes, deep underground engineering hazards, and land subsidence, this paper reviews the technological evolution from conventional machine learning and deep neural networks to geospatial foundation models. It examines the transition from static hazard detection and spatial susceptibility assessment to dynamic prediction, including displacement forecasting and the identification of seismic and microseismic signals. The spatiotemporal modeling characteristics of convolutional neural networks, Transformers, state space models, and graph neural networks are compared, with particular attention to their capabilities in spatial feature extraction, long-term temporal dependency modeling, and complex spatial interactions. Furthermore, the application potential of multi-source data fusion, generative downscaling, and physics-informed neural networks is discussed in relation to geohazard monitoring and prediction. Key challenges, including cross-regional generalization, physical consistency, uncertainty quantification, and false-alarm control, are systematically summarized. Finally, future directions in cascading multi-hazard modeling, physics-informed prediction, and operational early warning are explored, emphasizing the integration of multi-source sensing, artificial intelligence, and digital twin technologies to support more reliable geohazard prediction and risk management.

Keywords: Geohazard prediction; Artificial intelligence; Spatiotemporal modeling; Geospatial foundation models; Physics-informed neural networks; Early warning systems.

1 Introduction

Geohazard prediction spans rapid seismic signals, long-term slope deformation and land subsidence. Satellite observations, ground sensors and monitoring archives supply diverse inputs, but heterogeneous sampling, uncertain physical parameters and computational demands complicate interpretation. Artificial intelligence (AI) offers complementary ways to extract spatial features and model temporal dependence; integrating these capabilities with process understanding is central to data-driven Earth science [1].

This review relates hazard mapping, spatial susceptibility assessment and temporal forecasting to neural architectures, foundation models and physics-informed approaches. Applications cover landslides, earthquakes, underground engineering

hazards and land subsidence. The organizing question is how model structure, data and physical information determine what predictions can support, particularly across sites and in warning decisions. The studies illustrate methodological capabilities rather than establish a universally preferable architecture through a common benchmark.

2 Geohazard Modeling Tasks

2.1 Hazard Identification and Mapping

Hazard identification delineates observed extent and compiles inventories from imagery and terrain data. Landslide4Sense combines multispectral observations, elevation and slope for benchmark segmentation [2]. A regional ResNet34–U-Net study compares deep learning with conventional mapping methods [3]. Feature selection can reduce input redundancy while retaining segmentation performance in the evaluated settings [4]. Class imbalance and complex boundaries remain assessment challenges.

2.2 Spatial Susceptibility Assessment

Susceptibility describes relative likelihood across spatial units without specifying event timing. Grid cells and slope units represent terrain and geology differently, with comparative performance depending on the study setting [5]. This task differs from forecasting future deformation.

2.3 Temporal Deformation Forecasting

Forecasting must represent nonstationary deformation and delayed environmental forcing. Reservoir-level changes and rainfall can be associated with seasonal, step-like displacement [6]. Creep-based failure-time estimates depend on deformation assumptions and external forcing [7]. Decomposition–forecast–reconstruction methods separately model trend and fluctuating components [8]. Displacement accuracy alone establishes neither failure timing nor warning reliability. Table 1 distinguishes task targets and evaluation criteria; results are not directly interchangeable.

Table 1. Distinct geohazard modeling tasks and evaluation criteria.

Task	Prediction target	Typical inputs	Evaluation
Hazard identification and mapping	Observed extent and boundaries	Optical/SAR imagery; terrain	IoU, F1, precision, recall
Spatial susceptibility	Relative likelihood; no event timing	Terrain, lithology, land use, rainfall	ROC-AUC, PR-AUC
Temporal forecasting	Future displacement or deformation	GNSS/InSAR series; environmental drivers	RMSE, MAE, R^2; interval coverage

3 AI Architectures and Complementary Modeling Approaches

3.1 Neural Architectures for Spatiotemporal Modeling

Convolutional neural networks (CNNs) extract local and multiscale features from images or spectrograms; Transformers model context through attention. Swin Transformer uses shifted local windows to exchange information across high-resolution imagery [9]. These mechanisms balance local structure, context and computation differently, but architecture alone does not determine mapping accuracy.

Long short-term memory (LSTM) networks and gated recurrent units (GRUs) retain information through gated recurrence for displacement forecasting [10]. Attention offers an alternative representation of temporal context. Adaptive variable and time-step weighting can represent changing environmental lags in reservoir landslides [11]. Mamba uses selective state-space modeling [12]; VMD-Mamba combines variational mode decomposition with sequence modeling for step-like displacement [13]. Comparisons remain conditional on the selected sites, inputs and forecasting setup.

Graph neural networks (GNNs) aggregate information over specified edges connecting monitoring stations or slope units. Adjacency and environmental similarity offer alternative graph definitions for susceptibility assessment [14]. These definitions constrain represented relationships and do not automatically imply physical causation.

3.2 Pretraining and Generative Modeling

Geospatial foundation models learn representations before task-specific adaptation. Prithvi-EO-2.0 combines multitemporal observations with temporal and location embeddings, reporting transfer to landslide segmentation with limited labels [15]. Deployment depends on memory, energy, computation and software–hardware compatibility [16]. SeisMoLLM transfers a text-pretrained language model to seismic waveform tasks through waveform encoding, alignment and parameter-efficient adaptation [17]. These approaches pursue transfer learning through different modalities and pretraining designs.

Conditional diffusion can downscale precipitation and correct bias, potentially supplying finer environmental inputs for rainfall-driven hazard analysis [18], rather than directly forecasting hazards. Figure 1 provides a conceptual overview of selected neural architectures, complementary modeling approaches and illustrative tasks considered in this review. Physics-informed learning introduces physical priors into features, architectures or training objectives and is discussed further in Section 5. These approaches can be combined with different architectures; the task examples do not imply fixed model–task assignments.

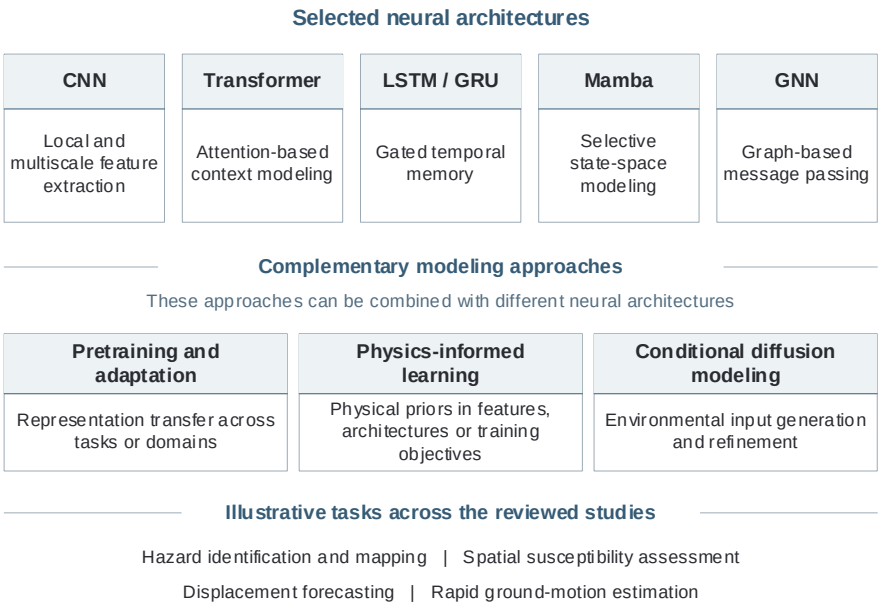


Fig. 1. Conceptual overview of selected neural architectures and complementary modeling approaches for geohazard analysis. Task examples are illustrative and non-exclusive; no performance ranking or fixed architecture–task correspondence is implied.

4 Applications in Representative Geohazard Settings

4.1 Landslides

Multitemporal interferometric synthetic aperture radar (InSAR) complements static inventories and susceptibility maps with deformation observations, as demonstrated in Three Gorges Reservoir hazard mapping [19]. Physics-informed studies use shear-zone thermal, hydraulic and mechanical equations to model basal temperature as a stability-related proxy [20], or incorporate Newmark displacement into coseismic susceptibility estimation [21]. These outputs connect observations with processes but differ in meaning: neither a modeled proxy nor susceptibility establishes a validated prediction of imminent failure.

4.2 Earthquakes

Early warning interprets signals after P-wave arrival. DLPGA estimates peak ground acceleration from a short single-station window. Japanese and Chilean tests reported higher prediction correlations and lower error dispersion with 3 s inputs than the conventional Pd method using 6 s [22]. These are observation windows, not measured warning lead times. EQTransformer jointly detects events and picks P- and S-wave phases using attention [23]. MST-FuseNet combines multiscale and temporal features

to address noise and domain shift in phase picking [24]. These complementary tasks have different outputs and datasets, precluding a common performance ranking.

4.3 Underground Engineering Hazards

Microseismic monitoring records fracture-related activity relevant to rockburst and mine-floor water-inrush assessment. Interpreting event patterns, waveforms and water pathways requires geological and mechanical context. A field study combined monitoring, mechanical analysis and numerical simulation to investigate floor fracturing and inrush risk [25]. Such evidence supports hazard interpretation but does not establish a transferable warning rule.

4.4 Land Subsidence and Sinkholes

Monitoring must distinguish deformation from atmospheric effects and decorrelation noise. InSAR time series offer wide-area observations, while identifying groundwater-related drivers requires independent evidence. Deformation detection and collapse forecasting remain distinct: an anomaly needs physical interpretation and validation before supporting a warning recommendation.

5 Reliability and Warning Decision Support

Multi-hazard analysis requires dependencies beyond a single prediction target. A Bayesian-network framework combines physical dependencies and remote sensing to jointly estimate post-earthquake cascading hazards and impacts [26]. These explicitly specified dependencies should not be conflated with causal discovery from observations alone.

Among the complementary approaches outlined in Fig. 1, physics-informed neural networks incorporate governing-equation, boundary and initial-condition residuals into training [27]. Imbalanced optimization can weaken these soft constraints, which do not guarantee physically consistent extrapolation. Phys-LSTM integrates discretized geomechanical equations into its recurrent formulation and objective. A six-site comparison reported mean site-level RMSE of approximately 1.78 mm and the lowest between-site RMSE standard deviation among compared models, 1.34 mm [28]. This mean is not an upper bound for each site, and the comparison does not establish generalization to all geological settings.

Operational assessment must jointly consider detection, false alarms and usable lead time. Validated prediction intervals and cross-sensor checks can flag observations for review, but deviations do not establish impending failure. Warning requires site-specific thresholds, expert interpretation and independent events and non-events for validation. Evaluation across sites, noise conditions and extreme forcing tests transfer beyond training conditions.

Figure 2 distinguishes the incorporation of physical information during modeling from subsequent reliability assessment. Where applicable, equation and boundary-

condition residuals can inform physics-based checks, alongside uncertainty assessment and cross-sensor consistency. Warning recommendations also require site-specific thresholds and expert interpretation. Multi-hazard reasoning and digital twins are scenario-dependent extensions requiring additional observations and physical information. Connected sensing and edge computation may support timely processing; digital twins require coupled observations and process models. Their contribution must be evaluated through warning performance rather than assumed operational readiness.

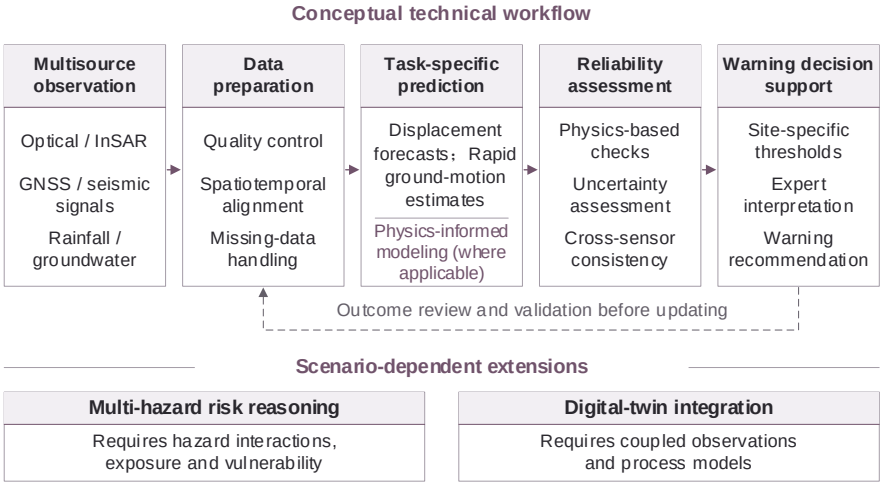


Fig. 2. Conceptual workflow for AI-assisted warning decision support. Solid arrows indicate the main processing sequence. Physics-informed modeling incorporates physical information where applicable, while physics-based checks may include equation and boundary-condition residuals. Warning recommendations combine reliability assessment, site-specific thresholds and expert interpretation. The dashed arrow denotes outcome review and validation before updating. Lower modules are scenario-dependent extensions. The workflow is a conceptual synthesis; warning dissemination, preparedness and response are outside the scope shown.

6 Conclusion

AI supports mapping, susceptibility assessment and temporal forecasting through complementary spatial, temporal and graph representations. Pretraining, generative inputs and physical constraints extend modeling options, but case-specific improvements do not establish warning reliability. Progress requires matching predictions to hazard processes, validating transfer and assessing uncertainty, false alarms and usable lead time. Multisource sensing and digital twins merit integration when independently evaluated decision-support gains justify their complexity.

References

1. Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat: Deep learning and process understanding for data-driven Earth system science. Nature 566, 195–204 (2019).
2. Ghorbanzadeh, O., Xu, Y., Ghamisi, P., Kopp, M., Kreil, D.: Landslide4Sense: Reference benchmark data and deep learning models for landslide detection. IEEE Transactions on Geoscience and Remote Sensing 60, 1–17 (2022).
3. Prakash, N., Manconi, A., Loew, S.: Mapping landslides on EO data: Performance of deep learning models vs. traditional machine learning models. Remote Sensing 12(3), 346 (2020).
4. Ahmad, A., Karakuş, O., Rosin, P.L.: Sequential feature selection for efficient landslide segmentation from multi-spectral data. Frontiers in Remote Sensing 7, 1877713 (2026).
5. Ba, Q., Chen, Y., Deng, S., Yang, J., Li, H.: A comparison of slope units and grid cells as mapping units for landslide susceptibility assessment. Earth Science Informatics 11(3), 373–388 (2018).
6. Yang, H., Song, K., Chen, L., Qu, L.: Hysteresis effect and seasonal step-like creep deformation of the Jiuxianping landslide in the Three Gorges Reservoir region. Engineering Geology 317, 107089 (2023).
7. Chang, C., Wang, G.: Creep of clayey soil induced by elevated pore-water pressure: Implication for forecasting the time of failure of rainfall-triggered landslides. Engineering Geology 296, 106461 (2022).
8. Gao, Y., Chen, X., Tu, R., Chen, G., Luo, T., Xue, D.: Prediction of landslide displacement based on the combined VMD-stacked LSTM-TAR model. Remote Sensing 14(5), 1164 (2022).
9. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin Transformer: Hierarchical vision transformer using shifted windows. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, pp. 9992–10002. IEEE (2021).
10. Nava, L., Carraro, E., Reyes-Carmona, C., Puliero, S., Bhuyan, K., Rosi, A., Monserrat, O., Floris, M., Meena, S.R., Galve, J.P., Catani, F.: Landslide displacement forecasting using deep learning and monitoring data across selected sites. Landslides 20, 2111–2129 (2023).
11. Weng, N., Fan, L., Tang, L.: An end-to-end deep learning framework with nonstationary weighting strategy for enhancing reservoir landslide displacement prediction. Scientific Reports 16, 24320 (2026).
12. Qu, H., Ning, L., An, R., Fan, W., Derr, T., Liu, H., Xu, X., Li, Q.: A survey of Mamba. ACM Transactions on Intelligent Systems and Technology 17(5), Article 116, 1–43 (2026).
13. Zhu, Q., Chen, W., Zeng, Q., Li, Y., Liu, S.: Step-type landslide displacement prediction method based on VMD-Mamba algorithm. Natural Hazards 121(8), 9339–9362 (2025).
14. Zhang, Q., He, Y., Zhang, L., Lu, J., Gao, B., Yang, W., Chen, H., Zhang, Y.: A landslide susceptibility assessment method considering the similarity of geographic environments based on graph neural network. Gondwana Research 132, 323–342 (2024).
15. Szwarcman, D., Roy, S., Fraccaro, P., Gíslason, Þ.E., Blumenstiel, B., Ghosal, R., et al.: Prithvi-EO-2.0: A versatile multitemporal foundation model for Earth observation applications. IEEE Transactions on Geoscience and Remote Sensing 64, 4400120 (2026).
16. Sang, H., Zhang, L., Chen, T., Guo, W., Zhang, Z.: Onboard deployment of remote sensing foundation models: A comprehensive review of architecture, optimization, and hardware. Remote Sensing 18(2), 298 (2026).

17. Wang, X., Liu, F., Su, R., Wang, Z., Fang, L., Zhou, L., Bai, L., Ouyang, W.: SeisMoLLM: Advancing seismic monitoring via cross-modal transfer with pretrained large language model. Geophysical Research Letters 53(8), e2025GL118505 (2026).
18. Aich, M., Hess, P., Pan, B., Bathiany, S., Huang, Y., Boers, N.: Conditional diffusion models for downscaling and bias correction of Earth system model precipitation. Geoscientific Model Development 19, 1791–1808 (2026).
19. Zhou, C., Cao, Y., Hu, X., Yin, K., Wang, Y., Catani, F.: Enhanced dynamic landslide hazard mapping using MT-InSAR method in the Three Gorges Reservoir Area. Landslides 19, 1585–1597 (2022).
20. Moeineddin, A., Seguí, C., Dueber, S., Fuentes, R.: Physics-informed neural networks applied to catastrophic creeping landslides. Landslides 20(9), 1853–1863 (2023).
21. Dahal, A., Lombardo, L.: Towards physics-informed neural networks for landslide prediction. Engineering Geology 344, 107852 (2025).
22. Liu, Y., Zhao, Q., Wang, Y.: Peak ground acceleration prediction for on-site earthquake early warning with deep learning. Scientific Reports 14, 5485 (2024).
23. Mousavi, S.M., Ellsworth, W.L., Zhu, W., Chuang, L.Y., Beroza, G.C.: Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nature Communications 11, 3952 (2020).
24. Ma, R., Wang, M., Zhou, L., He, J., Zeng, Z.: Noise-robust and transferable seismic phase picking for regional earthquake monitoring using a multi-scale temporal fusion network. Journal of Seismology 30(5), Article 88 (2026).
25. Ma, K., Sun, X.Y., Tang, C.A., Yuan, F.Z., Wang, S.J., Chen, T.: Floor water inrush analysis based on mechanical failure characters and microseismic monitoring. Tunnelling and Underground Space Technology 108, 103698 (2021).
26. Li, X., Xu, S.: Scalable variational learning for noisy-OR Bayesian networks with normalizing flows for complex cascading disaster systems. npj Natural Hazards 2, 30 (2025).
27. Raissi, M., Perdikaris, P., Karniadakis, G.E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 378, 686–707 (2019).
28. Zhou, J., Ju, N., He, C., Xie, M.: A physics-informed online learning framework for landslide displacement prediction. Applied Sciences 16(12), 6003 (2026).

Microwave-Induced Fracturing and Drillability Bottleneck Breakthrough in Deep Hard Rock: A Multiphysics Coupling and Geomechanical Perspective

Jiawei Lv[1]

[1] Xi'an University of Technology, Xi'an, Shaanxi 710048, China
Lvjiawei200@163.com

Abstract. Deep hard-rock excavation is constrained by high strength, high hardness, and high abrasivity, which cause rapid cutter wear, low penetration rate, and high specific energy. Microwave irradiation provides a non-contact pre-conditioning method by exploiting mineral dielectric contrasts to generate selective volumetric heating and thermal-expansion mismatch. This paper condenses the electromagnetic mechanism, thermo-hydro-mechanical damage amplification, mechanical degradation, and drillability evidence of microwave-treated hard rock. Results show that thermal-expansion mismatch drives intergranular-to-transgranular cracking, while pore-steam pressure and liquid-nitrogen quenching further amplify damage. Microwave treatment reduces static strength and fracture toughness, enhances impact energy dissipation, lowers Cerchar abrasivity, and decreases specific energy under controlled laboratory or reduced-scale cutter-test conditions. These quantitative improvements are condition-dependent and should not be directly extrapolated to full-scale deep hard-rock excavation under true-triaxial confinement. Deep confinement and arc discharge remain engineering barriers, requiring hybrid microwave–hydraulic or cryogenic strategies and arc-suppressed antenna systems.

Keywords: microwave-induced fracturing; deep hard rock; thermal expansion mismatch; THM coupling; drillability; specific energy

1 Introduction

Deep tunnels, mines, and geothermal wells increasingly encounter hard rocks such as granite, basalt, and diorite. These formations exhibit high in-situ stress, high strength, and high abrasivity, causing severe wear of TBM disc cutters, roadheader picks, and drill bits [1]. Conventional mechanical breaking therefore reaches an efficiency bottleneck: penetration rate decreases, downtime increases, and specific energy rises sharply.

Microwave-induced rock fracturing is a promising non-contact pre-conditioning technology [2]. Microwave energy penetrates rock and heats minerals selectively according to dielectric loss. Strongly absorbing minerals expand rapidly, while weakly absorbing minerals remain cooler, producing thermal-expansion mismatch and microcrack nucleation. With continued irradiation, cracks coalesce and macroscopic strength degrades. This paper summarizes the key mechanisms of microwave rock breaking, its THM coupling effects, mechanical degradation, drillability improvement, and remaining engineering barriers. The overall research roadmap, linking deep hard-rock drillability bottlenecks to the microwave fracturing mechanism, multi-scale characterization, multiphysics simulation, and engineering optimization, is presented in Fig. 1.

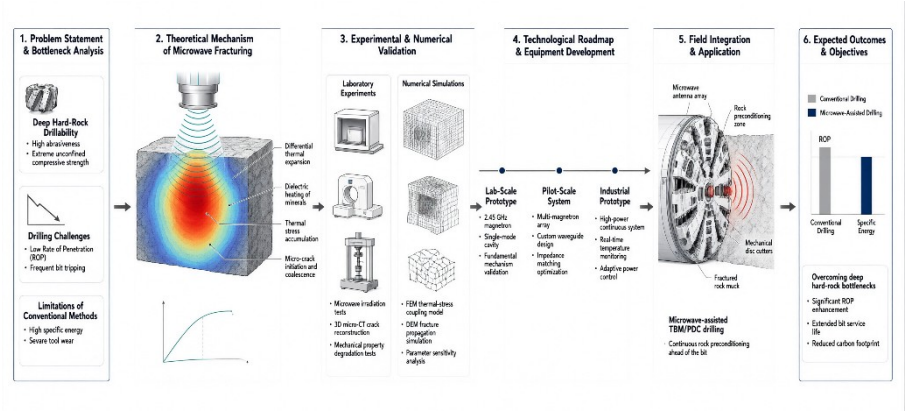


Fig. 1. Research framework and technological roadmap of microwave-induced rock fracturing for overcoming deep hard-rock drillability bottlenecks.

2 Microwave Energy Deposition and Thermomechanical Damage

2.1 Dielectric Heating and Penetration Depth

Microwave heating is governed by the complex relative permittivity:

$$\varepsilon^* = \varepsilon' - j\varepsilon'' \quad (1)$$

where ε'represents energy storage and ε''represents dielectric loss. The volumetric power density is:

$$p_v = 2\pi f \varepsilon_0 \varepsilon'' |E|^2 \quad (2)$$

where $f$is frequency, $\varepsilon_0$is vacuum permittivity, and $E$is electric field strength. Thus, heat generation depends on both mineral loss factor and local field intensity [3].

Rocks are heterogeneous multiphase materials. Metallic minerals such as pyrite and magnetite are strong microwave absorbers, whereas quartz, feldspar, and calcite are relatively microwave-transparent [4]. Table 1 summarizes typical dielectric parameters at 2.45 GHz.

Table 1. Dielectric parameters and microwave response of typical minerals at 2.45 GHz.

Mineral	ε'	ε''	Microwave response
Quartz	3.75–4.5	<0.01	Transparent, weak heat source
Biotite	5.9	0.002–0.044	Weak absorber
Hornblende	7.37–14.45	0.026–0.324	Moderate absorber
Pyrite	8.0–23.6	0.75–11.2	Strong absorber, hotspot
Magnetite	14.5	2.5	Extremely strong absorber

Penetration depth is the distance over which power density decays to 1/e, about 36.8%, of its surface value. Lower frequency gives deeper penetration. Therefore, 915 MHz microwaves are suitable for larger-volume pre-conditioning, while 2.45 GHz microwaves produce more concentrated shallow heating and thermal spalling [5].

2.2 Thermal Expansion Mismatch and Crack Evolution

Under irradiation, absorbing minerals heat rapidly and expand, while adjacent transparent minerals form cold barriers. This produces local stress incompatibility, or thermal-expansion mismatch. The damage evolves across scales:

a. Grain-boundary cracking: Thermal shear and tensile stresses concentrate at

mineral interfaces, initiating intergranular microcracks [6].

b. Transgranular cracking: In quartz-bearing rocks, the α–β quartz transition near 573 ℃ produces an abrupt volume change of about 0.8%–1.2%. Combined with thermal stress, this can fracture quartz grains directly [7].

c. Crack coalescence: Intergranular and transgranular cracks connect, forming a cross-scale damage network that reduces macroscopic strength.

The progressive failure sequence is: dehydration and microcrack initiation at 100–300 ℃; visible thermal spalling at 400–600 ℃; and local melting above approximately 1000 ℃. The spalling stage is usually the optimal intervention window for mechanical cutting, while excessive melting wastes energy and may clog cracks.

3 THM Coupling and Mechanical Degradation

The complete pathway from electromagnetic energy deposition, through thermo-hydro-mechanical coupling, to macroscopic mechanical degradation is illustrated in Fig. 2.

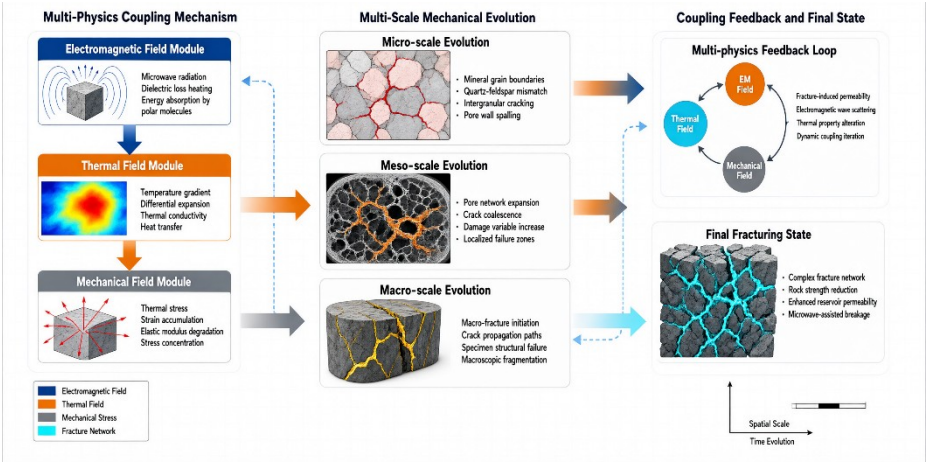


Fig. 2. Flowchart of multi-scale multi-physics coupling and mechanical evolution in microwave-induced rock fracturing.

3.1 Power, Time, Water, and Quenching Effects

Equivalent total microwave energy does not produce equivalent damage. High power–short duration irradiation is generally more effective than low power–long

duration irradiation [8]. High power creates steep local temperature gradients before heat conduction can dissipate energy, producing concentrated thermal stress and denser cracks. Low power tends to heat rock more uniformly, weakening thermal mismatch. Therefore, there is an optimal energy-efficiency range, often evaluated through weakening per unit microwave energy.

Water is a major damage catalyst. As a polar molecule, water undergoes dipole rotation in the microwave field and heats efficiently. This is especially important for rocks with low intrinsic microwave absorptivity, such as sandstone and dense limestone. In low-permeability hard rock, pore water vaporizes rapidly but cannot escape easily, generating high steam pressure. This endogenous pressure superposes with thermal-expansion stress, producing a THM “micro hydraulic fracturing” effect and increasing macroporosity and microcrack density [9]. Consequently, water-saturated rocks usually weaken more than dry rocks.

Cooling path also controls final damage. Natural air cooling preserves cracks but allows partial stress relaxation. Liquid-nitrogen quenching at −196 °C produces extreme surface contraction and secondary thermal shock. In conventional heating–LN_2 cooling cycle tests on granite, specimens heated to 300 °C and cooled by LN_2 for 20 cycles showed P-wave velocity reductions of 39%–60% and uniaxial compressive-strength reductions of up to approximately 49% under unconfined laboratory conditions [10]. These values are protocol-specific and should not be directly generalized to microwave-treated deep confined rock masses. This hot–cold hybrid route is especially useful for extremely hard rock and geothermal drilling.

3.2 Static and Dynamic Mechanical Degradation

Microwave-induced crack networks can substantially reduce static mechanical parameters measured under controlled laboratory conditions, including uniaxial compressive strength, Brazilian tensile strength, and point load index, depending on mineralogy, microwave power, irradiation time, water content, and cooling path. These specimen-scale reductions should not be directly extrapolated to deep confined rock masses. Fracture toughness also decreases because the fracture process zone broadens and crack paths become more tortuous. Weakened rock bridges allow cracks to propagate with lower driving energy. Soft-computing models can further predict residual strength and fracture toughness from easily measurable inputs such as UCS, BTS, and P-wave velocity, providing low-cost surrogates for field assessment [11].

Under dynamic loading, microwave-pre-damaged rocks exhibit enhanced energy dissipation [12]. Although rock strength generally increases with strain rate, the pre-existing microcrack network consumes impact energy through new crack formation, crack branching, fragment sliding, and friction. As a result, transmitted energy decreases, reflected and dissipated energies increase, and fragments become finer and more uniform. These laboratory observations suggest that microwave pre-conditioning may reduce oversized muck and cutter impact load during TBM or drill-bit excavation, although full-scale confined validation is required.

4 Drillability Verification and Engineering Constraints

4.1 Cerchar Abrasivity and Cutter Wear

Rock abrasivity controls cutter consumption in tunnel boring. The Cerchar abrasivity index is measured by dragging an HRC 55 steel stylus across a rough rock surface under 70 N load for 10 mm; the diameter of the wear flat is recorded in units of 0.1 mm [13]. CAI depends not only on compressive strength but also on equivalent quartz content, grain size, and cementation degree. Table 2 gives a common abrasivity classification.

Table 2. CAI-based rock abrasivity grading.

Abrasivity grade	CAI range	Typical lithologies
Extremely low	0.3–0.5	Mudstone, shale
Low to medium	0.5–2.0	Limestone, sandstone
High	2.0–4.0	Hard sandstone, some igneous rocks
Extremely high	4.0–6.0	Granite, basalt, quartzite

Microwave treatment reduces abrasivity by weakening grain-boundary continuity. Thermal cracks sever the load-bearing connection of hard minerals, so grains are more likely to detach as blocks rather than produce intense sandpaper-like abrasive wear. This mechanism is expected to reduce CAI in highly abrasive granites under matched laboratory conditions, although CAI tests do not fully represent cutter wear in confined, jointed, and continuous excavation. However, the improvement is mineral-dependent: quartz-rich rocks with low microwave absorptivity may require stronger irradiation to activate phase-transition thermal shock.

4.2 Reduced-scale cutter tests and specific energy

Full-scale linear cutting machines use real disc cutters, commonly 17 or 19 inches in diameter, to evaluate excavability. The key metric is specific energy:

$$SE = \frac{W}{V} \tag{3}$$

where $W$is cutting work and $V$is fragmented volume. Untreated high-strength granite or basalt requires high thrust to form a crushed zone. After microwave irradiation, the internal damage network acts as crack guides. Adjacent groove cracks coalesce more easily, allowing large flakes to spall at lower thrust. In the reduced-scale TBM cutter tests on Chifeng basalt reported by Lu et al. [14], specimens were irradiated at 2100 W for up to 15 min and naturally cooled; the specific energies associated with the driving and rolling forces decreased by approximately 53% and 52%, respectively, under controlled reduced-scale laboratory conditions. The concurrent increase in rock removal and more extensive crack coalescence support a shift from energy-intensive plastic comminution toward lower-energy brittle spalling in these tests. These results suggest that microwave pre-conditioning may increase penetration rate and reduce cutter wear under favorable laboratory or reduced-scale conditions, although full-scale confined validation is still required.

4.3 Deep Confinement, Arc Discharge, and Hybrid Strategies

At depth, rock is under true-triaxial in-situ stress. Confining pressure suppresses microwave-induced fracturing because thermal cracks are mainly Mode I opening cracks. High lateral stress forces new cracks closed, preventing coalescence into connected networks. Achieving equivalent damage therefore requires higher power density or longer irradiation.

This creates a second barrier: arc discharge. High-intensity microwaves delivered through waveguide or horn antennas near the rock face can ionize air around conductive minerals or field-distortion zones. Arcing blocks energy transmission, degrades impedance matching, and may damage the antenna [15]. Field-scale application therefore requires arc-suppressed antenna arrays, dynamic impedance matching, and robust operation in dusty, high-pressure environments.

Hybrid pre-fracturing can mitigate confinement. Microwave irradiation can first create near-wellbore thermal cracks, altering local permeability and stress continuity. Subsequent hydraulic fracturing then initiates at lower breakdown pressure and may

form complex three-dimensional fracture networks [16]. When coupled with cryogenic fluids, breakdown pressure can decrease substantially, improving stimulated reservoir volume. Additional hybrids, such as microwave-assisted laser kerfing or ultrasonic vibration, may further enhance local stress concentration and spalling efficiency.

5 Conclusions

a. Microwave rock breaking is driven by dielectric-loss contrast among minerals. Thermal-expansion mismatch produces intergranular cracking first, then transgranular cracking, especially when quartz phase transition is activated.

b. Pore-water vaporization and cryogenic cooling can amplify THM damage under controlled laboratory conditions. Cyclic heating–LN_2 cooling tests on granite, rather than microwave irradiation tests, showed P-wave velocity reductions of 39%–60% and uniaxial compressive-strength reductions of up to approximately 49% after 20 cycles [10]. These protocol-specific results should not be generalized to microwave-treated deep confined rock masses.

c. Microwave treatment reduces static strength, fracture toughness, and dynamic failure threshold, while enhancing impact energy dissipation and improving muck size distribution.

d. Reduced-scale cutter tests indicate lower specific energy under controlled laboratory conditions, while microwave-induced grain-boundary cracking is expected to reduce abrasivity where verified by CAI tests. The quantitative improvements depend on rock type, microwave power, exposure time, moisture state, cooling path, and test scale, and should not be presented as universally representative of deep, confined rock excavation.

e. Deep triaxial confinement and arc discharge remain major engineering barriers. Hybrid microwave–hydraulic or cryogenic strategies and arc-suppressed antenna systems are necessary for field-scale deployment.

References

[1] Rauch, R., Goliasch, R., Käsling, H., & Thuro, K. (2018). Does the in-situ stress state have an effect on tool wear in deep TBM tunnelling? Haben Gebirgsspannungen Einfluss auf den Abrasivverschleiß im tiefliegenden TBM-Vortrieb?. Geomechanics and Tunnelling, 11(2), 169-178.

[2] Gao, M. Z., Yang, B. G., Xie, J., Ye, S. Q., Liu, J. J., Liu, Y. T., ... & Zhou, X. M. (2022). The mechanism of microwave rock breaking and its potential application

to rock-breaking technology in drilling. Petroleum Science, 19(3), 1110-1124.

[3] Wei, W., Shao, Z., Zhang, Y., Qiao, R., & Gao, J. (2019). Fundamentals and applications of microwave energy in rock and concrete processing–A review. Applied Thermal Engineering, 157, 113751.

[4] He, C. L., Ma, S. J., Su, X. J., Mo, Q. H., & Yang, J. L. (2015). Comparison of the microwave absorption characteristics of hematite, magnetite and pyrite. Journal of Microwave Power and Electromagnetic Energy, 49(3), 131-146.

[5] Frasch, L. L., McLean, S. J., & Olsen, R. G. (1998). Electromagnetic properties of dry and water saturated basalt rock, 1-110 GHz. IEEE transactions on geoscience and remote sensing, 36(3), 754-766.

[6] Yuan, Y., Shao, Z., Qiao, R., Fei, X., & Wu, D. (2021). Quantifying damage evolution within olivine basalt based on crack propagation behavior under microwave irradiation. International Journal of Damage Mechanics, 30(10), 1617-1641.

[7] Li, Q., Li, X., & Yin, T. (2021). Effect of microwave heating on fracture behavior of granite: An experimental investigation. Engineering Fracture Mechanics, 250, 107758.

[8] Li, J., Kaunda, R. B., Arora, S., Hartlieb, P., & Nelson, P. P. (2019). Fully-coupled simulations of thermally-induced cracking in pegmatite due to microwave irradiation. Journal of Rock Mechanics and Geotechnical Engineering, 11(2), 242-250.

[9] Zhao, Q. H., Zhao, X. B., Zheng, Y. L., Li, J. C., He, L., & Zou, C. J. (2020). Microwave fracturing of water-bearing sandstones: heating characteristics and bursting. International Journal of Rock Mechanics and Mining Sciences, 136, 104495.

[10] Xue, Y., Wang, L., Liu, J., Ranjith, P. G., Gao, F., Cai, C., & Xie, H. (2024). Experimental study on the effect of heating and liquid nitrogen-cooling cyclic treatment on mechanical properties and fracturing characteristics of granite. International Journal of Rock Mechanics and Mining Sciences, 176, 105691.

[11] Hu, X., Liao, D., Ma, D., Xie, S., Xie, N., Hu, H., & Gong, X. (2023). Machine learning models for predicting rock fracture toughness at different temperature conditions. Case Studies in Construction Materials, 19, e02622.

[12] Wang, S., Xu, Y., Xia, K., & Tong, T. (2021). Dynamic fragmentation of microwave irradiated rock. Journal of Rock Mechanics and Geotechnical Engineering, 13(2), 300-310.

[13] Alber, M., Yaralı, O., Dahl, F., Bruland, A., Käsling, H., Michalakopoulos, T. N., ... & Özarslan, A. (2014). ISRM suggested method for determining the abrasivity of rock by the CERCHAR abrasivity test. In The ISRM suggested methods for rock characterization, testing and monitoring: 2007-2014 (pp. 101-106). Cham: Springer International Publishing.

[14] Lu, G., Ding, C., Zhou, J., Liu, H., & Liu, C. (2023). Influences of microwave irradiation on rock-breaking efficiency of a reduced-scale TBM cutter. Applied Sciences, 13(8), 4713.

[15] Ma, Z., Zheng, Y., Li, J., Zhao, X., & Zhao, J. (2024). Enhancing rock breakage efficiency by microwave fracturing: A study on antenna selection. Energy, 288, 129876.

[16] Zhu, Z., Xie, J., Du, Y., Ren, L., Yang, B., Zhou, Q., & Gao, M. (2026). Theoretical feasibility assessment of hydraulic fracturing with microwave assistance: Insights from a coupled electromagnetic-thermo-hydro-mechanical (ETHM) phase-field simulation. International Journal of Rock Mechanics and Mining Sciences, 204, 106563.

AI-Based Prediction of Mechanical Properties of Microwave-Pretreated Rock and Optimization of Rock Breaking

Zhuo Pan [1]

[1] Xi'an University of Technology, Xi'an, Shaanxi 710048, China

Abstract. Microwave pretreatment weakens rock by exploiting differences in the dielectric properties, thermal expansion coefficients, and phase-transition behaviors of constituent minerals, which induce localized temperature rises, thermal stress concentrations, and the initiation and propagation of microcracks. This paper reviews the intelligent prediction of the mechanical properties of microwave-pretreated rock and the optimization of rock-breaking processes. The physical and micromechanical mechanisms of microwave-induced rock damage are first analyzed, followed by a discussion of artificial intelligence methods, including XGBoost, Random Forest, CatBoost, metaheuristic optimization algorithms, and physics-informed neural networks, for predicting uniaxial compressive strength, tensile strength, and other mechanical parameters. In addition, SHAP, Sobol, and Morris methods are examined for model interpretation and sensitivity analysis. The applications of artificial intelligence to TBM-specific energy evaluation, cutter spacing and penetration optimization, Cerchar abrasivity index assessment, cutter wear control, and intelligent monitoring based on acoustic emission and microseismic signals are further discussed. The integration of artificial intelligence with microwave-induced damage mechanisms, multiphysics modeling, and engineering parameter optimization provides an effective technical framework for performance prediction, parameter design, and engineering applications of microwave-assisted rock breaking.

Keywords: Microwave pretreatment, Rock mechanics, Artificial intelligence, Machine learning, TBM, Rock-breaking optimization

1 Introduction

Rock properties govern tunnel boring machine (TBM) cutting and wear [1]. Microwave pretreatment modifies cutting behavior [2] through selective heating and thermally induced cracking [3].

Artificial intelligence (AI) supports uniaxial compressive strength (UCS) prediction after microwave treatment [4]. Random Forest (RF) strength models [5] and CatBoost Cerchar abrasivity index (CAI) models [6] provide methods for untreated rock; physics-informed neural networks (PINNs) incorporate governing equations [7].

Specific energy (SE) prediction [8] and cutter replacement classification [9] connect properties to engineering decisions. Microwave applications must distinguish pretreatment design inputs from post-treatment measurements [4].

Linking property prediction to cutting decisions, this paper reviews microwave damage, AI prediction, TBM optimization, model interpretation, and acoustic emission (AE) and microseismic (MS) monitoring.

2 Microwave-Induced Rock Damage

2.1 Dielectric Response and Selective Heating

Mineral dielectric properties govern microwave absorption. The complex relative permittivity (ε_r) can be expressed as:

$$\varepsilon_r = \varepsilon' - j\varepsilon'' \tag{1}$$

where ε' and ε'' denote the real part and dielectric loss factor, and j the imaginary unit. Heat conduction also raises low-loss mineral temperatures.

Table 1 compares mineral responses within the cited evidence. Silicate contrasts refer to the granite model [3]; oxide and sulfide comparisons refer to measured samples [10].

Table 1. Qualitative comparison of mineral responses

Mineral	Dielectric response	Thermal or modeling relevance	Source
Magnetite	More lossy than the cited pyrite sample	Sample-specific dielectric inputs	[10]
Pyrite	Less lossy than the cited magnetite sample	Specimen-dependent dielectric inputs	[10]
Biotite	Strongest loss among model silicates	Differential heating and thermal mismatch	[3]
Amphibole	Loss below biotite in the model	Differential heating and thermal mismatch	[3]
K-feldspar	Loss below biotite and amphibole	Thermal contrast with adjacent minerals	[3]
Quartz	Lowest loss among model silicates	Phase transition if heating is sufficient	[3]

2.2 Thermal Stress and Cracking

Mineral expansion and stiffness contrasts generate constrained thermal stress. For small strain, isotropy, and a constant expansion coefficient, the thermal strain tensor (ε^T) is:

$$\varepsilon^T = \alpha(T - T_0)I \tag{2}$$

where α denotes linear expansion coefficient; T and T_0, current and reference temperatures; and I, the identity tensor. Stress requires constitutive constraints; anisotropy, variable α, and phase changes require extensions.

The quartz $\alpha-\beta$ transition and local melting contribute only when the relevant temperature conditions are reached [3]. Figure 1 summarizes the mechanism.

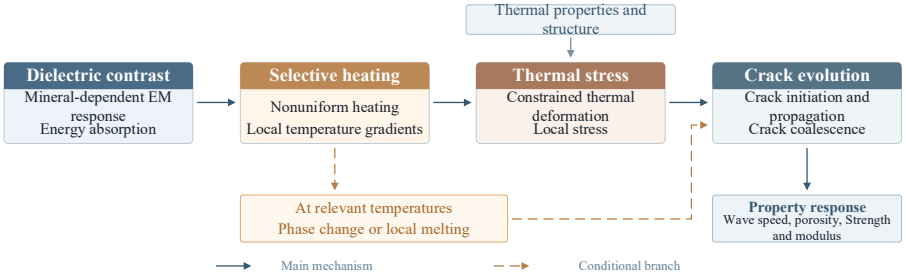


Fig. 1. Microwave-induced damage mechanisms and rock response. EM denotes electromagnetic; dashed arrows indicate the temperature-dependent phase-change or melting branch.

Microwave irradiation decreased X-ray micro-computed tomography (micro-CT) values and longitudinal-wave velocity in Fangshan granite [11]. Wave velocity and porosity support strength prediction [4]. These mechanisms guide inputs: mineralogy and initial properties characterize rock; irradiation conditions describe treatment; post-treatment wave velocity and porosity indicate damage.

3 AI Prediction of Rock Mechanical Properties

Microwave-treated UCS (UCSA) prediction provides direct evidence [4]; models for untreated rock require treatment-specific validation.

3.1 Mechanical Prediction and Optimization

XGBoost adds regularized trees; its objective function ($\mathcal{L}^{(t)}$) at boosting iteration t is [12]:

$$\mathcal{L}^{(t)}=\sum_{i=1}^{n} l\left(y_i,\hat{y}_i^{(t-1)}+f_t(x_i)\right)+\Omega(f_t) \tag{3}$$

where n is sample count; x_i and y_i denote inputs and observations; $\hat{y}_i^{(t-1)}$ is the preceding prediction, f_t the new tree, and l differentiable loss. The term $\Omega(f)$ penalizes tree complexity.

RF combines randomized trees for UCS prediction [5]. CatBoost uses ordered boosting and categorical-feature processing [13].

Using a dataset of 78 rock specimens, Gu et al. combined pied kingfisher optimization (PKO) with XGBoost and reported a fitting R^2 of 0.978 for predicting UCSA [4]. Post-treatment measurements support residual-strength assessment; prospective parameter design instead requires inputs available before irradiation or separately validated estimates of treatment-induced changes.

For untreated rock, XGBoost tuning used the sparrow search algorithm (SSA) [14] and artificial bee colony (ABC) optimization with mine measurements and published data [15].

Table 2 also includes moth-flame optimization (MFO)-RF rock mass rating (RMR) classification [16] and CatBoost CAI prediction [6]. These provide engineering extensions rather than direct evidence for microwave-treated strength prediction; their different targets and datasets preclude direct algorithm ranking.

Table 2. Representative prediction models and applicability

Model	Target	Evidence scope	Application limitation	Source
SSA-XGBoost	UCS	Untreated rock	Requires treatment-specific validation	[14]
XGBoost-ABC	UCS	Deep-mine rock	Requires treatment-specific validation	[15]
PKO-XGBoost	UCSA	Microwave-treated specimens	Post-treatment inputs limit design use	[4]
MFO-RF	RMR	TBM rock mass	Classification differs from strength prediction	[16]
CatBoost	CAI	Untreated igneous rock	Requires recalibration after treatment	[6]

Brazilian tensile strength (BTS) prediction used slake durability and specific gravity for evaporitic rocks [17]. Optimized extreme learning machines predicted UCS and modulus [18]; tree and neural models predicted Mode I fracture toughness from cracked chevron-notched Brazilian disc (CCNBD) tests [19]. Microwave transfer requires matched specimen and loading conditions and separate validation for each mechanical target.

3.2 PINNs and Multiphysics Constraints

PINNs constrain training with governing equations [7], informed by coupled electromagnetic-thermal-mechanical models [3]. With initial and boundary penalties combined, the total loss ($\mathcal{L}_{\text{total}}$) can be written schematically as:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{data}} + \lambda_{\text{PDE}}\mathcal{L}_{\text{PDE}} + \lambda_{\text{BC}}\mathcal{L}_{\text{BC}} \tag{4}$$

where $\mathcal{L}_{\text{data}}$, $\mathcal{L}_{\text{PDE}}$, and $\mathcal{L}_{\text{BC}}$ measure data, governing-equation, and combined initial and boundary residuals; nonnegative λ_{PDE} and λ_{BC} weight the constraints, with unit data weight.

Inferring residual strength from predicted temperature or stress also requires a calibrated damage-to-property relation and numerical or experimental validation, given PINN training difficulties [20].

3.3 Model Interpretation and Sensitivity

SHAP explains feature contributions relative to a background baseline [21], including microwave-strength predictions [4], without establishing causality.

Sobol decomposes output variance assuming independent inputs [22]; correlated properties require dependence-aware analysis. Morris elementary effects screen influential inputs but cannot distinguish nonlinearity from interactions [23].

4 TBM Optimization and Monitoring

Microwave-assisted TBM optimization connects weakening to cutting loads, energy, and maintenance. Reduced-scale experiments provide direct evidence [2].

4.1 Energy, Cutting Parameters, and Wear

Mechanical specific energy (SE_{mech}) is work per removed volume [24]. For circular excavation of radius R and advance l, approximately constant thrust and torque give:

$$SE_{\text{mech}} = \frac{F_v l + M\theta}{\pi R^2 l} \tag{5}$$

where F_v is thrust, M torque, and θ rotation (rad); $\pi R^2 l$ is removed volume. Units N, m, and $N \cdot m$ yield J/m^3. Variable loads require integration; local tests use actual removed volume.

In reduced-scale basalt cutting tests, microwave treatment at 2100 W for 15 min reduced mean driving and rolling forces by approximately 33% and 32%, respectively, relative to untreated rock [2]. These laboratory reductions do not establish whole-system energy savings.

Chaos-assisted SSA-RF predicted SE from conventional cutter experiments using strength, cutter geometry, spacing, and penetration [8]. Microwave transfer requires treatment descriptors.

Particle Flow Code (PFC) simulations link efficiency to the relative positions of cutters and treated zones [25]. Conventional tunnel data show that spacing-to-penetration recommendations depend on rock and water conditions [26]; microwave applications require separate calibration.

Optimization varies microwave power, exposure time, cutter spacing, and penetration to screen energy, load, and wear tradeoffs. Schemes must satisfy equipment and thermal limits and undergo independent cutting validation; whole-system comparisons include microwave electricity and maintenance.

CAI models for untreated igneous rock need recalibration after microwave treatment [6]. Kernel support vector machine (KSVM) replacement classification used Jilin water conveyance tunnel data [9]. CAI alone cannot determine cutter life.

4.2 Monitoring and Feedback

AE combined with digital image correlation (DIC) characterized damage in microwave-treated basalt [27]. The *b*-value describes the event-size distribution, but its different evolution in treated and untreated specimens requires calibration for treatment and loading conditions.

RA is rise time divided by peak linear amplitude; AF is threshold counts divided by duration [27]. Fracture-class boundaries require acquisition-specific calibration.

RF and K-nearest neighbors (KNN) classified mine microseismic and blasting events [28]. Classification performance alone does not demonstrate warning capability; event-based splits help limit leakage between training and evaluation.

Warnings additionally require prediction windows, false-alarm and miss rates, and lead times; monitoring supports verified updates to models and operation.

Figure 2 links treatment inputs, calibrated property prediction, cutting screening, and independent evaluation, with monitoring feedback.

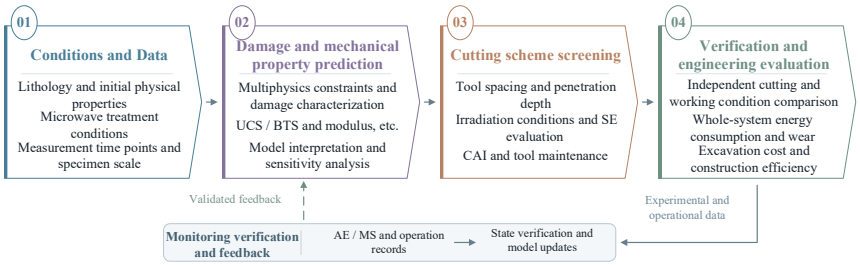


Fig. 2. Prediction and evaluation of microwave-assisted rock breaking

Scaling requires consistent specimen and irradiation records, treatment-depth checks, and coordination between irradiation and cutters.

Whole-system comparisons include microwave supply, TBM, and auxiliary electricity over consistent boundaries. Distinguishing electrical input, microwave output, and absorbed energy is essential; mechanical SE alone cannot establish energy or cost savings.

5 Conclusions

Selective microwave absorption and thermal-expansion contrasts induce cracking, while phase-change contributions depend on temperature and material state. Ensemble models support post-microwave UCS prediction, but other mechanical targets and PINN property mappings require treatment-specific validation. SHAP, Morris, and Sobol aid interpretation and sensitivity analysis under stated assumptions without proving causality. Engineering optimization requires independent cutting tests, CAI-wear calibration, and whole-system energy and cost evaluation. AE and MS monitoring likewise needs calibrated indicators and assessment of false alarms, missed events, and warning lead time. Future work should link irradiation, mechanical, cutting, and monitoring datasets and assess uncertainty across scales within the proposed mechanism-informed AI framework.

References

1. Jeong, H., Choi, S., Lee, Y.-K.: Evaluation of cutting performance of a TBM disc cutter and Cerchar abrasivity index based on the brittleness and properties of rock. Applied Sciences 13(4), 2612 (2023).
2. Lu, G., Ding, C., Zhou, J., Liu, H., Liu, C.: Influences of microwave irradiation on rock-breaking efficiency of a reduced-scale TBM cutter. Applied Sciences 13(8), 4713 (2023).
3. Su, X., Li, D., Zhao, J., Wang, M., Su, X., Zhou, A.: Numerical simulation of microwave-induced cracking and melting of granite based on mineral microscopic models. International Journal of Minerals, Metallurgy and Materials 31(7), 1512–1524 (2024).
4. Gu, Z., Xiong, X., Yang, C., Cao, M.: Investigation of micro-scale damage and weakening mechanisms in rocks induced by microwave radiation and their associated strength reduction patterns: Employing meta-heuristic optimization algorithms and extreme gradient boosting models. Mathematics 12(18), 2954 (2024).
5. Wang, M., Wan, W., Zhao, Y.: Prediction of the uniaxial compressive strength of rocks from simple index tests using a random forest predictive model. Comptes Rendus. Mécanique 348(1), 3–32 (2020).
6. Wani, S.R., Sharifi Teshnizi, E., Jalota, S.: Correlation between Cerchar abrasivity index and geotechnical properties of igneous rocks: A comprehensive analysis using machine learning algorithms and interpretative analysis. Measurement 257, 118989 (2026).
7. Raissi, M., Perdikaris, P., Karniadakis, G.E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 378, 686–707 (2019).
8. Zhou, J., Liu, Z., Li, C., Du, K., Yang, H.: Cutting-edge approaches to specific energy prediction in TBM disc cutters: Integrating COSSA-RF model with three interpretative techniques. Underground Space 22, 241–262 (2025).
9. Liu, Y., Huang, S., Wang, D., Zhu, G., Zhang, D.: Prediction model of tunnel boring machine disc cutter replacement using kernel support vector machine. Applied Sciences 12(5), 2267 (2022).
10. Gutiérrez-Cano, J.D., Catalá-Civera, J.M., López-Buendía, A.M., Plaza-González, P.J., Penaranda-Foix, F.L.: High-resolution detection of rock-forming minerals by permittivity measurements with a near-field scanning microwave microscope. Sensors 22(3), 1138 (2022).

11. Wang, S., Xu, Y., Xia, K., Tong, T.: Dynamic fragmentation of microwave irradiated rock. Journal of Rock Mechanics and Geotechnical Engineering 13(2), 300–310 (2021).
12. Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. ACM, New York (2016).
13. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A.: CatBoost: unbiased boosting with categorical features. In: Advances in Neural Information Processing Systems, vol. 31, pp. 6638–6648. Curran Associates, Red Hook (2018).
14. Xu, B., Tan, Y., Sun, W., Ma, T., Liu, H., Wang, D.: Study on the prediction of the uniaxial compressive strength of rock based on the SSA-XGBoost model. Sustainability 15(6), 5201 (2023).
15. Zhao, J., Li, D., Jiang, J., Luo, P.: Uniaxial compressive strength prediction for rock material in deep mine using boosting-based machine learning methods and optimization algorithms. Computer Modeling in Engineering & Sciences 140(1), 275–304 (2024).
16. Yang, B., Jahed Armaghani, D., Fattahi, H., Afrazi, M., Koopialipoor, M., Asteris, P.G., Khandelwal, M.: Optimized random forest models for rock mass classification in tunnel construction. Geosciences 15(2), 47 (2025).
17. Hassan, M.Y., Arman, H.: Comparison of six machine-learning methods for predicting the tensile strength (Brazilian) of evaporitic rocks. Applied Sciences 11(11), 5207 (2021).
18. Jin, X., Zhao, R., Ma, Y.: Application of a hybrid machine learning model for the prediction of compressive strength and elastic modulus of rocks. Minerals 12(12), 1506 (2022).
19. Wang, Y.T., Zhang, X., Liu, X.S.: Machine learning approaches to rock fracture mechanics problems: Mode-I fracture toughness determination. Engineering Fracture Mechanics 253, 107890 (2021).
20. Krishnapriyan, A.S., Gholami, A., Zhe, S., Kirby, R.M., Mahoney, M.W.: Characterizing possible failure modes in physics-informed neural networks. In: Advances in Neural Information Processing Systems, vol. 34, pp. 26548–26560. Curran Associates, Red Hook (2021).
21. Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, vol. 30, pp. 4765–4774. Curran Associates, Red Hook (2017).
22. Sobol, I.M.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Mathematics and Computers in Simulation 55(1–3), 271–280 (2001).
23. Morris, M.D.: Factorial sampling plans for preliminary computational experiments. Technometrics 33(2), 161–174 (1991).
24. Teale, R.: The concept of specific energy in rock drilling. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts 2(1), 57–73 (1965).
25. Wen, S., Du, L., Kong, Q., Zhang, M., Ding, X.: Study on microwave-assisted TBM double-edged cutter rock-breaking efficiency and its positional relationship. Eksploatacja i Niezawodność - Maintenance and Reliability 26(2), 186447 (2024).
26. Wang, W., Yan, C., Guo, J., Zhao, H., Li, G., Yao, W., Ren, T.: Improving tunnel boring machine tunneling performance by investigating the particle size distribution of rock chips and cutter consumption. Buildings 14(4), 1124 (2024).
27. Yang, J., Wang, C., Zhao, Y., Bi, J.: Acoustic emission characteristics and damage evolution of basalt by microwave irradiation. Scientific Reports 13, 6006 (2023).
28. Peng, K., Tang, Z., Dong, L., Sun, D.: Machine learning based identification of microseismic signals using characteristic parameters. Sensors 21(21), 6967 (2021).

Research on a Dynamic Site Selection and Evaluation System for Underground Gas Storage Based on Hydro-Mechanical Coupling Dynamics

Miao Zhang[1]

[1] Xi'an University of Technology, Xi'an, Shaanxi 710048, China

Abstract. Underground gas storage (UGS) facilities serve as critical infrastructure for natural gas peak shaving and strategic supply assurance, playing an irreplaceable role in global energy decarbonization. However, the high-frequency cyclic injection and production operations induce complex hydro-mechanical (HM) coupling effects, rendering traditional static-parameter-based site selection and evaluation methods inadequate for safe and efficient operations. This paper develops a theoretical dynamic site selection and evaluation framework in which coupled HM dynamics supply the indicators of a multi-criteria decision model; the objective is a reproducible engineering workflow rather than a purely descriptive review. First, the nonlinear evolution of effective stress, stress-dependent Biot coefficient, and rock fatigue damage under cyclic loading are discussed, and a saturating cycle-dependent power-law damage law is formulated to link the fatigue state of the rock to the degradation of porosity and permeability that controls well deliverability. Second, relative permeability hysteresis, the Land trapping model, and thermo-hydro-mechanical-chemical (THMC) coupling risks associated with alternative cushion gas transport are analysed in complex pore networks. Third, dynamic leakage risks of the reservoir-caprock-fault system are systematically assessed, including fatigue degradation of caprock sealing capacity and fault activation induced by pore-pressure diffusion and poroelastic stress transfer. Fourth, a 4D (three-dimensional space plus time) hydro-mechanical simulator built on the unconditionally stable fixed-stress split algorithm, together with a four-tier multiscale upscaling chain, is presented. Finally, a dynamic indicator system organized in four aspects: safety, stability, flow assurance, and economic performance, is proposed and integrated with a multi-criteria decision-making workflow combining the Analytic Hierarchy Process (AHP), entropy weighting, and the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), together with a VIKOR robustness check. The findings provide theoretical and technical foundations for scientific site selection, safe operation, and lifecycle management of UGS under complex geological conditions.

Keywords: Underground gas storage, Hydro-mechanical coupling, Dynamic site selection, 4D numerical simulation, AHP-TOPSIS.

1 Introduction

Natural gas remains the main bridge between fossil fuels and renewables, and its share of global primary energy keeps rising. Demand, however, swings sharply with season, region and extreme weather, whereas supply stays comparatively rigid. Underground gas storage (UGS)—for natural gas peak shaving as well as emerging underground hydrogen storage (UHS)—serves as the primary infrastructure capable of delivering large-scale, cross-seasonal flexibility and strategic supply security [1,2].

Depleted reservoirs dominate. Cushion gas typically occupies 20–70 % of total storage capacity; physically, it limits water invasion, maintains reservoir pressure and supplies the elastic energy needed for rapid withdrawal. Conventionally the cushion is native methane, which locks up both gas and capital. To improve the investment structure, operators are testing cheaper inert or greenhouse gases as substitute cushion gas; carbon dioxide, whose phase behaviour is the best documented of the candidates [3], is the most widely considered option, while nitrogen has so far been tested mainly in pilot projects. Between 5 and 13 MPa and at 303.15–323.15 K, CO_2 is roughly three to eight times denser than methane and one and a half to five times more viscous, both ratios increasing with pressure. This density and viscosity contrast damps gravity override and viscous fingering in the pore network and suppresses convective mixing, and so preserves the purity and calorific value of the produced gas [4].

Production is a slow, one-way pressure decline; storage operation is a high-frequency cycle of large injection and withdrawal. Over annual or seasonal cycles, pore pressure rises and falls repeatedly, so the reservoir frame, the caprock and the bounding faults are subjected to cyclic stress disturbance for decades. Traditional site selection has predominantly focused on static volumetric indicators such as porosity, permeability, and structural relief. However, long-term cyclic pressure swings induce significant hydro-mechanical damage, as recent experimental studies have demonstrated substantial permeability degradation in reservoir sandstones [5] and dynamic sealing fatigue along caprocks under multi-cycle alternating loads [6].

These observations have led the rock-mechanics and petroleum-engineering communities to a common conclusion: sealing capacity and mechanical stability measured under static conditions represent only an upper bound for a site that will be cycled for decades, so site selection and performance evaluation must rest on coupled HM dynamics rather than on static indices alone [6,7]. A high-fidelity 4D simulator coupled with multi-criteria decision analysis is, in our view, the appropriate route to a quantitative and reproducible assessment of sealing capacity, mechanical stability, and economics over the full life cycle. The proposed theoretical framework (Fig. 1) is organized as five sequential steps: (1) multi-source characterisation, in which three-dimensional seismic, well logs, core and fluid data, the injection–withdrawal history and the microseismic catalogue are assembled into a static model of the trap, caprock and faults; (2) rock-physics calibration, in which digital-core and pore-network models provide the stress-dependent Biot coefficient, the fatigue damage factor, the Land coefficient and the pseudo-relative-permeability curves; (3) 4D hydro-mechanical simulation with the fixed-stress split solver, including fault contact mechanics and the multi-cycle injection–withdrawal schedule over the project life; (4) dynamic indicator

extraction, covering caprock breakthrough pressure, fault slip tendency, wellbore shear failure and interface plastic shear strain, productivity decline, cushion-gas mixing width and near-wellbore temperature and salt risk; and (5) the decision layer, in which AHP pairwise comparison and entropy weighting feed TOPSIS ranking with a VIKOR robustness check. Steps (1) – (2) form the characterisation stage, step (3) the simulation stage and steps (4) – (5) the evaluation stage. The framework is closed by a feedback loop: when the ranked scenarios show that the operating envelope is unacceptable, the revised pressure scenarios are returned to step (3) and the simulation is repeated until a feasible envelope is obtained.

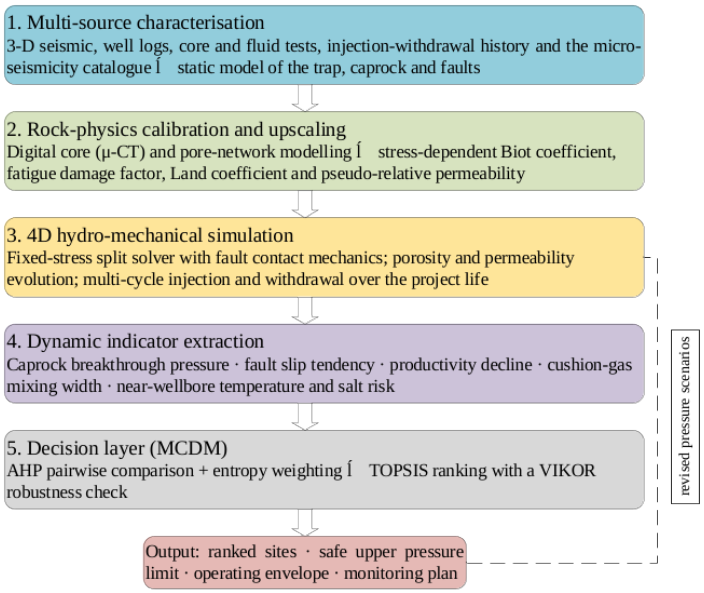


Fig. 1. Workflow of the proposed 4D dynamic site selection and evaluation framework.

2 Hydro-mechanical response and damage under cyclic injection and withdrawal

2.1 Effective stress and the stress dependence of the Biot coefficient

The deformation and failure of a porous frame are controlled by effective stress. In the classical poroelastic framework of Terzaghi and Biot, the effective stress tensor of a saturated porous medium is written in terms of the total stress tensor and the pore pressure: Throughout this paper stresses are taken positive in compression, so that $\sigma' = \sigma - \alpha p \mathbf{I}$; the opposite (tension-positive) convention only reverses the sign of every stress term and leaves the relations themselves unchanged.

$$\sigma'_{ij} = \sigma_{ij} - \alpha p \delta_{ij} \tag{1}$$

Here α is the Biot coefficient, the key dimensionless parameter that measures the strength of hydro-mechanical coupling. For an isotropic medium it follows from the drained bulk modulus of the frame, K_{dr}, and the bulk modulus of the solid grains, K_s:

$$\alpha = 1 - K_{dr}/K_s \tag{2}$$

It is widely reported that α is highly non-linear and stress-dependent under cyclic operations. During the gas withdrawal stage, the decline in pore pressure leads to an increase in the effective confining stress. Consequently, micro-cracks and compliant pores undergo mechanical closure, causing the rock frame to stiffen, the drained bulk modulus K_{dr} to increase, and α to decrease noticeably. Conversely, during high-pressure gas injection, the elevated pore pressure reduces the effective confining stress, promoting micro-crack reopening and poroelastic matrix dilation, which lowers K_{dr} and pushes α back toward unity. Ignoring this cyclic, stress-path-dependent evolution of α in a 4D geomechanical model will lead to severe underestimations of reservoir expansion strains, stress concentrations along the wellbore, and the potential for fault reactivation.

2.2 Fatigue damage and plastic strain accumulation under cyclic loading

Over several decades of operation a storage site experiences several tens of large injection–withdrawal pressure cycles, up to a few per year at high-turnover facilities, and both the reservoir rock and its caprock remain in a state of fatigue damage.

To represent this weakening in a field-scale flow model, a cycle-dependent fatigue damage factor, m, is introduced into the conventional stress-sensitivity formulation. Consistent in spirit with the cyclic-loading experiments on reservoir sandstone reported in [5], in which core permeability declines as the number of stress cycles increases, the damage factor is represented here by a saturating form. The form is an assumption for field-scale modelling rather than a fit to site-specific data, and it is written in terms of the injection–withdrawal cycle number N:

$$m(N) = m_\infty[1 - (1 + N/N_0)^{-\beta}] \tag{3}$$

where m_∞ is the asymptotic (saturation) damage factor approached as $N \to \infty$, N_0 is a reference cycle number and β is the fatigue exponent that controls the rate at which damage accumulates. Since $m(0) = 0$ and m rises monotonically to $m_\infty < 1$ for $\beta > 0$, the factor $[1 - m(N)]$ in Eq. (4) decreases from unity towards the residual value $1 - m_\infty$. The values of m_∞, N_0 and β are to be identified from cyclic permeability measurements on cores of the candidate formation following the protocol of [5]; no site-specific fit is claimed. The damage factor enters the stress-sensitive permeability law as

$$k(\sigma', N) = k_0[1 - m(N)] \exp[-\gamma(\sigma' - \sigma'_0)] \tag{4}$$

where k_0 is the permeability of the intact frame at the reference effective stress σ'_0 and γ is the stress-sensitivity coefficient of the intact frame; porosity is advanced with the volumetric strain of the same time increment, so that φ and k remain consistent. The

limiting cases follow from Eq. (3): for N = 0, or for any N when β = 0, m(N) ≡ 0 and Eq. (4) reduces to the classical reversible law $k = k_0 \exp[-\gamma(\sigma' - \sigma'_0)]$; for β > 0 the damage accumulates irreversibly and permeability falls monotonically towards $k_0(1 - m_\infty)$. Two points deserve emphasis. The irreversibility of Eqs. (3) – (4) is a cycle-accumulating, not a hysteretic, effect: the response depends on N and does not by itself produce a loading–unloading loop, which arises instead from the stress-path dependence of α (Section 2.1) and from the relative-permeability hysteresis of Section 3. Second, the law is phenomenological: it is meant to represent, not to derive, the degradation seen in cyclic tests, and the resulting loss of injection and withdrawal capacity is expected to be gradual and saturating rather than unbounded.

3 Multiphase flow and hysteresis in complex pore networks

3.1 Physical origin of hysteresis and the Land trapping model

During injection, the non-wetting phase displaces the wetting phase from the pore-throat network. During withdrawal, reservoir pressure falls and water imbibes spontaneously, re-invading the pore space and snapping off the continuous gas pathways, so that much of the gas is trapped by capillarity and remains as residual gas saturation, S_{gr}. Contact-angle hysteresis, together with pore-throat snap-off, substantially reduces the gas relative permeability during imbibition relative to that during drainage. This is the well-established hysteresis effect of multiphase flow in porous media; for CO_2 enhanced recovery and sequestration its magnitude has been measured experimentally by Sedaghatinasab et al. [8].

On the basis of extensive core drainage–imbibition experiments, Land proposed an empirical relation between the maximum historical gas saturation, S_{gi}, and the residual gas saturation reached at the end of imbibition [9]: Because trapping acts on the connected, mobile gas phase, the relation is written in terms of the effective (normalized) gas saturation $S_g^* = S_g/(1 - S_{wc})$, the form that makes S_{gr} independent of the connate-water content:

$$S_{gr} = S_{gi} /(1+C \cdot S_{gi}) \tag{5}$$

The constant C is the Land trapping coefficient, which measures the capacity of the pore network to trap the non-wetting phase. Using absolute instead of effective saturation can overestimate the trapped gas, and hence the hysteresis loss, in reservoirs with a high connate-water content.

Using Land's estimate of trapped gas, Killough and Carlson built the hysteresis models used in three-dimensional reservoir simulators. Killough anchors the imbibition branch on the Land residual saturation and reconstructs gas relative permeability by interpolation [10], whereas Carlson generates the imbibition curve by shifting the drainage curve in parallel [11]. Obtaining C and the hysteresis curves for a specific rock has been transformed by pore network modelling (PNM) and digital-core technology. PNM avoids the cost and duration of core-flood experiments and, more importantly,

reveals how the coordination number and connectivity of the pore network control macroscopic hysteresis [12]. A comprehensive comparison of the mainstream relative permeability hysteresis models, including their core physical assumptions, boundary conditions, and primary input parameters, is summarized in Table 1.

Table 1. Relative permeability hysteresis models commonly used in UGS numerical simulation

Model	Core physical assumption and application scenario	Key input parameters
Land model	Describes the single-valued relationship between residual gas saturation and initial gas saturation; applicable to strongly water-wet rocks.	Land trapping coefficient
Killough model	Based on the Land residual gas prediction, interpolates between the drainage main branch and the imbibition end point; applicable to gas-water two-phase imbibition.	Relative permeability curve of the drainage period, Land coefficient
Carlson model	Assumes that the imbibition curve is obtained by parallel translation of the drainage curve, and relies on a single experimental datum for correction.	One known experimental imbibition test point, Land coefficient
PNM upscaling model	Solves the pore-scale Navier-Stokes or quasi-static capillary equilibrium based on the real topological structure of the digital core.	Contact angle, pore-throat radius distribution, pore connectivity, coordination number

3.2 Multicomponent transport of cushion gas and thermo-hydro-mechanical-chemical risk

As noted above, CO_2 or N_2 is increasingly used as a substitute cushion gas. Injecting a non-native fluid into a depleted reservoir involves not only multiphase flow but also multicomponent gas mixing, thermodynamics and geochemistry, that is, coupled thermo-hydro-mechanical-chemical behaviour. Where CO_2 is used as cushion gas, geology, operating parameters and rock properties all control the thickness of the mixing zone, which is therefore retained as an explicit decision indicator.

Thermal effects are equally important, but their origin depends on the operation. During withdrawal, high-pressure gas expands from reservoir pressure to the much lower wellbore pressure through the completion and any downhole choke; the Joule–Thomson effect governs this throttling process, and the associated cooling of the near-wellbore zone can be large enough to promote hydrate formation and freeze-up. During high-pressure injection the injected gas is normally warmer than the formation, and heat exchange is dominated by convection with the gas stream and conduction to the rock frame, so compression heats the gas rather than cooling it. In both stages the near-wellbore temperature field is coupled to the mechanical state through thermoelastic stress.

Salt precipitation is not driven by cooling. When dry or under-saturated gas invades a water-bearing formation, water evaporates into the gas phase to satisfy the vapor-

pressure deficit between the two phases; the resident brine then becomes over-saturated and halite crystallizes in the pore throats [13]. Evaporation is itself endothermic and adds a secondary cooling term, but the controlling potential is the vapor-pressure deficit, so the risk is highest for the driest injected gases and the highest flow rates. Because salt crystals block pore throats close to the well, the permeability falls abruptly and the injectivity can be lost at a cliff edge, which is a direct threat to flow assurance; for hydrogen and CO_2-rich systems the magnitude of this damage has been quantified [14].

4 Dynamic leakage risk of the reservoir–caprock–fault system

4.1 Fatigue degradation of caprock sealing capacity

The caprock is the final barrier that prevents buoyant escape of the stored gas. Its sealing capacity is controlled mainly by the capillary breakthrough pressure, which is set by the fine pore structure of the rock. In a conventional accumulation that has been at rest for millions of years, the static breakthrough pressure is sufficient to hold the gas column. Storage operation breaks this static equilibrium.

Cyclic loading changes effective stress in the reservoir and also subjects the base of the caprock, near the reservoir–caprock interface, to repeated stress disturbance. Evidence of such fatigue degradation has moved the design of the maximum operating pressure away from empirical rules and towards dynamic criteria based on in-situ stress monitoring and limit-state analysis. A widely applied criterion is that the maximum injection pressure must not exceed a fixed fraction, usually 0.8, of the minimum principal stress.

4.2 Fault reactivation and poroelastic stress transfer

Fault rupture triggered by injection results from the combined action of pore-pressure diffusion and poroelastic stress transfer [15]. Slip risk is traditionally quantified by the slip tendency (ST), defined as the ratio of shear stress to effective normal stress acting on the fault plane, where reactivation approaches when ST reaches the critical friction coefficient of the fault gouge. Furthermore, to account for operational and geomechanical uncertainties, probabilistic fault slip potential (FSP) assessment workflows have been integrated into dynamic screening protocols [16]. Because the FSP formulation of [16] was developed for monotonic wastewater injection, its transfer to a storage site requires two corrections: the Coulomb stress change must be evaluated along the cyclic effective-stress path rather than on a single pressurization ramp, and the friction coefficient must be allowed to degrade with the number of cycles through the fatigue factor m of Eq. (3). With these corrections the same workflow yields a cycle-dependent slip-probability envelope that can be screened against the operating pressure schedule.

(1) The induced earthquake sequence at the Castor storage project, Spain. A few days after large-scale cushion-gas injection began in September 2013, a sequence of thousands of micro-earthquakes started. It peaked several weeks after injection stopped and included three events of local magnitude 4.1, which led the Spanish government to close the facility permanently [17].

(2) Microseismicity at the Hutubi storage facility, China. Since it began operating in 2013, the monitoring network has recorded clear seismicity during each seasonal injection and withdrawal cycle, with a maximum magnitude of 3.0 and hypocentres concentrated near the reservoir and down to about 6 km below it [18]. Slip tendency reached 0.43 ± 0.04 when small slip nucleated on the fault and micro-earthquake clusters appeared [19]. Joint interpretation of GPS surface deformation and geomechanical modelling confirms that bulk expansion of the gas reservoir, which produced larger horizontal extension than vertical uplift, generated the poroelastic stress perturbation that reactivated pre-existing faults nearby [20].

5 4D hydro-mechanical simulation based on the fixed-stress split algorithm

5.1 Governing equations and the coupling problem

(1) Mass conservation for each fluid phase, accounting for component exchange, compressibility and multiphase flow:

$$\partial(\varphi\rho_\alpha S_\alpha)/\partial t + \nabla\cdot[\rho_\alpha(k_{r\alpha}K/\mu\alpha)(\nabla p_\alpha - \rho_\alpha g\nabla z)] = Q_\alpha \tag{6}$$

where φ is the dynamic porosity; ρ_α, S_α, $k_{r\alpha}$ and μ_α are the density, saturation, relative permeability and viscosity of phase α (gas or water); K is the absolute permeability tensor; Q_α is the source or sink term; g is the gravitational acceleration and z is the vertical coordinate, positive upwards.

(2) Momentum balance for the solid frame, describing equilibrium of the rock mass under effective stress and pore pressure in the same compression-positive convention:

$$-\nabla\cdot(\sigma' + \alpha pI) = \rho_b g + f \tag{7}$$

α is the Biot coefficient, p is the pore pressure, I is the second-order identity tensor, ρ_b is the bulk density, g is the gravity vector and f is the body-force term. Inserting Eq. (1) into Eq. (7) shows that the momentum balance and the fluid potential of Eq. (6) use the same sign convention; had stresses been taken positive in tension, both equations would reverse sign together, so the effective-stress relation itself is unaffected.

5.2 The fixed-stress split

Among the available decoupling strategies, drained, undrained and fixed-strain splits, the fixed-stress split is widely used for deep HM simulation because of its mathematical properties and of the stability results established for linear poroelasticity [21]. The scheme is sequential: the flow and mechanics problems are solved one after the other,

and it must not be confused with a fully implicit (monolithic) formulation, in which the two sets of unknowns are assembled into a single Jacobian and advanced simultaneously. The framework proposed herein is designed to be implemented with the sequential fixed-stress algorithm of [21]; fully implicit coupling is mentioned only as the reference against which the split is benchmarked.

Within each time step the flow equation is solved first, assuming that the increment of total mean stress is zero; the updated pressure field is then used to solve mechanical equilibrium. To remove the non-physical oscillations that decoupling introduces in a strongly coupled system, a stabilization term is added to the flow equation to compensate for the change in volumetric strain. Denoting by r the iteration index within a time step, the semi-discrete pressure equation reads

$$(c_0 + \beta_{FS})\partial p^r/\partial t - \nabla\cdot[(K/\mu)\nabla p^r] = Q_f - \alpha\partial\varepsilon_v^{\,r-1}/\partial t + \beta_{FS}\partial p^{r-1}/\partial t \quad (8)$$

where c_0 is the pore and fluid compressibility, β_{FS} is the fixed-stress stabilization parameter, and the last two terms represent the drained volumetric-strain coupling and the pressure correction that removes the splitting error of the previous iteration. β_{FS} is usually taken as

$$\beta_{FS} = \alpha^2/K_{dr} \quad (9)$$

with K_{dr} the drained bulk modulus of the frame, a combination of the Lamé parameters. Because the stabilization parameter is non-negative for $\alpha \in [0, 1]$, the stabilized flow operator remains positive definite and the scheme is unconditionally stable regardless of the time-step size [21]. This result, however, is established for linear poroelasticity with constant coefficients and does not carry over unchanged to the present setting, in which α is stress-dependent, the damage factor of Eq. (3) evolves within the time step, and fault contact imposes unilateral constraints on the stress state. The argument of [21] therefore holds only locally, for the coefficients frozen at the current iteration: β_{FS} is evaluated from the current damage-degraded drained modulus, the damage update is lagged by one time step, fault contact is resolved by an active-set iteration, and unconditional stability is expected under quasi-linear conditions and verified numerically rather than inherited as a theorem.

5.3 Multiscale upscaling

Computing pore-scale hysteresis and stress concentration directly on a field grid of millions to tens of millions of cells is impractical, so microscopic mechanisms must reach the field simulator through a rigorous upscaling chain (Fig. 2). The chain crosses four spatial tiers—pore network (10^{-6}–10^{-3} m), core and representative elementary volume (10^{-2}–10^{-1} m), grid block (10^{1}–10^{2} m) and field model (10^{3}–10^{4} m)—through three upscaling steps. (i) Pore-scale PNM and digital-core simulations solve the Navier–Stokes and capillary-equilibrium problems to extract directional relative-permeability hysteresis, capillary-pressure curves and the Land coefficient. (ii) These microscopic constitutive relationships are upscaled to representative elementary volume (REV) rock-physics models, which combine the dynamic cyclic triaxial test

data with the pore-scale results to deliver the damage factor m(N) and the equivalent poroelastic stiffness tensor. (iii) Dynamic pseudo-functions (pseudo-relative-permeability and pseudo-capillary-pressure curves) and coarse-grid transmissibility multipliers are mapped onto the field-scale reservoir–caprock–fault numerical grid, bridging the micro-mechanisms to multi-decade operational forecasts. The transmissibility multipliers are essential because they carry the sub-grid heterogeneity that the coarse grid block cannot resolve.

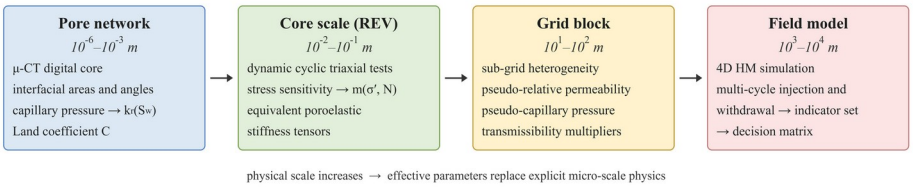


Fig. 1. Upscaling chain from pore network to field model.

6 A four-aspect dynamic indicator system and the AHP–TOPSIS decision workflow

6.1 Indicator system

The dynamic indicators extracted from the HM analysis are listed in Table 2. They cover trap and caprock safety, fault stability, reservoir deformation, well productivity, cushion-gas economics and near-wellbore THMC risk. The four primary dimensions are aspects of site performance, not spatial dimensions; the 4D character of the framework comes from the three-dimensional space plus time resolution of the simulation beneath them.

Table 1. Proposed dynamic evaluation index system for UGS site selection, organized in four aspects: safety, stability, flow assurance and economic performance

Primary Dimension	Secondary Criteria	Physical Significance & Evaluation Objective
Safety	(1) Dynamic capillary breakthrough pressure retention (P_b/P_{b0}) (2) Maximum Fault Slip Tendency (ST_{max})	Evaluates caprock sealing fatigue degradation and prevention of catastrophic induced seismicity or containment breach
Stability	(1) Cumulative plastic shear strain along caprock-reservoir interface (ε^p) (2) Wellbore shear failure and casing deformation index	Quantifies long-term fatigue deformation of the rock frame and structural integrity of the wellbore-casing-cement sheath system
Flow Assurance	(1) Relative permeability hysteresis trapping loss (S_{gr} via Land model) (2) Deliverability decay factor under cyclic stress (m)	Assesses the sustained deliverability of injection-withdrawal wells and mitigates cliff-edge injectivity loss caused by THMC effects
Economic Performance	(1) Working-gas-to-total-capacity ratio (V_w/V_t) (2) Alternative cushion gas (N_2/CO_2) replacement efficiency	Balances cushion gas cost savings against gas calorific dilution, operating costs, and cyclic turnover capacity

6.2 AHP–TOPSIS decision workflow

Raising the upper pressure limit increases storage capacity, but it raises the risk of fault slip and caprock failure exponentially. Rational and transparent trade-offs therefore require a formal multi-criteria decision pipeline. The pipeline used here combines AHP pairwise comparison [22], entropy weighting [23] and TOPSIS ranking [24], with a VIKOR compromise check [25]; previous screening studies based on AHP–FCE [26] or on hybrid AHP–entropy–TOPSIS workflows provide the methodological background. For completeness the steps of the pipeline are stated below in closed form.

Step 1: fusion of subjective and objective weights. A judgement matrix is first built by expert elicitation using AHP [22]. Pairwise comparisons on the 1–9 scale quantify geological experience and the principal eigenvector of the judgement matrix gives the subjective weight vector; logical contradictions are screened with the consistency ratio

$$CR = CI/RI < 0.1, \quad CI = (\lambda_{max} - n)/(n - 1) \tag{10}$$

where λ_{max} is the principal eigenvalue of the n × n judgement matrix and RI is the random-consistency index. To limit the bias of purely subjective judgement, entropy weighting [23] is used in parallel: for a decision matrix normalized to p_{ij} over m candidate sites, the entropy and the objective weight of indicator j are

$$e_j = -(1/\ln m)\Sigma_{i=1}^{m} p_{ij} \ln p_{ij}, \quad w_j^{o} = (1 - e_j)/\Sigma_k(1 - e_k) \tag{11}$$

and the two weight vectors are fused by a linear combination

$$w_j = \lambda w_j^s + (1 - \lambda) w_j^o, \quad 0 \le \lambda \le 1 \tag{12}$$

where λ expresses the trust placed in expert judgement; λ = 1 recovers the pure AHP weighting and λ = 0 the pure entropy weighting, so that the subjective–objective balance of the evaluation is documented explicitly instead of being hidden in an ad hoc choice of weights.

Step 2: ranking by distance to the ideal solution. Logs and three-dimensional seismic data for each candidate site would be fed into the fixed-stress HM model described above and simulated over several decades of cyclic operation. TOPSIS was introduced by Hwang and Yoon [24] and defines a positive ideal solution, which combines the best value of every indicator (lowest slip tendency, highest working gas volume and so on), and a negative ideal solution, which combines the worst. The weighted Euclidean distance of each site or pressure scenario to both solutions is computed, and the closeness coefficient ranks the alternatives

$$C_i = D_i^- / (D_i^+ + D_i^-), \quad 0 \le C_i \le 1 \tag{13}$$

where D_i^+ and D_i^- are the weighted distances of alternative i from the positive and negative ideal solutions. Because the ranking of near-tied candidates can be sensitive to the weights, a VIKOR compromise check [25] is run in parallel: for each alternative the group utility S_i, the individual regret R_i and the compromise index Q_i are evaluated as

$$Q_i = v(S_i - S^*)/(S^- - S^*) + (1 - v)(R_i - R^*)/(R^- - R^*) \tag{14}$$

where S^*, S^-, R^*, R^- are the best and worst values of S and R and $v \in [0, 1]$ weights the group utility against the individual regret. A candidate is accepted as robust only when it ranks first in both C_i and Q_i and satisfies the two VIKOR acceptance conditions; otherwise the pressure scenario is revised and returned to step (3) of Fig. 1, which closes the loop between decision and simulation.

This workflow, built on rigorous HM multi-physics and driven by the mathematics of AHP, entropy weighting and TOPSIS, is proposed as a structured alternative to the qualitative judgement and single-parameter veto that still dominate much of storage practice. Once the constitutive parameters and the weights are calibrated for a given basin, it is intended to bring the reproducibility of industrial reservoir-software workflows to the screening of very deep, very large traps; no such calibration is attempted here, so the workflow is a specified procedure awaiting application rather than established practice. Because every step is specified as an explicit formula, the pipeline is reproducible and auditable: identical inputs always produce an identical ranking, and the sensitivity of that ranking to the subjective weights can be quantified through the parameter λ of Eq. (12).

6.3 Scope and prospective verification

The theoretical framework is formulated as a research contribution rather than a literature review: every element is stated as a closed-form or numerically solvable

relation and the decision workflow is fully specified. Its physical core can be checked against published field observations. At the Hutubi UGS, geomechanical inversions report a slip tendency of 0.43 ± 0.04 at the onset of fault nucleation and magnitudes not exceeding 3.0 [18,19], while GPS observations show horizontal extension exceeding vertical uplift [20]; these values are proposed as prospective calibration targets against which the fault-stability and deformation indicators of Table 2 could later be checked. No simulation, calibration or ranking is performed in this paper, so the indicator set is not claimed to reproduce the observed critical state: the Hutubi values only indicate the order of magnitude that a calibrated model would have to match before it is applied to a greenfield candidate. A full multi-site application of the pipeline is the subject of ongoing work.

7 Conclusions

This study has analysed the HM dynamics driven by cyclic stress and multiphase flow, and has built a 4D evaluation system and decision workflow. The main conclusions are as follows.

(1) Rock fatigue damage and fluid hysteresis are the foundation of any credible deliverability forecast. Cyclic effective stress drives microscopic plastic deformation and crack reorganisation, so the Biot coefficient evolves non-linearly, while the cycle-dependent damage factor m(N) of Eq. (3) makes porosity and permeability decrease monotonically and irreversibly, towards the residual value set by the asymptotic damage factor.

(2) Long-term pore-pressure disturbance degrades the capillary breakthrough pressure of the caprock through fatigue, and poroelastic stress transfer—the elastic stress change transmitted from the deforming reservoir to faults located outside it—can raise the shear stress on those faults, pushing slip tendency past the critical friction limit and producing earthquake clusters such as those at Castor and Hutubi. Operation must therefore be capped at a fixed fraction of the minimum principal stress, and probabilistic fault slip potential monitoring corrected for cyclic loading must be part of the operating plan.

(3) Sequential 4D hydro-mechanical simulation based on the fixed-stress split, combined with a multi-criteria decision framework, is proposed here as a systematic and auditable route to site evaluation that complements, rather than supersedes, current practice. Converting model output into indicators of mechanical stability, seal integrity, flow assurance and cushion-gas economics, and feeding them into an AHP–entropy–TOPSIS pipeline with a VIKOR robustness check, is designed to close the loop from the governing physical equations to operations-research optimization.

References

1. Al-Shafi, M., Massarweh, O., Abushaikha, A. S., & Bicer, Y. (2023). A review on underground gas storage systems: Natural gas, hydrogen and carbon sequestration. *Energy Reports*, *9*, 6251-6266.

2. Zivar, D., Kumar, S., & Foroozesh, J. (2021). Underground hydrogen storage: A comprehensive review. *International Journal of Hydrogen Energy*, *46*(45), 23436-23462.
3. Oldenburg, C. M. (2003). Carbon dioxide as cushion gas for natural gas storage. *Energy & Fuels*, *17*(1), 240-246.
4. Sadeghi, S., & Sedaee, B. (2022). Mechanistic simulation of cushion gas and working gas mixing during underground natural gas storage. *Journal of Energy Storage*, *46*, 103885.
5. Zhou, J., Deng, G., Tian, S., Xian, X., Yang, K., Zhang, C., & Dong, Z. (2023). Experimental study on the permeability variation of sandstone at cyclic stress: Implication for underground gas storage. *Journal of Energy Storage*, *60*, 106677.
6. Mi, L., Guo, Y., Li, Y., Zeng, D., Lu, C., & Zhang, G. (2024). Evaluation of the dynamic sealing performance of cap rocks of underground gas storage under multi-cycle alternating loads. *Energy Geoscience*, *5*(4), 100319.
7. Zhu, S., Zheng, D., Sun, J., Wei, G., Wu, Z., Wang, J., ... & Li, C. (2022). Experimental study on dynamic sealing capacity and safe threshold of caprock in underground gas storages. *Journal of Natural Gas Science and Engineering*, *101*, 104521.
8. Sedaghatinasab, R., Kord, S., Moghadasi, J., & Soleymanzadeh, A. (2021). Relative permeability hysteresis and capillary trapping during CO2 EOR and sequestration. *International Journal of Greenhouse Gas Control*, *106*, 103262.
9. Land, C. S. (1968). Calculation of imbibition relative permeability for two-and three-phase flow from rock properties. *Society of Petroleum Engineers Journal*, *8*(02), 149-156.
10. Killough, J. E. (1976). Reservoir simulation with history-dependent saturation functions. *Society of Petroleum Engineers Journal*, *16*(01), 37-48.
11. Carlson, F. M. (1981, October). Simulation of relative permeability hysteresis to the nonwetting phase. In *SPE Annual Technical Conference and Exhibition* (pp. SPE-10157). SPE.
12. Blunt, M. J., Bijeljic, B., Dong, H., Gharbi, O., Iglauer, S., Mostaghimi, P., ... & Pentland, C. (2013). Pore-scale imaging and modelling. Advances in Water Resources, 51, 197–216.
13. Pruess, K., & Müller, N. (2009). Formation dry-out from CO2 injection into saline aquifers: 1. Effects of solids precipitation and their mitigation. *Water Resources Research*, 45(3), W03402.
14. Bashtani, F., & Mahani, H. (2026). Salt precipitation-induced formation damage near the wellbore: Implications for underground hydrogen storage. *Earth Energy Science*, *2*(2), 100056.
15. Segall, P., & Lu, S. (2015). Injection-induced seismicity: Poroelastic and earthquake nucleation effects. *Journal of Geophysical Research: Solid Earth*, *120*(7), 5082-5103.
16. Walsh III, F. R., & Zoback, M. D. (2016). Probabilistic assessment of potential fault slip related to injection-induced earthquakes: Application to north-central Oklahoma, USA. *Geology*, *44*(12), 991-994.
17. Cesca, S., Grigoli, F., Heimann, S., González, Á., Buforn, E., Maghsoudi, S., ... & Dahm, T. (2014). The 2013 September–October seismic sequence offshore Spain: a case of seismicity triggered by gas injection?. *Geophysical Journal International*, *198*(2), 941-953.
18. Zhang, B., Wang, B., Wei, B., Wu, Z., Wu, N. E., Lu, R., ... & Li, L. (2022). Spatiotemporal evolution of seismicity during the cyclic operation of the Hutubi underground gas storage, Xinjiang, China. *Scientific Reports*, *12*(1), 14427.
19. Liu, C., Zhang, F., Wang, Q., Wang, B., Zhang, Q., & Xu, B. (2023). Evaluation of fault stability and seismic potential for Hutubi underground gas storage due to seasonal injection and extraction. *Underground Space*, *13*, 74-85.
20. Jiang, G., Qiao, X., Wang, X., Lu, R., Liu, L., Yang, H., ... & Wong, T. F. (2020). GPS observed horizontal ground extension at the Hutubi (China) underground gas storage facility

and its application to geomechanical modeling for induced seismicity. *Earth and Planetary Science Letters*, *530*, 115943.

21. Kim, J., Tchelepi, H. A., & Juanes, R. (2011). Stability and convergence of sequential methods for coupled flow and geomechanics: Fixed-stress and fixed-strain splits. *Computer Methods in Applied Mechanics and Engineering*, *200*(13-16), 1591-1606.
22. Saaty, T.L. (1980). *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation.* McGraw-Hill, New York.
23. Shannon, C.E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379-423.
24. Hwang, C.L., Yoon, K. (1981). *Multiple Attribute Decision Making: Methods and Applications*. Springer, Berlin.
25. Opricovic, S., & Tzeng, G. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *European Journal of Operational Research*, 156(2), 445–455.
26. He, Y., Zhao, G., Tang, Y., Song, J., & Chen, Y. (2024). Available target evaluation of underground gas storage based on AHP-FCE methods. *Geoenergy Science and Engineering*, *242*, 213255.

Exploring Correlation of Body Weight with Depth of 3D Footwear Impression for Forensic Investigations

Anwar AMA Salem[1] and Loong Chuen Lee[2][0000-0002-3144-9877]

[1] Forensic Science Program, CODTIS, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
P148860@springer.com

[2] Forensic Science Program, CODTIS, Faculty of Health Sciences, Universiti Kabangsaan Malaysia, Bangi, Selangor, Malaysia
lc_lee@ukm.edu.my

Abstract. Three-dimensional (3D) footwear impressions are frequently encountered at outdoor crime scenes, particularly on soft substrates such as muddy soil in oil palm plantations. Although matching a footwear impression to a person of interest (POI) provides critical evidence, estimating the POI's biological traits can offer crucial investigative leads, especially when a comprehensive reference database is unavailable. This preliminary study explores the correlation between Body Mass Index (BMI) status and the depth profile of 3D footwear impressions using image processing tools. Seven adult Kuwaiti participants (three classified as obese and four as normal weight) voluntarily provided 3D impressions of both feet. A portable 3D scanner (Revopoint MIRACO Plus) was used to digitize the impressions, generating a total of 14 .STL files. The depth profile for each .STL file was estimated using an M-estimator, from which depth features capturing overall penetration depth and surface dispersion were extracted. The results demonstrate that subjects could be differentiated by their BMI status (normal vs. obese). Furthermore, inter-BMI variation was more pronounced in left footwear impressions than in their right counterparts. In conclusion, these preliminary findings suggest a potential correlation between BMI status and the numerical depth profile of 3D footwear impressions.

Keywords: Three-dimensional (3D) Footwear Impression, Body Weight Index (BMI), Forensic Sciences.

1 Introduction

Footwear impressions-whether three-dimensional (3D) or two-dimensional (2D)-are frequently deposited on both soft and hard surfaces at crime scenes. Despite their high prevalence and substantial evidential value, forensic footwear examination remains tedious and methodologically challenging. Standard protocols typically involve recovering impressions using techniques such as photography or casting, followed by a comparative analysis between the recovered mark and a reference impression seized from

a suspect (Kellett et al. 2024). However, in real-world investigations, a suspect is rarely identified during the initial stages.

Rather than relying solely on comparative analysis, investigators can extract a wealth of investigative intelligence directly from the impression itself (Kaur & Sodhi 2025). Class characteristics such as the make, model, and size of the footwear can be inferred, aiding in source attribution down to the individual shoe level. Alternatively, the spatial dimensions of an impression correspond to the wearer's foot size, while its depth profile potentially reflects the individual's weight (Svabova (nee Uhrova) et al. 2022).

Over the past few decades, significant research has focused on integrating statistical, image processing, and deep learning algorithms into forensic footwear analysis. For instance, Jang et al. (2025) demonstrated the utility of deep learning approaches for the quantitative assessment of pattern evidence, while Kim et al. (2026) deployed deep segmentation algorithms to automate comparative footwear analysis. Additionally, Hassan et al. (2021) conducted the pioneering study investigating the relationship between wearer age and footwear impression characteristics.

Building upon these advancements, this paper explores the correlation between Body Mass Index (BMI) and the depth profile extracted from 3D-scanned footwear impressions. Utilizing advanced image processing tools, various numerical features were extracted from 3D surface files (.STL) to model this relationship.

2 Methodology

2.1 Sample Collection

Seven Kuwaiti adults participated voluntarily in this study (Table 1). None of the participants presented with physical deformities or gait abnormalities that could potentially affect their movement. The Nike Air Max 270 Sneakers (rubber outsole) shoes of varying sizes were purchased from the local store in Kuwait. All the shoes were worn by volunteers for at least six months (ensuring 10K steps achieved in each week) before the subjects worn them for creating a number of 3D footwear impressions. All the subjects were instructed to walk normally at least eight steps before stepping on a shallow tray filled with fine sands.

Table 1. Details of the subjects.

Subject ID	Sex	BMI value	BMI status
S1	Male	28.56	Obese
S2	Male	29.43	Obese
S3	Male	33.88	Obese
S4	Male	22.34	Normal
S5	Female	22.86	Normal
S6	Female	23.73	Normal
S7	Female	24.13	Normal

2.2 3D Scanner

A portable 3D digital scanner with photogrammetric metrology capabilities (Revopoint MIRACO Plus) was utilized in this study to record the 3D footwear impressions formed on the sandy substrate. The device features a single-frame accuracy of up to 0.04 mm and a photogrammetric length accuracy expressed as 0.02 mm + 0.05 mm × L (m), where L represents the measured distance in meters.

2.3 Data Analysis

All image processing procedures and statistical analyses were performed in R (R Core Team 2025) using functions implemented in the following packages: 'tidyverse' (Wickham 2026), 'e1071' (Meyer et al. 2026), 'quantmod' (Ryan et al. 2026), 'Rvcg' (Schlager et al. 2026), 'rgl' (Murdoch et al. 2026), and 'MASS' (Ripley et al. 2026).

For each 3D mesh (.STL file), the 3D vertex coordinates (X, Y, Z) were extracted. To define a reference surface, a robust plane-fitting technique using M-estimation (via the MASS package) was applied to model the dominant flat ground plane. The plane's normal vector was derived based on Equation (1):

$$Z = AX + BY + C \tag{1}$$

where A, B, C denote the coefficient of X, Y and intercept vectors. Subsequently, the perpendicular distance (d) from each point to the fitted plane was calculated using Equation (2).

$$Distance\ (d) = \frac{|AX+BY-Z+C|}{\sqrt{(A^2+B^2+1)}} \tag{2}$$

To isolate points strictly within the impression cavity, vertices with observed Z-values exceeding the predicted baseline Z-surface were removed, retaining only negative displacement points (i.e., those penetrating below the ground surface). From the isolated cavity points, ten key depth features were extracted for each impression (Table 2).

Table 2. Numerical features describing the depth of the footwear impression.

Feature	ID	Descriptions
Maximum depth value	1	Denoting peak penetration
Mean depth value	2	Summarizing overall cavity depth
Median depth value	3	A robust measure of central depth
Range of depth	4	Representing total vertical span
Interquartile range (IQR) of depth	5	Reflecting the variability of the middle 50% of depth data
Standard deviation of depth value	6	Degree of variation of individual depth points around the mean depth
Skewness of depth value	7	A measure of asymmetry in the depth distribution

Kurtosis of depth value	8	Highlighting the presence of extreme depth outliers or sharp depth transitions.
Topographic tread depth	9	Revealing the structural height difference between tread elements and grooves.
Volumetric displacement	10	A proxy metric for applied force, weight and distance distribution.

For illustrative purposes, a depth waveform profile was generated for each .STL file to visualize the longitudinal depth distribution from heel to toe. To standardize variations in footwear length across subjects, the spatial profile was partitioned along the Y-axis into 20 equal intervals (bins), with values normalized from 0 (heel) to 100 (toe). The resulting one-dimensional (1D) cross-sectional depth waveform represents the continuous longitudinal transition across the impression. This visualization effectively highlights key biomechanical features, including structural tread transitions, arch gaps, and localized pressure distribution patterns along the length of the footwear.

3 Results and Discussions

3.1 Univariate Summary Statistics

The discriminatory ability of the numerical features was first explored using univariate boxplots. For brevity, only features showing satisfactory separation between the two BMI groups are presented in Fig. 1. Topographic tread depth and volumetric displacement were excluded due to substantial overlap between the two groups. Initial observations indicate that BMI-related differences were more pronounced in left footwear impressions than in right impressions, particularly across mean, maximum, median and range values. Conversely, distribution metrics such as kurtosis and skewness yielded less distinct separation between the groups. Regarding intra-group variability, right impressions from obese subjects exhibited lower dispersion compared to those from the normal-weight group. In contrast, right impressions from normal-weight group demonstrated relatively high intra-group variation, most notably in skewness, kurtosis, range and maximum depth metrics.

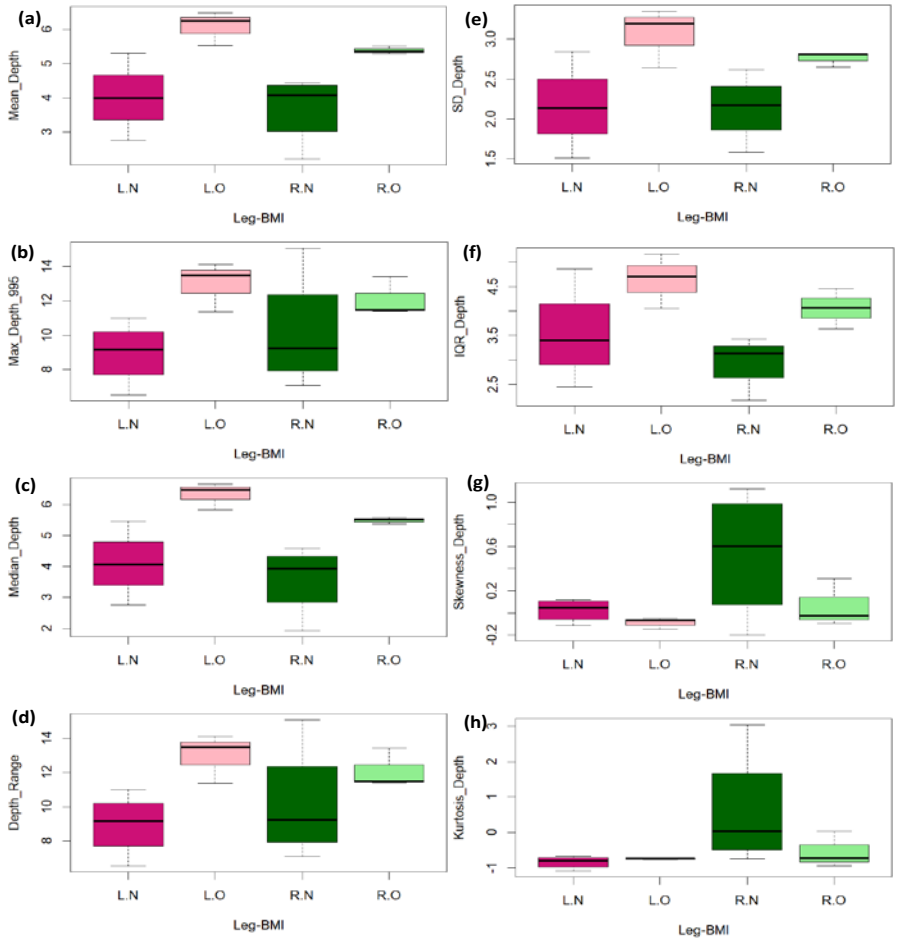


Fig. 1. Comparative boxplots showing the distributions of 3D footwear impression parameters across the four study groups: left leg of normal-weight subjects (L.N), left leg of obese subjects (L.O), right leg of normal-weight subjects (R.N), and right leg of obese subjects (R.O). Distributions are displayed for (a) mean depth, (b) maximum depth, (c) median depth, (d) range of depth, (e) SD of depth, (f) IQR of depth, (g) skewness of depth, and (h) kurtosis of depth.

3.2 1D Depth Waveform Profiles

Consistent with previous findings indicating that inter-BMI variation is more pronounced in the subjects' left legs, 1D waveform depth profiles of the six selected subjects were constructed to further elucidate the biomechanical and morphological differences (Fig. 2). Specifically, normal-weight subjects tended to exhibit shallower overall depth, with maximum depth values ranging around 7 to 12 units, whereas obese subjects demonstrated markedly deeper impressions ranging from approximately 12 to 15 units. Furthermore, peak pressure in normal-weight individuals was primarily concentrated across the rearfoot-midfoot region, whereas obese subjects displayed peak

pressure concentration predominantly at the forefoot. Additionally, normal-weight subjects presented moderate offset between mean and maximum depth, while obese subjects exhibited a substantially wider divergence. Finally, normal-weight subjects demonstrated gradual depth transitions from heel-strike to toe-off, whereas obese subjects were characterized by steeper depth gradients across transition zones.

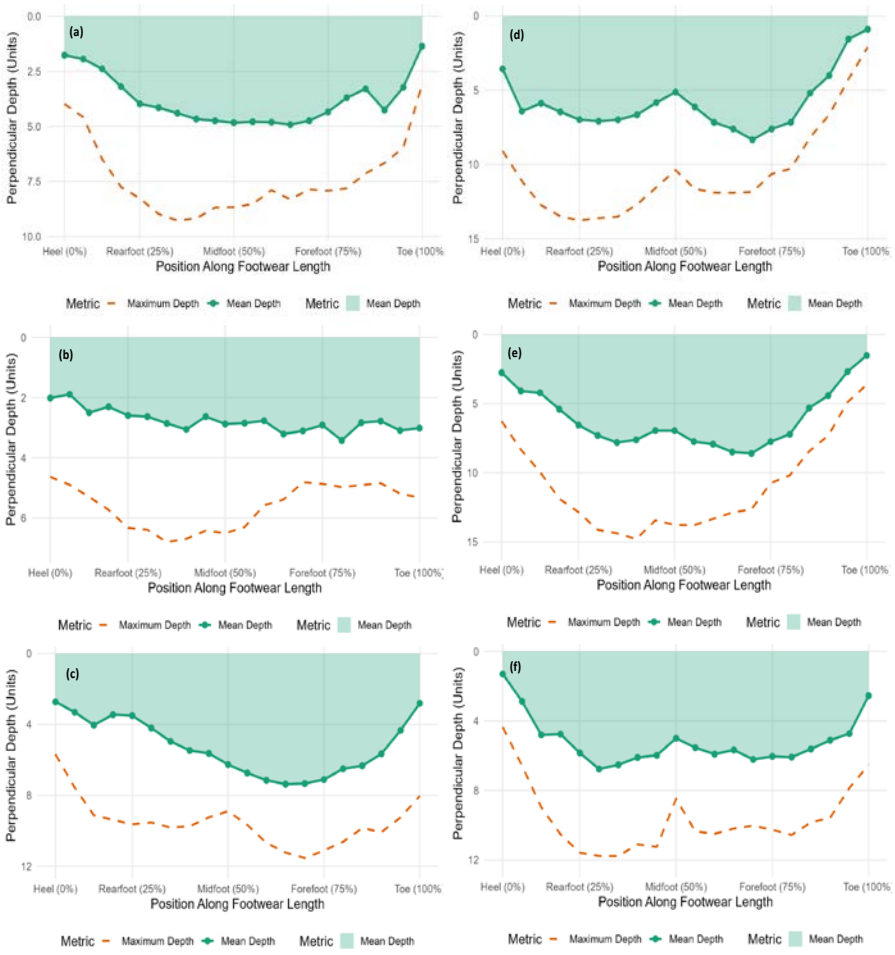


Fig. 2. Waveform plots illustrating the 1D cross-sectional depth profile of normal: (a) S4, (b) S5, (c) S6, and obese subjects: (d) S1, (e) S2 and (f) S3. All are of left-leg.

4 Conclusion

This study demonstrates the utility of integrating 3D surface scanning with quantitative image processing tools to extract biomechanical intelligence directly from forensic footwear impressions. By evaluating 3D vertex coordinates from .STL mesh models

using robust M-estimation plane fitting, ten distinct depth parameters and 1D longitudinal depth waveforms were successfully derived. The empirical findings reveal satisfactory degree of morphological and biomechanical differences between normal-weight and obese individuals. Specifically, left footwear impressions and 1D depth waveforms exhibited high sensitivity to weight-induced variation, with obese subjects characterized by slightly higher impression depths. Despite that, this work acknowledged that inter-BMI depth variation may also be influenced by sex, foot size, gait biomechanics, and plantar pressure distribution. Future studies involving a larger cohort are warranted to validate the impact of BMI on 3D footwear impression depth profiles, while isolating these effects from confounding factors such as sex, foot size and gait.

Acknowledgement

Anwar AMA Salem gratefully acknowledges the financial support from the Government of Kuwait through a PhD scholarship.

References

1. Hassan, M., Wang, Y., Wang, D., Li, D., Liang, Y., Zhou, Y., Xu, D.: Deep learning analysis and age prediction from shoeprints. Forensic Science International 327, 110987 (2021).
2. Jang, M., Carriquiry, A., Park, S.: Enhancing forensic shoeprint analysis: Application of the Shoe-MS algorithm to challenging evidence. Science & Justice 65(4), 101255 (2025).
3. Kaur, J., Sodhi, G.S.: Forensic Importance of footwear impression evidence: A review. Journal of Indian Academy of Forensic Medicine 46(3), 425-430 (2025).
4. Kellett, D., Zolghadriha, S., Morgan, R., Lagnado, D., Nakhaeizadeh, S.: Forensic footwear examination: A systematic review of the existing literature. Forensic Science International 365, 112295 (2024).
5. Kim, Y., Carriquiry, A., Park, S.: Enhancing automated shoeprint comparison via synthetic data generation and deep segmentation. Science & Justice 66, 101434 (2026).
6. R Core Team: A Language and Environment for Statistical Computing. R version 4.5.0 (2025-04-11), R Foundation for Statistical Computing, Vienna, Austria (2025).
7. Svabova (nee Uhrova), P., Caplova, Z., Benus, R., Chovancova (nee Kondekova), M., Masnicova, S.: Estimation of stature and body weight from static and dynamic footprints-Forensic implications and validity of non-colouring cream method. Forensic Science International 330, 111105 (2022).
8. Wickham, H.: R package 'tidyverse', v.2.0.0 (2023). https://cran.r-project.org/web/packages/tidyverse/tidyverse.pdf
9. Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C-C., Lin, C-C.: R package 'e1071', v.1.7-17 (2025). https://cran.r-project.org/web/packages/e1071/e1071.pdf
10. Ryan, J.A, Ulrich, J.M., Smith, E.B., Thielen, W., Teetor, P., Bronder, S.: R package 'quantmod', v. 0.4.29 (2026). https://cran.r-project.org/web/packages/quantmod/quantmod.pdf
11. Schlager, S., Francois, G., Schaefer, T., Wang, Z.: R package 'Rvcg', v. 0.25 (2025). https://cran.r-project.org/web/packages/Rvcg/Rvcg.pdf

-

12. Murdoch, D., Adler, D., Nenadic, O., Urbanek, S., Chen, M., Gebhardt, A., Bolker, B., Csardi, G., Strzelecki, A., Senger, A.: R package ‘rgl’, v. 1.3.36 (2026). https://cran.r-project.org/web/packages/rgl/rgl.pdf
13. Ripley, B., Venables, B., Bates, D.M., Hornik, K., Gebhardt, A., Firth, D.: R package ‘MASS’, v. 7.3-66 (2026). https://cran.r-project.org/web/packages/MASS/MASS.pdf

Synthesis and Characterization of Fe_3O_4/rGO/Ag Nanocomposite in Extracting Opioids

Enas Amoori Hadi Hadi[1], Wijdan Shakir Khayoon[2] and Loong Chuen Lee[3[0000-0002-3144-9877]]

[1] Forensic Science Program, CODTIS, Faculty of Health Sciences, Universiti Kabangsaan Malaysia, Bangi, Selangor, Malaysia
P127374@siswa.ukm.edu.my
[2] Department of Chemistry, College of Sciences, University of Baghdad, Iraq
wijdan.khayoon@sc.uobaghdad.edu.iq
[3] Forensic Science Program, CODTIS, Faculty of Health Sciences, Universiti Kabangsaan Malaysia, Bangi, Selangor, Malaysia
lc_lee@ukm.edu.my

Abstract. A novel ternary Fe_3O_4/rGO/Ag (magnetite/reduced graphene oxide/silver) nanocomposite was synthesized via sonication-assisted physical processes driven by non-covalent interactions. The structural, morphological, and magnetic properties of the synthesized nanocomposite were comprehensively characterized using Field Emission Scanning Electron Microscopy (FE-SEM) and Energy-Dispersive X-ray Spectroscopy (EDX) mapping, Fourier-transform infrared spectroscopy (FTIR) and Vibrating Sample Magnetometry (VSM). FE-SEM micrographs revealed the typical morphology of the nanocomposite: spherical magnetite nanoparticles (50-100 nm) forming dense clusters alongside individual quasi-spherical Ag nanoparticles (15-25 nm) dispersed across the basal planes of the rGO sheets. EDX elemental mapping confirmed that iron (48.6 wt%), carbon (21.8 wt%) and oxygen (26.4 wt%) were the primary elements of the nanocomposite. The Fe_3O_4/rGO/Ag nanocomposite exhibited a saturation magnetization of approximately 55.19 emu/g and displayed an S-shape hysteresis loop, confirming its superparamagnetic behavior. Additionally, the IR spectrum verified the presence of rGO via characteristic bands at 3440 and 1034 cm^{-1}. Overall, the magnetic nanocomposites were successfully synthesized as a nano-adsorbent for magnetic solid-phase extraction (MSPME). Its potential for extracting opioids from urine samples will be evaluated in future work.

Keywords: Fe3O4 nanoparticles, reduced graphene oxide (rGO), silver nanoparticles, magnetic separation.

1 Introduction

Opioids are among the most widely abused illicit drugs worldwide (Cragg et al. 2019). According to the World Health Organization, nearly 80% of drug-related deaths are attributed to opioids. Opioid overdose can lead to severe psychological and physical

dependence, causing adverse reactions such as nausea, vomiting, sweating, and diarrhea (Leung et al. 2022).

Consequently, detecting trace amounts of opioids in complex biological matrices, such as urine, is essential for monitoring usage and identifying abuse trends (Razlansari et al. 2022). Although analytical instruments like liquid chromatography offer high sensitivity, their reliability heavily depends on sample preparation, which critically impacts selectivity, recovery and detection limits. Thorough sample preparation is imperative when detecting opioids in biological fluids due to their low concentrations and the presence of interferences-such as proteins and lipids-that can obscure drug detection (Al, Kul & Sagirli 2026).

Solid phase extraction (SPE) is a well-established sample preparation technique widely applied in the trace-level analysis of drugs in complex matrices (Cliville-Cabre et al. 2025). However, conventional SPE workflows suffer from several major drawbacks: they require complex, multi-step procedures (activation, loading, washing and elution) that are both time-consuming and solvent-intensive. Advances in nanotechnology have enabled the development of magnetic nano adsorbents for magnetic SPE (MSPE) (Vasconcelos & Fernandes 2017). Compared to traditional sorbents, magnetic nano adsorbents facilitate rapid liquid-solid separation under an external magnetic field, eliminating tedious centrifugation and filtration steps to significantly increase sample throughput.

Typically, magnetic nano adsorbents consist of two main components: a magnetic core-most commonly magnetite (Fe3O4)-and a functional shell designed to stabilize and protect the core (Deivayanai et al. 2025). Surface functionalization mitigates the inherent limitations of pure magnetite while enhancing adsorbent performance (Manyangadze et al. 2020). Tailored surface modifications improve both the sensitivity and selectivity of the detection method. For example, Soltani et al. (2018) synthesized magnetic carbon nanotube (MCNT) nano composites to preconcentrate morphine and codeine from human urine and blood serum, demonstrating enhanced operational simplicity and convenience.

In the present study, a novel ternary Fe_3O_4/rGO/Ag (magnetite/reduced graphene oxide/silver) nanocomposite was synthesized by adapting established methodologies. The morphological and structural properties of the nano adsorbents were carefully determined. Its potential as adsorbent for the simultaneous preconcentration of multiple opioids (including morphine, codeine, 6-MAM, oxycodone, hydromorphone, thebaine) from urine samples will be evaluated in a future study.

2 Methodology

2.1 Reagents and Apparatuses

The consumables used in this study are detailed in Table 1. Ultrasonic treatment was performed using a stainless-steel digital ultrasonic washer cleaning machine (LK-D31S, 5L, 200W, 110V/220V, 50/60Hz, Zhengzhou Linker Medical Equipment Co., Ltd, China). Magnetic stirring was carried out using and SH-2 magnetic stirrer (HYCC, China).

Table 1. Details of consumables.

Chemicals	Details
Iron (II, III) oxide nano powder	50-100 nm particle size, 97% trace metal basis, 637106-25G, Sigma-Aldrich (USA)
Reduced graphene oxide	777684-250MG, Sigma-Aldrich (Spain)
Silver nano powder	99.9% purity (APS: 20nm), 7440-22-4, Nanoshel (Punjab, India)

2.2 Nano Adsorbent Synthesis

The synthesis of the Fe_3O_4/rGO/Ag nanocomposite was adapted from previously reported methods (Abdolmohammad-Zadeh et al. 2018, Soltani et al. 2018) with minor modifications. The ternary nano composite was prepared via sonication-assisted physical assembly driven by non-covalent interactions.

First, 100 mg of reduced graphene oxide (rGO) was dispersed in 100 mL of HPLC-grade deionized water and sonicated for 30 min until a stable, homogeneous black suspension was obtained (Soltai et al. 2018). Subsequently, 200 mg of Fe3O4 nanoparticles were added gradually to the rGO suspension under continuous stirring at room temperature for 1 h. The mixture was further sonicated for 20 min to facilitate the formation of the binary Fe3O4/rGO magnetic composite (Taghvimi et al. 2017). The resulting composite was collected using an external NdFeB magnet, washed twice with deionized water and once with ethanol to eliminate residual impurities and unreacted materials, and dried in an oven at 60 oC for 8 h (Taghvimi et al. 2017).

To synthesized the ternary composite, 100 mg of the dried Fe3O4/rGO material was redispersed in 50 mL of deionized water and sonicated for 15 min. Thereafter, 28 mg of silver nanoparticles (AgNPs) was gradually added under continuous stirring. The synthesized Fe3O4/rGO/Ag nanocomposites were isolated with an external magnet, washed thoroughly with deionized water followed by ethanol, and dried at 50-60 oC for 6-8 h to yield the purified final product (Abdolmohammad-Zadeh et al. 2019).

2.3 Instrumentation

Surface morphology and elemental distribution of the nanocomposites were evaluated using a field-emission scanning electron microscope equipped with energy-dispersive X-ray spectroscopy (FESEM-EDS, Carl Zeiss, Supra 55VP/Merlin Compact). Surface functional groups were identified via Fourier transform infrared spectroscopy (FTIR) in the range of 4000 to 650 cm-1 using a PerkinElmer Spectrum 400 FTIR spectrometer equipped with a universal attenuated total reflectance (UATR) accessory. Magnetic properties and hysteresis behavior were determined using a vibrating sample magnetometry (VSM, Lakeshore 7404 Series) under magnetic field of ± 10000 Oe at room temperature.

3 Results

3.1 Magnetic profile

As depicted in Fig. 1, the magnetic properties of the synthesized Fe_3O_4/rGO/Ag nanocomposite were evaluated using VSM. The nanocomposite exhibited a saturation magnetization (Ms) of approximately 55.19 emu/g, confirming the successful incorporation and high loading density of magnetic iron oxide nanoparticles anchored onto the graphene matrix. The narrow hysteresis loop accompanied by a small coercivity (Hc=93.5 G) indicates that the anchored magnetite nanoparticles are small in size (typically around 10-20 nm). Although the coercivity (Hc) and remanent magnetization (Mr) are non-zero, their near-zero values demonstrate soft ferromagnetic behavior bordering on superparamagnetism. Overall, this magnetic profile confirms the preserved magnetic integrity of the Fe3O4 nanoparticles within the composite structure, ensuring rapid magnetic separation under an external magnetic field.

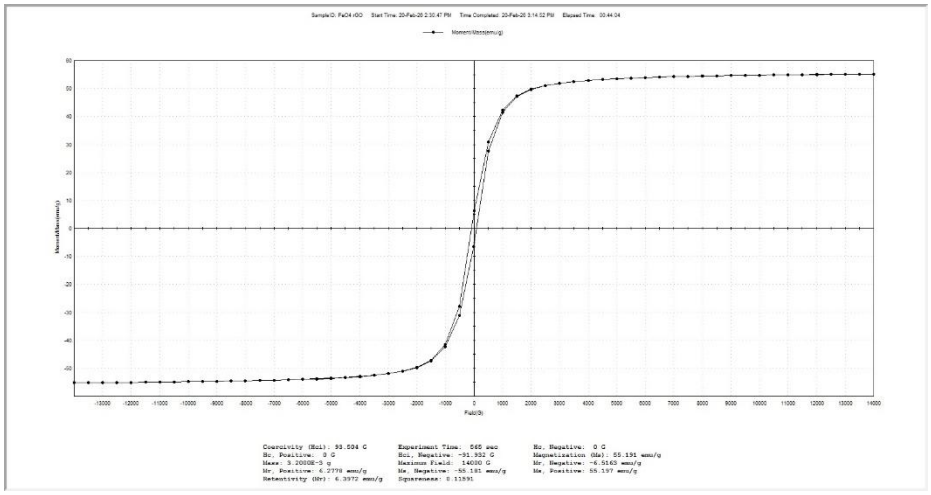


Fig. 1. The magnetic property of the Fe_3O_4/rGO/Ag nanocomposite by vibrating sample magnetometer.

3.2 Morphological and structural characterization

The surface morphology of the Fe3O4/rGO/Ag nanocomposite was characterized by FESEM (Fig. 2). The micrograph clearly displays three distinct structural components corresponding to the thin reduced graphene oxide (rGO) sheets, polyhedral Fe3O4 magnetite, and spherical/aggregated Ag nanoparticles (AgNPs). In the background and underlying regions, characteristic thin, wrinkled, and folded sheet-like structures resembling rGO are evident. The well-defined, faceted geometric structures exhibiting sharp edges and flat crystal faces correspond to magnetite nanoparticles. Meanwhile, granular spherical aggregates attached to the surface and edges of the composite represent AgNP

clusters. The presence and distribution of AgNPs across the nanocomposite surface re further corroborated by the EDX analysis (Table 2), which confirms an average silver content of 1.3±0.3wt%. The relatively low Ag content detected by EDX, compared with the initial amount of Ag nanoparticles added during synthesis (28 mg), may be attributed to the localized nature of EDX analysis. EDX provides elemental information from a limited analyzed area and therefore may not fully represent the overall bulk composition of the synthesized nanocomposite. In addition, a non-uniform distribution of Ag nanoparticles across the Fe_3O_4/rGO surface may have contributed to the observed Ag percentage.

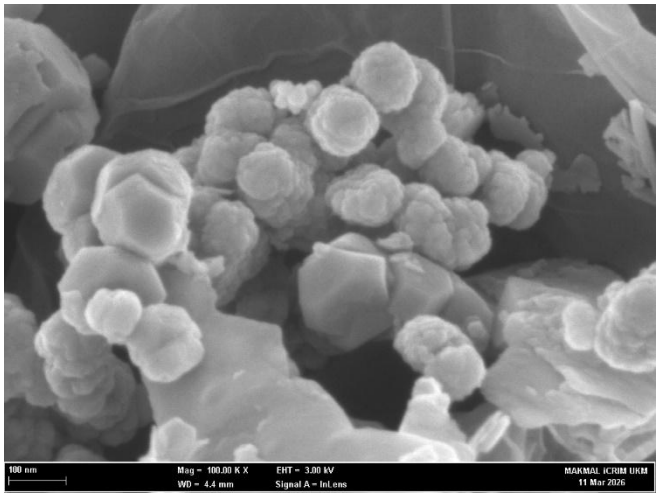


Fig. 2. A high-resolution FESEM micrograph illustrating the morphology of the Fe_3O_4/rGO/Ag nanocomposite at magnification 100 000×.

Table 2. Composition of Fe_3O_4/rGO/Ag nanocomposite based on EDX analysis

Element	Wt% (standard deviation)
Fe	48.6 (0.6)
O	26.4 (0.5)
C	21.8 (0.6)
Ag	1.3 (0.3)
Mn	0.9 (0.2)
Mg	0.7 (0.1)
Si	0.2 (0.1)

3.3 Surface Functional Groups

The FTIR spectrum of the Fe_3O_4/rGO/Ag nanocomposites is shown in Fig. 3, with key characteristic absorption bands summarized in Table 3. Major absorption bands

corresponding to functional groups on the rGO sheets are clearly observed. However, characteristic vibrations associated with the metal oxides and lattice modes-specifically the Fe-O stretching of magnetite and metal-based vibrations of Ag-are absent due to the spectral cutoff at 650 cm-1. Inorganic metal-oxygen and metal-surface lattice vibrations typically appear below 650 cm-1. For example, the stretching vibration of Ag generally manifests as a sharp peak in the range of 620-585 cm-1 (Kayed, Issa & Al-ourabi 2024), whereas the characteristics Fe-O stretching of Fe3O4 typically occurs near 578 cm-1.

Table 3. Characteristic IR band defined by functional groups

Vibration mode	Wavenumber (cm^{-1})	Details (Component)
-OH deformation	3400-3450 (broad), 1390	Residual hydroxyl groups or adsorbed moisture (rGO)
-C=O stretching	1720	Residual carbonyl or carboxylic groups (rGO)
C=C stretching	1640-640	Skeletal vibrations of unoxidized sp2 carbon domains (rGO)
-C-O stretching	1220-1002	Epoxy, alkoxy, or alcoholic C-O bonds that diminished during reduction (rGO)
Ag stretching	620-585	Silver nanoparticles (AgNPs)
Fe-O stretching (Octahedral site)	620-585	Fe3O4 NPs
Fe-O stretching (Tetrahedral site)	600-560	Fe3O4 NPs

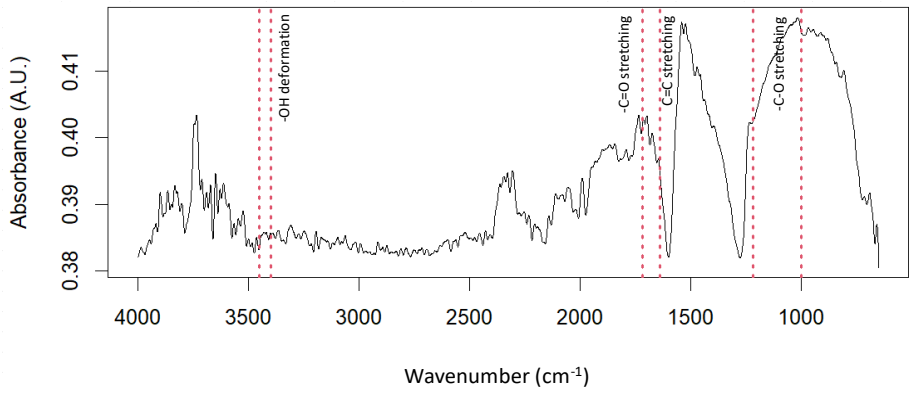


Fig. 3. Infrared spectrum showing the functional group present in the Fe_3O_4/rGO/Ag nanocomposites.

4 Conclusion

Overall, the ternary Fe3O4/rGO/Ag nanocomposite was successfully synthesized and characterized as an effective nano adsorbent for magnetic solid-phase microextraction extraction (MSPME). Its performance and applicability for the simultaneous extraction and preconcentration of trace opioids from urine samples will be systematically evaluated in future work.

Acknowledgement

Enas Amoori Hadi Hadi gratefully acknowledges the financial support provided by the Government of Iraq through a PhD scholarship.

References

1. Abdolmohammad-Zadeh, H., Zamani, A., Shamsi, Z.: Preconcentration of morphine and codeine using a magnetite/reduced graphene oxie/silver nano-composite and their determination by high-performance liquid chromatography. Journal of Chromatography A 1590, 2-9 (2019).
2. Al, S., Kul, A., Sagirli, O.: LC-MS/MS methods for synthetic opioids in urine: analytical and clinical evaluation. Clinica Chimica Acta 588, 120991 (2026).
3. Cliville-Cabre, P., Rosendo, L.M., Borrull, F., Aguilar, C., Calull, M.: A comparative study of SPE- and DLLME-based methods for the determination of opioids and benzodiazepines in urine samples using LC-MS/MS. Microchemical Journal 208, 112354 (2025).
4. Cragg, A., Hau, J.P., Woo, S.A., Kitchen, S.A., Liu, C., Doyle-Waters, M.M., Hohl, C.M.: Risk factors of misuse of prescribed opioids: A systematic review and meta-analysis. Annals of Emergency Medicine 74(5), 634-646 (2019).
5. Deivayanai, V.C., Thamarai, P., Karishma, S., Saravanan, A., Vickram, A.S., Yaashikaa, P.R., Sonali, S.: A comprehensive review on impregnated magnetic nanoparticles in advanced wastewater treatment: An in-depth technical review and future directions. Sustainable Chemistry for the Environment 9, 100220 (2025).
6. Kayed, K., Issa, M., Al-ourabi, H. The FTIR spectra of Ag/Ag2O composites doped with silver nanoparticles. Journal of Experimental Nanoscience 19(1), 2336227 (2024).
7. Leung, J., Santo Jr., T., Colledge-Frisby, S., Mekonen, T., Thomson, K., Degenhardt, L., Connor, J.P., Hall, W., Stjepanovic, D.: Mood and anxiety symptoms in persons taking prescription opioids: A systematic review with meta-analyses of longitudinal studies. Pain Medicine 23(8), 1442-1456 (2022).
8. Manyangadze, M., Chikuruwo, N.H.M., Narsaiah, T.B., Chakra, C.S., Radhakumari, M., Danha, G.: Enhancing adsorption capacity of nano-adsorbents via surface modification: A review. South African Journal of Chemical Engineering 31, 25-31 (2020).
9. Razlansari, M., Ulucan-Karnak, F., Kahrizi, M., Mirinejad, S., Sargazi, S., Mishra, S., Rahdar A., Diez-Pascual, A.M.: Nanobiosensors for detection of opioids: A review of latest advancements. European Journal of Pharmaceutics and Biopharmaceutics. 179, 79-94 (2022).
10. Soltani, M.D., Taher, M.A., Behzadi, M., Fazelirad, H.: Synthesis, characterization and application of magnetic carbon nanotubes for the simultaneous extraction and high

performance liquid chromatographic determination of codeine and morphine in human urine, blood serum, opium and tablet samples. Sensors and Actuators A: Physical 280, 31-37 (2018).
11. Taghvimi, A., Hamishehkar, H.: Carbon coated magnetic nanoparticles as a novel magnetic solid phase extraction adsorbent for simultaneous extraction of methamphetamine and ephedrine from urine samples. Journal of Chromatography B 1041-1042, 113-119 (2017).
12. Vasconcelos, I., Fernandes, C.: Magnetic solid phase extraction for determination of drugs in biological matrices. TrAC Trends in Analytical Chemistry 89, 41-52 (2017).

Application of Image Processing in the Discrimination of Red Fabrics for Forensic Investigations

Wan Adam Harris Wan Azmirrudin[1] and Loong Chuen Lee[2[0000-0002-3144-9877]]

[1] Forensic Science Program, CODTIS, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
a202473@siswa.ukm.edu.my
[2] Forensic Science Program, CODTIS, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
lc_lee@ukm.edu.my

Abstract. Fabric composed of microscopic fibers can serve as valuable corroborative evidence in forensic investigation by establishing a link between a suspect, a victim and a crime scene. Consequently, the reliability of fiber color matching methods is critical. This study aimed to evaluate image processing tools for discriminating between red cotton and red polyester fabrics. Two XL-sized t-shirts-one cotton and one polyester-were purchased from an online store and cut into 35 equal-sized swatches (6cm×4 cm). A handheld Dino-Lite digital microscope was used to capture four images per swatch (i.e., two from the front side and two from the rear side), yielding 140 images in .bmp format for each fabric type. The images were preprocessed by cropping the edges, applying a Median Blur filter for denoising, and removing glare. Principal Component Analysis (PCA) score plots were then used to evaluate sample clustering by fabric type. The results demonstrated that red cotton-based swatches were clearly discriminated from red polyester-based swatches. Furthermore, the front and rear sides of the cotton-based swatches were clustered in two distinct groups. In conclusion, image processing tools are effective for differentiating red fabrics based on their material origin. These insights highlight the potential of digital image analysis coupled with image processing tools as a non-destructive and reliable approach to enhance fiber evidence evaluation in forensic casework.

Keywords: Red Fabric, Color Discrimination, Forensic Investigation.

1 Introduction

Textile fibers can be produced from natural, regenerated or synthetic sources and are available in various forms, including clothing and personal belongings. Consequently, they are commonly encountered at virtually all types of crime scenes, such as rape, kidnapping and homicide (Goodpaster & Liszewski 2009). As one of the most prevalent types of microtraces, fibers contribute significantly to crime scene reconstruction due to their widespread distribution in the environment and relatively high transfer and

persistence rates (Lim et al. 2022). Thus, they serve as valuable associative or exclusionary evidence in elucidating the past actions of a perpetrator.

By showing that a questioned fiber (typically recovered from the victim) shares the same source as a known fiber (from the suspect), one can postulate that the victim and suspect were once in close physical contact (Powell et al. 2021). The source of a questioned fiber is determined by examining its physical and chemical characteristics. However, precise color determination remains challenging due to the vast diversity of fibers manufactured worldwide. While a wide array of non-destructive optical and spectral techniques area available for fiber color analysis, many of these methods are costly and instrumentally demanding despite their merits (Nayak et al. 2020).

Microscopic color evaluation is the most rapid and accessible approach (Hu et al. 2020). Nevertheless, visual microscopic inspection remains highly susceptible to cognitive bias and subjective interpretation (Kunkler & Roy 2023). To mitigate subjectivity, image processing and computer vision tools have been integrated to analyze fiber colors from digital images. Numerical feature describing color can be extracted from these images, enabling objective assessment with enhanced efficiency compared to manual inspection (Muratbekova et al. 2026). Although image processing tools have been widely applied to digital image color evaluation across various domains-particularly food quality assessment-their potential in forensic fiber examination remains largely unexplored. Therefore, the present work aims to evaluate the color of polyester and cotton fibers using digital image analysis techniques.

2 Methodology

2.1 Sample

Two red T-shirts were purchased from an online retailer and subsequently cut into 35 equal-sized swatches (6cm×4cm). Further details are presented in Table 1.

Table 1. Details of the studied samples

Fiber Types	Side (ID)	Details
Cotton	Outer (C.F)	XL size (chest: 54.10cm, length: 71.12cm)
	Inner (C.R)	
Polyester	Inner (P.F)	L size (chest: 55.88cm, length: 74.93cm)
	Outer (P.R)	

2.2 Data Collection

Images of the swatches were acquired using a handheld Dino-Lite digital microscope (USB 2.0, AMA4515-ZT-EDGE). Four images were captured per swatch-two from the inner side and two from the inner side-yielding a total of 140 images in .BMP format for each fiber type. All samples were imaged under uniform lighting exposure, at a fixed working distance, and at a 90-degree angle. The samples were captured in color

using the manufacturer's recommended acquisition procedures, relying solely on the microscope's built-in illumination without supplementary light sources.

2.3 Statistical Analysis

All the statistical analyses were performed using the R software (R Development Core R Team 2025).

The raw BMP images were carefully preprocessed prior to interpretation to convert them into standardized, clean pixel data. This preprocessing step aimed to eliminate subtle shading, specular highlights, uneven illumination, and edge fraying, which could otherwise obscure genuine dye variations. First, the outer boundaries of each swatch were cropped by 10% on all sides, Spatial median filtering was then applied for noise reduction; a 3×3 sliding window traversed each pixel, replacing its intensity value with the median value of its eight neighboring pixels. This filter effectively isolated hot pixels while preserving the microstructure and sharp yarn boundaries of the fabric weave. Next, the images were temporarily converted into single-channel intensity (grayscale) maps. Pixels exceeding a brightness threshold of 95% maximum luminance ($L > 0.95$) were identified as glare and zeroed out across all RGB channels. A binary mask M (x, y) was created according to Equation (1). Then, each color channel (R, G, B) is multiplied element-wise by M.

$$M(x,y) = \begin{cases} 1 \text{ if Grayscale}(x,y) < 0.95 \\ 0 \text{ if Grayscale}(x,y) \leq 0.95 \end{cases} \quad (1)$$

The preprocessed images were subsequently transformed into the CIE L^*a^*b color space, where L^* represents lightness, a^* denotes the red-green chromatic axis, and b^* denotes the yellow-blue chromatic axis. From this color space, ten distinct statistical moments were extracted for analysis (Table 2).

Table 2. Descriptions of ten statistical moments extracted from the digital images

Properties	Feature (ID)	Details
Core shade & tone discrimination	Mean red-green value (V1)	Measure the average color shift along the green-to-red axis. More vivid red color denoting by the higher value.
	Mean yellow-blue value (V2)	Positive value denoting warm red; negative value denoting cool red.
	Mean lightness (V3)	Measures the overall brightness or intensity. For identifying bright red (high value) versus dark red (low value).
	Mean color saturation (V4)	Measure overall color saturation/purity. Quantifying how vivid or neutral the color is, independent of lightness.

Pattern & texture	Standard deviation of lightness (V5)	Measures the contrast or variation in brightness. High values denote uneven surface texture; low values indicate smooth uniform surface.
	Standard deviation of red-green value (V6)	Measure the variance in green-red intensity across pixels. High value indicates patchy redness rather than a uniform tone.
	Skewness of red-green value (V7)	Measure the asymmetry in the distribution of red-green pixel values. Positive skew denotes toward red; negative skew indicates toward green.
Defect detection	Percentage of high-red pixels (V8)	Percentage of pixels in upper red quartiles.
	Mean red-to-green ratio (V9)	Red-to-green ratio. Captures subtle changes in yellowing or green undertones independently of total lighting
	Mean red-to-blue ratio (V10)	Red-to-blue ratio. Captures subtle changes in yellowing or green undertones independently of total lighting

The numerical features describing the color of the swatches were subsequently evaluated using Principal Component Analysis (PCA). PCA is a robust multivariate exploratory method that constructs a reduced number of uncorrelated variable, or principal components, from the original input variables (Lee & Jemain 2021). Sample clustering based on fiber type and swatch side (inner vs. outer) was inspected using score plots, while the relative contribution of the numerical features was assessed through the corresponding biplot.

3 Results and Discussion

3.1 Descriptive Statistics

Figure 1 shows the examples of swatches studied in this work. Table 3 shows the summarizes the descriptive statistics for the four classes across the ten numerical features.

At first glance, both the polyester and cotton swatches exhibited subtle variations in their red coloration, with the inner and outer side of the same fiber type demonstrating high color similarity. Although both fiber types were characterized by a warm red hue, as indicated by the mean yellow-blue values (V2), the cotton swatches achieved slightly higher men red-green values (V1) than the polyester swatches, suggesting that the cotton possessed a more vivid red color.

Furthermore, the two fiber types could be differentiated based on their surface texture. Polyester swatches exhibited a higher standard deviation of lightness (V5) and a higher standard deviation of the red-green value (G6), indicating a less uniform surface texture compared to the cotton swatches. Additionally, polyester swatches exhibited markedly higher mean red-to-green (V9) and mean red-to-blue (V10) ratios than the cotton swatches.

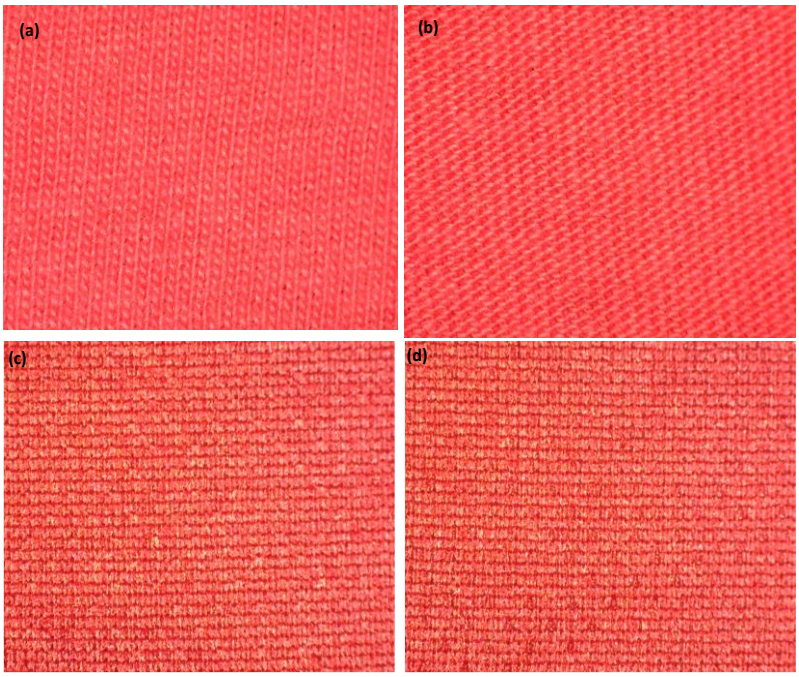


Fig. 1. Examples of (a-b) cotton, (c-d) polyester-based swatches by (a, c) inner, and (b, d) outer sides.

Table 3. Summary statistics of the features.

Samples	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
C.F	62.5	34.4	60.3	71.5	4.0	7.3	-0.9	25.0	32.3	38.5
C.R	63.5	37.2	59.9	73.6	4.4	8.6	-0.9	25.0	26.3	20.0
P.F	55.5	35.0	59.1	65.8	9.0	10.1	-1.1	25.0	91.1	1614
P.R	54.8	35.0	58.9	65.2	9.6	10.4	-1.1	25.0	166.4	2769

3.2 Principal Component Analysis

Figure 2 shows the PCA biplot constructed using the ten numerical features described above. The first two principal components accounted for more than 78% of the total variance in the dataset.

As illustrated, the 6cmx4cm polyester and cotton swatches were clearly differentiated based solely on the numerical features extracted from their digital images. Notably, the polyester swatches exhibited higher intra-shirt variability compared to the cotton swatches. Furthermore, the inner and outer sides of the cotton swatches were distinctly separated, whereas the inner and outer sides of the polyester swatches could not be readily discriminated. In line with the descriptive statistics, cotton swatches were

characterized by relatively higher values of V1 and V4, while polyester swatches were uniquely distinguished by higher values of V5, V6 and V10.

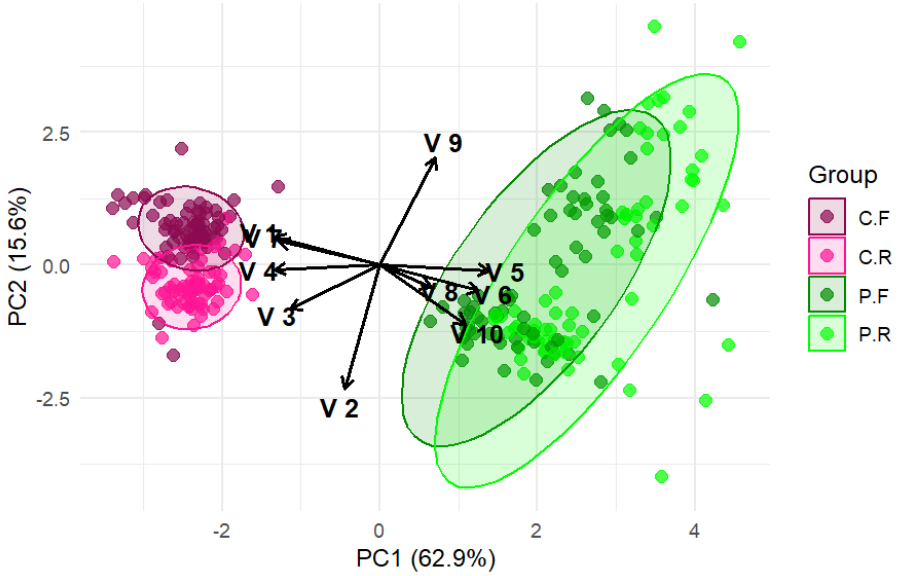


Fig. 2. Biplot (bottom) of PCA estimated using the ten numerical features of the swatches.

4 Discussion and Conclusion

The findings demonstrate the preliminary feasibility of coupling digital microscopy with image processing tools to distinguish between cotton- and polyester-based red swatches. Quantitative features extracted from digital micrographs enable rapid, objective differentiation between the two fabric types—achieving an analytical precision unattainable through conventional visual examination alone.

Overall, automated image analysis provides an objective, non-destructive, and accessible methodology for classifying red textile fibers by material composition, highlighting its potential to enhance forensic fiber evidence evaluation. Nevertheless, these results remain preliminary; rigorous validation across larger, independent fabric cohorts is required prior to operational implementation in forensic casework.

Future work will expand upon these findings by evaluating a broader spectrum of fiber classes and examining how environmental degradation—such as prolonged submersion in diverse aquatic environments—affects dye stability and algorithmic feature extraction. In addition, higher-order colorimetric and textural descriptors will be investigated to characterize subtle shade variations and surface morphologies more comprehensively.

Acknowledgement

Loong Chuen Lee gratefully acknowledges the financial support of CRIM, Universiti Kebangsaan Malaysia (UKM) through Geran Ganjaran [GP-K016373].

References

1. Goodpaster, J.V., Liszewski, E.A.: Forensic analysis of dyed textile fibers, Analytical and Bioanalytical Chemistry 394, 2009-2018 (2009).
2. Hu, C., Mei, H., Guo, H., Zhu, J.: Color analysis of textile fibers by microspectrophotometry. Forensic Chemistry 18, 100221 (2020).
3. Kunkler, K.S., Roy, T.: Reducing the impact of cognitive bias in decision making: Practical actions for forensic science practitioners. Forensic Science International: Synergy 7, 100341 (2023).
4. Lee, L.C., Jemain, A.A.: On overview of PCA application strategy in processing high dimensionality forensic data. Microchemical Journal 169, 106608 (2021).
5. Lim, S.J., Park, Y-K., Kim, H., Kwon, J., Moon, H.M., Lee, Y., Watanabe, A., Teramae, N., Ohtani, H., Kim, Y-M.: Selective solvent extraction and quantification of synthetic microfibers in textile laundary wastewater using pyrolysis-gas chromatography spectrometry, Chemical Engineering Journal 434, 134653 (2022).
6. Muratbekova, M., Toganas, N., Igali, A., Shagyrov, M., Kadyrgali, E., Yerkin, A., Shamoi, P.: Color models in image processing: a review and experimental comparison. Discover Applied Sciences 8, 494 (2026).
7. Nayak, R., Houshyar, S., Khandual, A., Padhye, R., Fergusson, S.: Identification of natural textile fibres. In: Handbook of Natural Fibres, pp. 503-534. Elsevier, Amsterdam, The Netherlands (2020).
8. Powell, R., Collins, P., Horsley, G., Coumbaros, J., van Bronswijk, W.: Enhancing the evidential value of textile fibres Part 2: Application of a database-driven fibre comparison strategy to a cold-case investigation. Forensic Science International 325, 110894 (2021).
9. R Core Team: A Language and Environment for Statistical Computing. R version 4.5.0 (2025-04-11), R Foundation for Statistical Computing, Vienna, Austria (2025).
10. Yerly, M., Bonvin, J., Lim-Hitchings, Y.C., Schaufelbuhl, S., Hess, S., Lunstroot, K., Massonnet, G.: Textile fibres in the respiratory tract part 1: A background study, Forensic Science International 369, 112410 (2025).

Self-Attention-Enhanced 3D Gaussian Splatting for Deformable Endoscopic Surgical Scene Reconstruction

Veena K V and Supriya M H

Department of Electronics, Cochin University of Science and Technology (CUSAT), Kochi, Kerala, India
veenakv@cusat.ac.in, supriyadoe@gmail.com

Abstract. Reconstructing a deformable surgical scene from a single endoscopic camera is one of the more stubborn open problems in computer-assisted surgery. 3D Gaussian Splatting (3DGS) has recently made real-time rendering of such scenes practical by representing the scene as an explicit set of learnable Gaussian primitives, but deformable 3DGS methods typically predict the motion of every Gaussian independently, which discards a useful piece of structure: neighbouring patches of tissue in an endoscopic scene rarely move on their own. Building on SurgicalGaussian, we insert a lightweight local self-attention module (Attn-A) into the deformation network, so that each Gaussian's deformation features are refined by attending to its $K = 16$ nearest neighbours in canonical space. The module adds 46k parameters, is zero-initialised so that training starts exactly at the baseline, and leaves the training objective unchanged. We compare it against our own reproduction of SurgicalGaussian, with every configuration trained with two random seeds using identical code on the same GPU. On the EndoNeRF pulling sequence, Attn-A improves PSNR from 38.39 ± 0.18 to 38.78 ± 0.03 dB ($+0.39$ dB), SSIM from 0.964 to 0.967 and LPIPS from 0.044 to 0.039, and it improves on the baseline in every metric for both seeds. On the cutting sequence it has no measurable effect (37.07 ± 0.20 vs. 37.11 ± 0.05 dB, with the per-seed difference changing sign). The learned attention is close to uniform, which suggests the module acts as a local smoothing prior that helps the coherent motion of tissue retraction but not the discontinuous motion of an incision. This comes at about 40% more training time.

Keywords: 3D Gaussian Splatting · Self-attention · Deformable scene reconstruction · Endoscopic surgery · Surgical data science

1 Introduction

Building a 3D model of the operative field from endoscopic video has obvious uses — surgical training, skill assessment, and augmented-reality guidance during the procedure itself — and it is also a genuinely hard vision problem. Tissue deforms non-rigidly as it is handled, specular highlights move together with

the camera and the light source, instruments periodically block large parts of the view, and depth has to be inferred from a single lens rather than a stereo pair. Neural Radiance Fields (NeRF) [1] and the surgical variants built on top of it produce detailed reconstructions, but their per-pixel volumetric rendering is too slow for anything close to real time. 3D Gaussian Splatting (3DGS) [2] sidesteps this by representing the scene as a set of anisotropic Gaussian primitives rendered through a fast, differentiable rasteriser, and it has quickly become the representation of choice wherever rendering speed matters — surgery being an obvious case in point [23].

The catch is that deformable 3DGS methods usually query their deformation network for each Gaussian on its own. That is a reasonable simplification, but it throws away a fairly obvious prior in surgical video: tissue deformation is spatially coherent. When a fold of tissue is retracted or a plane is opened up with cautery, the Gaussians representing nearby tissue tend to move together rather than independently. We try to give the model direct access to that coherence, using a local self-attention layer so that each Gaussian's deformation estimate is informed by its spatial neighbourhood instead of being computed alone.

Concretely, this paper makes three contributions. First, we add a local self-attention module to the deformation network of SurgicalGaussian [8] that lets deformation features propagate between neighbouring Gaussians, at a cost of 46k extra parameters and with no change to the loss. Second, we compare it against our own reproduction of SurgicalGaussian on both EndoNeRF sequences, with every configuration trained with two random seeds using identical code on the same GPU, and find a consistent gain on the pulling sequence (+0.39 dB PSNR on average, with SSIM and LPIPS improving for both seeds) and no measurable effect on the cutting sequence. Third, we analyse why: the learned attention is close to uniform, so the module behaves as a local smoothing prior, which suits coherent tissue retraction but not the discontinuous motion of an incision (Sect. 5).

2 Related Work

2.1 Neural rendering for endoscopic scenes

EndoNeRF [3] was the first serious attempt at adapting NeRF to monocular endoscopic video, using tool-mask-guided ray sampling together with depth supervision so the network would not waste capacity trying to reconstruct the instrument itself. EndoSurf [4] layered a neural implicit surface constraint on a similar pipeline to recover sharper geometry, and LerPlane [5] decomposed the 4D deformation field into a set of learnable planes, trading some representational power for a large speed-up. All three inherit NeRF's core bottleneck: per-ray volumetric integration is expensive, and none of them approach rates that would be usable inside an operating theatre.

2.2 Gaussian Splatting and its dynamic extensions

3DGS [2] replaced implicit volumetric fields with an explicit set of Gaussian primitives that are rasterised directly, reaching real-time frame rates at high visual quality for static scenes. Extending it to dynamic content has taken several forms. Some methods track a canonical set of Gaussians through time with a learned deformation field [22]; others adapt ideas first developed for dynamic NeRF variants such as Nerfies [16], D-NeRF [17], K-Planes [18], and HexPlane [19] onto a Gaussian backbone [20]; still others treat temporal correspondence itself as the main objective, using persistent tracking across frames [21]. Within surgery specifically, EndoGaussian [6] was among the first to bring 3DGS to endoscopic video, tracking Gaussians over time with a deformation field built on a voxel encoding and a lightweight decoder; Deform3DGS [7] models deformation at the level of individual Gaussians with flexible per-Gaussian deformation parameters shared across time; and SurgicalGaussian [8] uses a forward-mapping deformation MLP together with depth initialisation, tool-mask-guided training, and a regularisation that encourages locally consistent motion of Gaussians. Our work builds directly on SurgicalGaussian.

2.3 Attention mechanisms

Self-attention became the default building block for sequence modelling after the original transformer paper [9], and was later shown to work just as well over spatial feature maps in vision — first through non-local blocks [10], and subsequently through Vision Transformers [11] and hierarchical variants such as Swin [13]. Channel-wise attention, introduced by Squeeze-and-Excitation networks [12], is a lighter alternative that we considered but did not pursue, since our features are indexed by Gaussian identity rather than by feature channel. Attention-based backbones have also found their way into exactly the kind of auxiliary signals — segmentation masks [14] and monocular depth [15] — that surgical reconstruction pipelines already depend on, although we do not use either directly here. To our knowledge, no prior 3DGS-based surgical reconstruction method has applied self-attention over the Gaussian feature set itself, which is the gap this paper addresses.

2.4 Surgical scene understanding more broadly

Beyond dense reconstruction, a substantial body of work addresses related problems in the same domain: instrument detection and tracking [24], monocular SLAM for endoscopy [26,27], and transformer-based depth perception for dynamic surgical scenes [25]. Maier-Hein et al. [23] give a broader account of why this kind of surgical data science matters clinically, and their framing is largely what motivates the introduction above.

3 Method

3.1 Preliminaries

3DGS [2] represents a scene as N Gaussian primitives, each parameterised by a mean position, a covariance matrix, an opacity, and a colour (spherical-harmonic coefficients in the original formulation). Rendering projects the 3D Gaussians onto the image plane and alpha-composites them front to back through a differentiable rasteriser, which is what makes the whole pipeline trainable end-to-end directly from posed RGB images.

3.2 SurgicalGaussian baseline

We build on SurgicalGaussian [8], which keeps a set of canonical Gaussians $\{\mu_i, \Sigma_i, \alpha_i, c_i\}$ and deforms them to each frame time t with an MLP. Positional encodings $\gamma(\mu_i)$ and $\gamma(t)$ pass through an 8-layer trunk of width 256 with a skip connection, producing a feature $h_i \in \mathbb{R}^{256}$ per Gaussian, from which three linear heads predict the position, rotation and scale offsets $(\Delta\mu_i, \Delta r_i, \Delta s_i)$. Training uses SurgicalGaussian's objective unchanged: a tool-masked photometric term, a depth term against the provided depth maps, and its deformation regularisers. These regularisers already encourage nearby Gaussians to move consistently, but only through the loss: the network itself computes h_i from Gaussian i's own position and t alone, so no information passes between neighbours when a deformation is predicted. Attn-A adds that exchange to the architecture.

3.3 Local self-attention module (Attn-A)

Attn-A is inserted between the trunk and the heads (Fig. 1). Full self-attention over all Gaussians is not feasible here — N reaches 10^5–10^6, so an $N \times N$ attention matrix would not fit in memory — so each Gaussian instead attends to its $K = 16$ nearest neighbours $\mathcal{N}(i)$ in canonical space. The neighbour graph is recomputed every 100 iterations and whenever densification adds or removes Gaussians.

The trunk features are normalised and projected to $D = 64$ dimensions, $f_i = W_{\text{in}}\,\text{LN}(h_i)$, and split into queries, keys and values for 4 heads of dimension $d = 16$. For each head, Gaussian i attends over its neighbours with a learned relative-position bias $b_{ij} = \text{MLP}(\tilde{\mu}_j - \tilde{\mu}_i)$, where the offsets are normalised by the largest neighbour distance:

$$a_{ij} = \text{softmax}_{j\in\mathcal{N}(i)}\left(\frac{q_i^\top k_j}{\sqrt{d}} + b_{ij}\right), \qquad o_i = \sum_{j\in\mathcal{N}(i)} a_{ij}\, v_j. \tag{1}$$

The concatenated head outputs are projected back to 256 dimensions and added as a residual,

$$h_i' = h_i + W_{\text{out}}\, o_i, \tag{2}$$

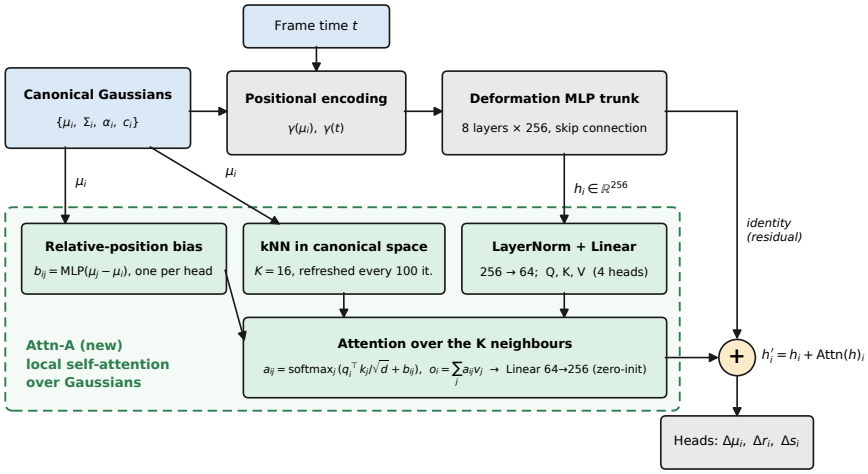


Fig. 1. Deformation network with Attn-A. The SurgicalGaussian trunk (grey) produces a feature h_i per Gaussian; Attn-A (green) lets each Gaussian attend to its $K = 16$ nearest canonical-space neighbours with a relative-position bias, and adds the result back to h_i through a zero-initialised projection before the deformation heads.

and h_i' replaces h_i as input to the three heads. W_{out} is zero-initialised, so at iteration 0 the network is exactly the SurgicalGaussian deformation MLP and attention is learned purely as a residual correction. The module adds 46,340 parameters to the 500,234-parameter deformation network (+9.3%) and introduces no new loss term.

3.4 Training procedure

Fig. 2 summarises one training iteration. Both the baseline and Attn-A are trained for 40,000 iterations with the Adam optimiser [28], using SurgicalGaussian's own configuration for each sequence (learning rates, densification schedule and loss weights), and test PSNR is logged every 1,000 iterations. The two variants differ only in whether the attention module is enabled.

4 Experiments

4.1 Dataset and setup

We evaluate on the EndoNeRF dataset [3], which provides two sequences captured with a da Vinci robotic surgical system at 640×512: *pulling_soft_tissues* (63 frames, 8 held-out test frames, indices $\{1, 9, \ldots, 57\}$) and *cutting_tissues_twice* (156 frames, 20 held-out test frames, every eighth frame from index 1), following the test splits in SurgicalGaussian's released configurations. Both sequences include per-frame RGB, depth, and tool-segmentation masks.

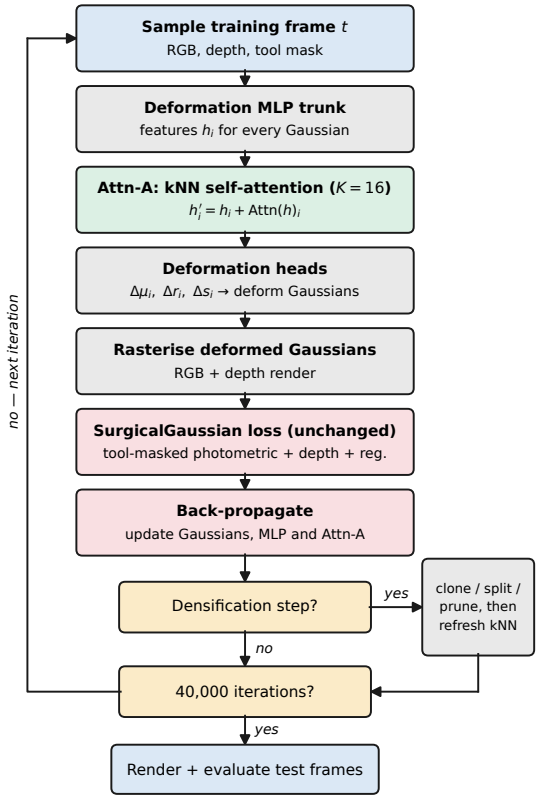


Fig. 2. One training iteration. Attn-A sits inside the deformation network; the loss and the densification schedule are those of SurgicalGaussian, with the kNN graph refreshed after densification.

4.2 Implementation details

All results in Tables 1 and 2 come from one software and hardware environment: SurgicalGaussian's public code in a Singularity container (PyTorch 2.0.1, CUDA 11.8) on a single NVIDIA A30 GPU. Every configuration (baseline and Attn-A, pulling and cutting) is trained with two random seeds (0 and 1), which set the Python, NumPy and PyTorch generators, and we report the mean and standard deviation over the two seeds. Metrics are computed with the repository's own evaluation script on the final (40k) checkpoint, with tool pixels masked out. We report PSNR, SSIM, LPIPS (AlexNet backbone) and inverse LPIPS (iLPIPS = 1 − LPIPS), and the final checkpoint rather than the best logged iteration, so as not to select on the test set. Training to 40k iterations took 1 h 10 min for the baseline and 1 h 36 min for Attn-A on pulling, and 1 h 27 min and 2 h 03 min

Table 1. Comparison on the EndoNeRF pulling and cutting sequences. Published rows are as reported in SurgicalGaussian's comparison table [8]. "Reproduced" and Attn-A are trained and evaluated by us with identical code on the same GPU; each is the mean of two random seeds (standard deviations in Table 2). Bold = best per column.

	Pulling			Cutting		
Method	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
EndoNeRF [3]	34.217	0.938	0.160	34.186	0.932	0.151
EndoSurf [4]	35.004	0.956	0.120	34.981	0.953	0.106
LerPlane [5]	36.241	0.950	0.102	35.580	0.955	0.101
EndoGaussian [6]	37.308	0.958	0.070	**38.287**	**0.962**	0.058
SurgicalGaussian [8] (published)	**38.783**	**0.970**	0.049	37.505	0.961	0.062
SurgicalGaussian (reproduced)	38.386	0.964	0.044	37.111	0.960	0.033
Ours – Attn-A	38.776	0.967	**0.039**	37.074	0.960	**0.032**

on cutting (seed means where two runs were timed), so the kNN attention adds roughly 40% to wall-clock training time.

4.3 Comparison with prior methods

Table 1 places Attn-A alongside published NeRF-based and Gaussian-based methods. All published numbers are taken from SurgicalGaussian's own comparison table [8], which reports the two sequences separately. Deform3DGS [7] publishes only a single aggregate over the whole dataset, so it is not included.

Averaged over two seeds, our reproduction of SurgicalGaussian falls 0.40 dB PSNR below the published number on both sequences (38.39 vs. 38.78 dB on pulling, 37.11 vs. 37.51 dB on cutting). A gap of the same size on both sequences points to differences in software versions, hardware and random seed rather than to a fault in either sequence, and it is the reason we draw conclusions about the attention module from the like-for-like rows below the rule (Sect. 4.4) rather than from the published ones. With that caveat, Attn-A recovers the published SurgicalGaussian PSNR on pulling (38.78 vs. 38.78 dB) and stays below it on cutting (37.07 vs. 37.51 dB), where EndoGaussian remains the strongest method. Attn-A is well ahead of the NeRF-based methods on both sequences, by 1.5–2.5 dB PSNR. Our LPIPS values, for both the reproduced baseline and Attn-A, are lower than all published ones; since we have not confirmed that the published LPIPS uses the same backbone and masking, we do not interpret that difference as a perceptual improvement.

4.4 Ablation: baseline vs. Attn-A

Table 2 and Fig. 3 isolate the effect of the attention module. On the pulling sequence, Attn-A improves on the reproduced baseline for both seeds and in every metric: on average by +0.39 dB PSNR (+0.54 and +0.24 dB for seeds 0 and 1), +0.0025 SSIM and a 12% reduction in LPIPS (0.044 → 0.039). The

Table 2. Ablation on the EndoNeRF sequences at 40k iterations: mean ± standard deviation over two random seeds. Both rows of each block use identical code, configuration and GPU, and differ only in the attention module. Δ: Attn-A minus baseline for the same seed (seed 0 / seed 1).

	Method	PSNR↑	SSIM↑	LPIPS↓
Pulling	Baseline	38.39 ± 0.18	0.9644 ± 0.0009	0.0444 ± 0.0014
	+ Attn-A	$\mathbf{38.78 \pm 0.03}$	$\mathbf{0.9669 \pm 0.0007}$	$\mathbf{0.0389 \pm 0.0013}$
	Δ	+0.54 / +0.24	+0.0037 / +0.0013	−0.0073 / −0.0036
Cutting	Baseline	$\mathbf{37.11 \pm 0.05}$	0.9601 ± 0.0001	0.0327 ± 0.0005
	+ Attn-A	37.07 ± 0.20	$\mathbf{0.9602 \pm 0.0007}$	$\mathbf{0.0322 \pm 0.0015}$
	Δ	−0.14 / +0.07	−0.0003 / +0.0005	+0.0002 / −0.0011

Attn-A mean (38.78 dB) lies above both baseline runs (38.26 and 38.52 dB), and the two Attn-A runs agree to within 0.04 dB.

On the cutting sequence the two variants are indistinguishable. The per-seed difference changes sign (−0.14 dB for seed 0, +0.07 dB for seed 1), the means differ by 0.04 dB, and SSIM and LPIPS differ only in the fourth decimal place. We therefore read the cutting result as no measurable effect of the attention module. Interestingly, the spread between seeds is not consistent either: Attn-A varies less than the baseline on pulling (0.03 vs. 0.18 dB) but more on cutting (0.20 vs. 0.05 dB), so with two seeds we draw no conclusion about the module's effect on training stability.

Convergence on pulling (Fig. 4). For seed 0 the gain is not an artefact of the final checkpoint. Attn-A is ahead of the baseline at 38 of the 40 logged test evaluations, the exceptions being 4k and 5k, before the first densification steps take effect, and at every evaluation from 20k onwards. Averaged over the last 10k iterations its test PSNR is 0.56 dB higher (38.66 vs. 38.10 dB), and it first reaches 38 dB at 20k iterations against 30k for the baseline. Both runs show the same transient dip at 23k, which therefore reflects the shared training schedule rather than the attention module. Training PSNR rises by more than test PSNR (+1.00 vs. +0.54 dB at 40k), so Attn-A's train–test gap is wider than the baseline's (2.58 vs. 2.13 dB): part of the added capacity goes into fitting the training views.

Convergence on cutting (Fig. 5). Unlike pulling, the two curves stay close throughout. For seed 0, Attn-A runs slightly below the baseline over the second half of training (0.16 dB lower on average over the last 10k iterations); for seed 1 the order reverses at the final checkpoint (+0.07 dB), so neither variant is consistently ahead. Convergence is slower and steadier than on pulling, and training and test PSNR stay within 0.2–0.3 dB of each other: on this sequence the model is limited by its ability to fit the deformation rather than by overfitting.

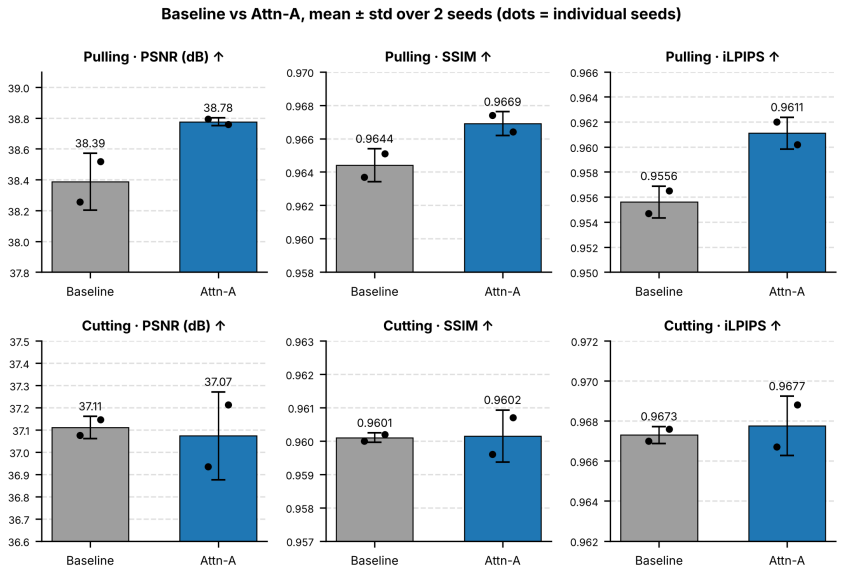


Fig. 3. Final-checkpoint metrics of the reproduced baseline and Attn-A on both sequences: bars are the mean over two seeds, error bars the standard deviation, dots the individual seeds (same code and GPU).

5 Discussion

The two sequences behave in a way that the mechanism of the module explains. The mean attention entropy stays between 2.767 and 2.772 nats throughout training in all Attn-A runs, against a maximum of $\ln 16 \approx 2.773$ for uniform weights over $K = 16$ neighbours. The learned weights are therefore close to uniform, and Attn-A acts mainly as a learned local aggregation of neighbour features: a smoothing prior on the deformation field rather than a selective attention mechanism.

Such a prior matches pulling, where the tissue is retracted as one coherent sheet and neighbouring Gaussians really do move together. It does not match cutting, for three reasons. First, an incision is a motion discontinuity: tissue on either side of the cut separates, and averaging deformation features across the cut line works against exactly the motion that has to be modelled. Second, neighbours are chosen in canonical space, so Gaussians that start adjacent and end up on opposite sides of the incision remain linked by the attention graph. Third, cutting is limited by fitting rather than by overfitting (a train–test gap of 0.2–0.3 dB, against 2.1–2.6 dB on pulling), and an added smoothing prior reduces rather than adds to the effective flexibility of the deformation field. The instrument also covers the incision site for long stretches, so uniform attention can mix tool-driven apparent motion into neighbouring tissue.

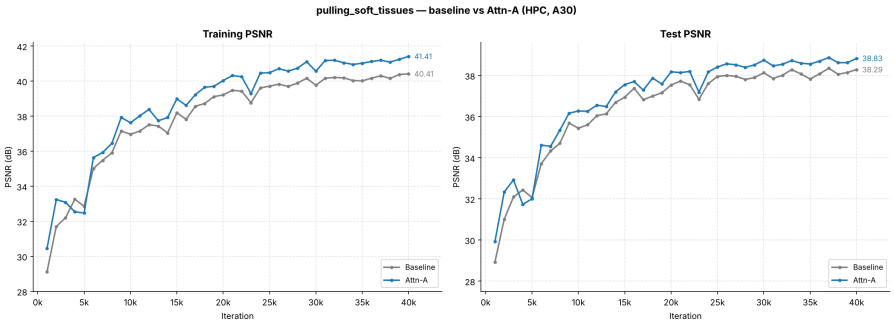


Fig. 4. Training (left) and test (right) PSNR versus iteration on the EndoNeRF pulling sequence for the reproduced baseline (grey) and Attn-A (blue), seed 0, same code and GPU as Table 2. Values are the logged test PSNR, which differs slightly from the evaluation-script value.

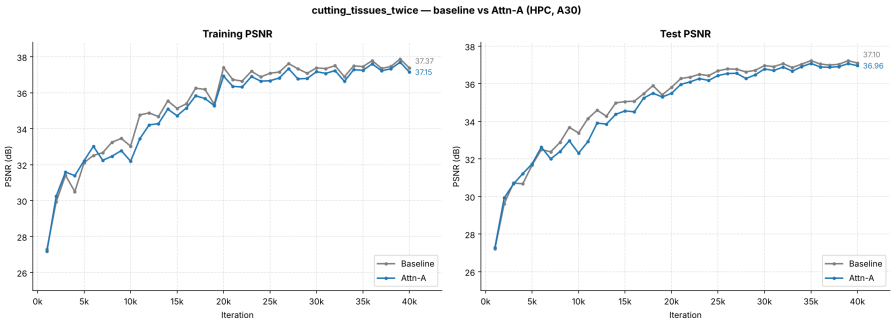


Fig. 5. Training (left) and test (right) PSNR versus iteration on the EndoNeRF cutting sequence for the reproduced baseline (grey) and Attn-A (blue), seed 0, same code and GPU as Table 2.

These explanations are hypotheses at this stage. We are testing them with region-wise error maps (near vs. far from the instrument, high vs. low tissue motion) and with the attention entropy and attention mass on separated neighbours measured per region. The pulling gain itself is modest, and its size depends on the seed (+0.24 to +0.54 dB), so more seeds would tighten the estimate; that it appears for both seeds and in every metric is the stronger part of the evidence. The module is also not free: it adds about 40% to training time, and its effect on rendering speed remains to be measured. Several extensions follow directly. Recomputing neighbours in the deformed rather than the canonical space would stop separated tissue from being linked. Mask-gating, suppressing attention from Gaussians that project onto tool pixels, would stop instrument motion leaking into tissue deformation. Temporal attention across nearby frames could help with the fragmented deformation of the cutting sequence. Finally, the near-

uniform weights suggest that a cheaper neighbourhood-averaging layer might recover most of the pulling gain at a lower cost, which we plan to test directly.

6 Conclusion

We presented Attn-A, a lightweight local self-attention module inserted into the deformation network of SurgicalGaussian, in which each Gaussian attends to its 16 nearest canonical-space neighbours. Against our own reproduction of SurgicalGaussian, with two random seeds per configuration and identical code and hardware, Attn-A improves the EndoNeRF pulling sequence by 0.39 dB PSNR on average (38.78 ± 0.03 vs. 38.39 ± 0.18 dB), with SSIM and LPIPS also improving for both seeds, and has no measurable effect on the cutting sequence (37.07 ± 0.20 vs. 37.11 ± 0.05 dB), at about 40% more training time. The near-uniform attention weights indicate that the module acts as a smoothing prior on deformation, which helps coherent retraction but not the discontinuous motion of an incision. Open questions we plan to follow up on are region-wise evidence for that explanation, neighbour selection in deformed space, a cheaper averaging alternative, rendering-speed cost, and whether the module generalises to other surgical datasets.

References

1. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: NeRF: Representing scenes as neural radiance fields for view synthesis. In: Computer Vision – ECCV 2020. LNCS, vol. 12346, pp. 405–421. Springer (2020)
2. Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G.: 3D Gaussian splatting for real-time radiance field rendering. ACM Trans. Graph. 42(4), 139:1–139:14 (2023)
3. Wang, Y., Long, Y., Fan, S.H., Dou, Q.: Neural rendering for stereo 3D reconstruction of deformable tissues in robotic surgery. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2022. Springer (2022)
4. Zha, R., Cheng, X., Li, H., Harandi, M., Ge, Z.: EndoSurf: Neural surface reconstruction of deformable tissues with stereo endoscope videos. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. Springer (2023)
5. Yang, C., Wang, K., Wang, Y., Yang, X., Shen, W.: Neural LerPlane representations for fast 4D reconstruction of deformable tissues. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. Springer (2023)
6. Liu, Y., Li, C., Yang, C., Yuan, Y.: EndoGaussian: Real-time Gaussian splatting for dynamic endoscopic scene reconstruction. arXiv preprint arXiv:2401.12561 (2024)
7. Yang, S., Li, Q., Shen, D., Gong, B., Dou, Q., Jin, Y.: Deform3DGS: Flexible deformation for fast surgical scene reconstruction with Gaussian splatting. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. LNCS, vol. 15006, pp. 132–142. Springer (2024)
8. Xie, W., Yao, J., Cao, X., Lin, Q., Tang, Z., Dong, X., Guo, X.: SurgicalGaussian: Deformable 3D Gaussians for high-fidelity surgical scene reconstruction. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. LNCS, vol. 15006, pp. 617–627. Springer (2024)

9. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. In: Advances in Neural Information Processing Systems (NeurIPS), vol. 30 (2017)
10. Wang, X., Girshick, R., Gupta, A., He, K.: Non-local neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
11. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N.: An image is worth 16x16 words: Transformers for image recognition at scale. In: International Conference on Learning Representations (ICLR) (2021)
12. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
13. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin Transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2021)
14. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P., Girshick, R.: Segment Anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2023)
15. Ranftl, R., Lasinger, K., Hafner, D., Schindler, K., Koltun, V.: Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer. IEEE Trans. Pattern Anal. Mach. Intell. 44(3) (2022)
16. Park, K., Sinha, U., Barron, J.T., Bouaziz, S., Goldman, D.B., Seitz, S.M., Martin-Brualla, R.: Nerfies: Deformable neural radiance fields. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2021)
17. Pumarola, A., Corona, E., Pons-Moll, G., Moreno-Noguer, F.: D-NeRF: Neural radiance fields for dynamic scenes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021)
18. Fridovich-Keil, S., Meanti, G., Warburg, F.R., Recht, B., Kanazawa, A.: K-Planes: Explicit radiance fields in space, time, and appearance. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2023)
19. Cao, A., Johnson, J.: HexPlane: A fast representation for dynamic scenes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2023)
20. Wu, G., Yi, T., Fang, J., Xie, L., Zhang, X., Wei, W., Liu, W., Tian, Q., Wang, X.: 4D Gaussian splatting for real-time dynamic scene rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
21. Luiten, J., Kopanas, G., Leibe, B., Ramanan, D.: Dynamic 3D Gaussians: Tracking by persistent dynamic view synthesis. In: International Conference on 3D Vision (3DV) (2024)
22. Yang, Z., Gao, X., Zhou, W., Jiao, S., Zhang, Y., Jin, X.: Deformable 3D Gaussians for high-fidelity monocular dynamic scene reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
23. Maier-Hein, L., Vedula, S.S., Speidel, S., et al.: Surgical data science for next-generation interventions. Nat. Biomed. Eng. 1(9) (2017)

24. Bouget, D., Allan, M., Stoyanov, D., Jannin, P.: Vision-based and marker-less surgical tool detection and tracking: A review of the literature. Med. Image Anal. 35 (2017)
25. Long, Y., Li, Z., Yee, C.H., Ng, C.F., Taylor, R.H., Unberath, M., Dou, Q.: E-DSSR: Efficient dynamic surgical scene reconstruction with transformer-based stereoscopic depth perception. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. Springer (2021)
26. Recasens, D., Lamarca, J., Fácil, J.M., Tardós, J.D., Civera, J.: Endo-Depth-and-Motion: Reconstruction and tracking in endoscopic videos using depth networks and photometric constraints. IEEE Robot. Autom. Lett. 6(4) (2021)
27. Ozyoruk, K.B., Gokceler, G.I., Bobrow, T.L., et al.: EndoSLAM dataset and an unsupervised monocular visual odometry and depth estimation approach for endoscopic videos. Med. Image Anal. 71 (2021)
28. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: International Conference on Learning Representations (ICLR) (2015)

A Critical Analysis of Artificial Intelligence Adoption In Higher Education

Samuel Adeniyi Adekunle[1[0000-0002-9230-2924]]; Solace Ameyibor[1[0009-0000-2904-7353]]; Simon Nyema[1[0009-0005-4578-9740]]; Temitope Olajide[1[0009-0006-8471-4087]]; Olajide Oluwatosin[1[0009-0009-5848-2448]], Daniel Katu[1[0009-0007-5074-7582]], Babatunde Ogunbayo[1[0000-0002-4334-9136]]

[1]William VS Tubman University, Harper City, Liberia.
Sadekunle135@gmail.com

Abstract. The rapid advancement of digital technologies has catalysed transformative changes across multiple sectors, with artificial intelligence (AI) emerging as a particularly influential innovation. Despite its widespread application across diverse sectors, the systematic adoption of AI in tertiary education remains comparatively underexplored. This study investigates the landscape of AI adoption in higher education by examining how AI technologies are implemented and identifying the key domains in which they are applied. To achieve this aim, a bibliometric research design was employed. Scholarly publications were retrieved from the Scopus database using carefully defined search keywords, followed by a rigorous data cleaning process to ensure analytical robustness. The refined dataset was subsequently analysed using VOSviewer to map co-occurrence networks, research clusters, and thematic trends within the existing literature. The findings reveal distinct clusters that reflect major areas of AI adoption in higher education. Additionally, the analysis highlights emerging and under-researched themes that indicate potential future trajectories for AI integration. By providing a comprehensive synthesis of the extant literature, this study contributes to a deeper understanding of the evolving role and impact sphere of AI in higher education and identifies critical research gaps for future investigation. The findings offer strategic insights for researchers, institutional leaders, policymakers, and technology developers seeking to inform evidence-based decision-making and sustainable AI adoption within the higher education sector.

Keywords: Higher education, AI, technology adoption, HEI

1 Introduction

Artificial intelligence (AI) has moved from a peripheral experiment to a structural force in higher education. The field spans intelligent tutoring systems, learning analytics, automated assessment and, more recently, generative large language models. Early conceptual work framed AI in education as a means of personalising instruction and scaling support beyond what human educators alone can offer (Luckin *et al.*, 2016; Holmes, Bialik and Fadel, 2019). Evidence of this expansion is now substantial. Reviews show that AI applications in higher education cluster around profiling and prediction, assessment and evaluation, adaptive systems and intelligent tutoring (Zawacki-Richter *et al.*, 2019; Chen, Chen and Lin, 2020). The release of ChatGPT in late 2022 accelerated this trajectory, putting generative AI directly in the hands of students and academics and prompting institutions to rethink teaching, assessment and academic integrity (Kasneci *et al.*, 2023; UNESCO, 2023). AI adoption is therefore no longer a question of whether universities will engage with the technology, but of how, to what extent and with what consequences.

Despite this momentum, the growth of the literature has not been matched by conceptual or methodological maturity. Zawacki-Richter *et al.* (2019) found that educators were largely absent from AI-in-education research, which was dominated by computer science perspectives. Bond *et al.* (2024), in a meta systematic review of AI in higher education, identified persistent weaknesses in ethical reflection, interdisciplinary collaboration and methodological rigour. Crompton and Burke (2023) likewise report that research remains concentrated in particular

regions and disciplines. Critical scholars add that much of the discourse is technologically deterministic and underplays the institutional, political and commercial interests shaping AI in education (Selwyn, 2019; Williamson and Eynon, 2020). Together, these observations suggest that the field is expanding quickly but unevenly, and that the pattern of adoption itself deserves critical scrutiny.

The nature of adoption is only partly captured by prior work. Established frameworks explain why individuals and organisations take up new technologies, including the Technology Acceptance Model (Davis, 1989), the Unified Theory of Acceptance and Use of Technology (Venkatesh *et al.*, 2003), Diffusion of Innovations (Rogers, 2003) and the Technology-Organisation-Environment framework (Tornatzky and Fleischer, 1990). However, these models were developed for earlier and less autonomous technologies. AI raises distinct concerns about algorithmic bias, data privacy, transparency, authorship and the redefinition of academic roles (Slade and Prinsloo, 2013; Popenici and Kerr, 2017). Adoption in higher education therefore cannot be understood only through individual intention. It must also be examined through institutional readiness, governance and ethics, as well as equity of access across contexts.

Evidence on the impact of AI is similarly divided. On the one hand, AI is credited with improving personalisation, feedback, administrative efficiency and early identification of at-risk students (Ouyang, Zheng and Jiao, 2022; Baidoo-Anu and Owusu Ansah, 2023). On the other, scholars warn of threats to academic integrity, over-reliance on automated outputs, erosion of critical thinking and widening digital divides (Cotton, Cotton and Shipway, 2024; Kasneci *et al.*, 2023). Bates *et al.* (2020) argue that whether AI transforms higher education depends less on the technology than on how institutions and educators choose to use it. This tension between promise and risk makes a systematic account of impact necessary, so that evidence can be separated from assertion.

Against this background, a clear gap emerges. While individual reviews address AI applications, ethics or generative AI in isolation, few studies jointly examine the patterns of adoption, the evidence on impact, and the resulting research priorities using a large-scale bibliometric lens. This leaves scholars, institutional leaders and policymakers without an integrated, evidence-based map of the field to guide future inquiry and decision-making. This study addresses that gap through a bibliometric analysis of the literature on AI in higher education. It pursues two objectives: (1) to explore the adoption of AI in higher education; (2) to identify areas for future research on AI adoption in higher education. The study contributes in three ways. First, it provides an empirical account of the structure, growth and geographic distribution of the field. Second, it offers a critical synthesis of how adoption and impact are conceptualised and evidenced. Third, it sets out a research agenda that can direct scholars toward under-researched and ethically significant areas.

2 AI adoption in higher education

Research on AI adoption in higher education has developed along technological, individual and institutional lines. Early reviews show that applications clustered around profiling and prediction, assessment, adaptive systems and intelligent tutoring, often with weak pedagogical grounding (Zawacki-Richter *et al.*, 2019). This suggests that uptake was initially driven by technical capability rather than educational need. At the individual level, studies draw mainly on the Technology Acceptance Model (Davis, 1989) and the Unified Theory of Acceptance and Use of Technology (Venkatesh *et al.*, 2003). Granić and Marangunić (2019) confirm that perceived usefulness and ease of use consistently predict the use of educational technology. Applied to AI, Chatterjee and Bhattacharjee (2020) found that attitude, social influence and facilitating conditions shaped adoption intention among higher education stakeholders, while Strzelecki (2023) showed that performance expectancy and habit influenced students' use of ChatGPT. These results explain individual intention, yet they say little about the organisational conditions that enable or constrain adoption. Tsai *et al.* (2019) argue that institutional leadership, strategic alignment and stakeholder engagement determine whether data-driven innovations become embedded in practice. Critical scholars add that adoption is also shaped by commercial and political interests, as vendors and policy actors promote AI through narratives of efficiency and

inevitability (Williamson and Eynon, 2020). Adoption is geographically uneven too, with research concentrated in a limited number of regions (Crompton and Burke, 2023). Taken together, the literature portrays adoption as a multi-level process, but one that remains fragmented across theories and contexts. This fragmentation justifies a field-wide mapping of how adoption has been studied.

3 Impact of AI in higher education

Evidence on the impact of AI in higher education points to both pedagogical gains and substantial risks. On the positive side, Ouyang, Zheng and Jiao (2022) reviewed empirical studies in online higher education and reported that AI was mainly used for prediction, assessment and evaluation, with several studies noting improved learning performance and engagement. Kasneci *et al.* (2023) similarly argue that large language models can support personalised feedback, content generation and lesson preparation while reducing routine workload. Such benefits indicate that AI can extend support beyond what instructors alone can provide, although the evidence base remains uneven in quality (Zawacki-Richter *et al.*, 2019). Concerns centre on academic integrity and learning quality. Cotton, Cotton and Shipway (2024) show that generative tools complicate plagiarism detection and challenge conventional assessment, and Sullivan, Kelly and McLaughlan (2023) highlight the risk of over-reliance weakening independent learning. Ethical impacts are equally prominent. Slade and Prinsloo (2013) identify issues of privacy, consent and the unfair classification of students through data, concerns that intensify with opaque algorithmic systems (Selwyn, 2019). Dwivedi *et al.* (2023) further warn of bias, misinformation and unclear authorship in generative AI outputs. Bates *et al.* (2020) contend that outcomes depend less on the technology than on how institutions design and govern its use. The literature therefore shows that impact is contingent and contested, and few studies trace long-term effects on learning outcomes, equity or academic roles. This motivates a systematic account of where evidence exists and where it is absent.

4 Methodology

This study adopted a bibliometric analysis to map the intellectual structure of research on AI in higher education. Bibliometrics applies quantitative techniques to bibliographic data, which allows large bodies of literature to be examined objectively and reproducibly (Zupic and Čater, 2015; Donthu *et al.*, 2021). The approach suits the three objectives because it reveals publication performance, collaboration patterns and thematic structures, from which adoption, impact and research gaps can be inferred.

Data were retrieved exclusively from Scopus. Scopus was selected for its broad, curated coverage of peer-reviewed journals and its rich citation metadata, which suit science mapping (Baas *et al.*, 2020; Pranckutė, 2021). Using a single database also avoids duplicate records and inconsistent metadata formats, although it limits coverage, as acknowledged in the limitations. The search used the keywords "AI" or “artificial intelligence” and "higher education", applied to article titles, abstracts and author keywords. Records were then filtered by language and a total 9876 publications were adopted by the study. The period covered was 1986 to 2025. This is because the study did not restrict the start date only the end date (2025) was restricted. Identification, screening and eligibility decisions followed the logic of the PRISMA flow (Page *et al.*, 2021) to keep the process transparent.

The final dataset was exported with full records and cited references in CSV and BibTeX formats and analysed using VOSviewer (van Eck and Waltman, 2010). Science mapping, comprising keyword co-occurrence, co-citation and thematic mapping, identified conceptual clusters and their evolution. The clusters were interpreted critically to establish what the field emphasises, what it neglects and where evidence is thin, which informed the research agenda in the third objective. Cluster labels were assigned through interpretive reading of the constituent keywords and highly cited documents, and the author reviewed them to reduce subjectivity.

5 Findings

Artificial intelligence occupies the centre of the network and links all clusters, which shows that the field is conceptually connected despite its breadth. The red cluster, *machine learning and predictive analytics*, reflects the algorithmic foundation of the field and supports earlier findings that prediction and profiling dominate applications (Zawacki-Richter *et al.*, 2019). The purple cluster, *technology acceptance and adoption*, shows that adoption research is still anchored in acceptance models (Davis, 1989; Venkatesh *et al.*, 2003), which suggests that adoption is conceptualised mainly at the individual level. The orange cluster, *generative AI, ChatGPT and AI literacy*, marks a recent shift towards user perceptions and literacy (Kasneci *et al.*, 2023; Strzelecki, 2023). The cyan, dark blue and green clusters capture impact on teaching, assessment, ethics and learner outcomes. Ethical terms such as *ethical issues* and *fairness* appear small and peripheral, which indicates limited attention to the ethical dimensions raised by critical scholars (Slade and Prinsloo, 2013; Selwyn, 2019). Among the labels shown, only a few countries appear, which points to uneven geographic coverage that the country analysis should confirm.

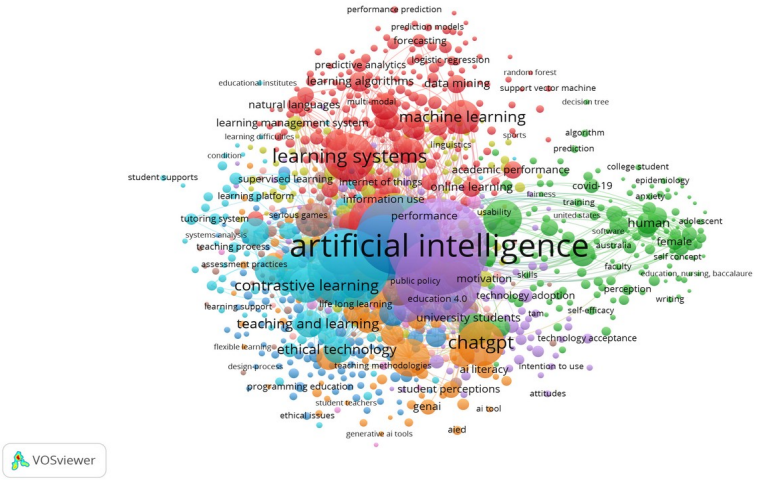


Figure 1: Visualisation of clusters

Table 1: Cluster explanation summary

Colour/theme	Keywords	Explanation
Red **Machine learning and predictive analytics**	learning systems, machine learning, learning algorithms, data mining, predictive analytics, prediction models, forecasting, logistic regression, random forest, support vector machine, decision tree, natural languages, learning management system	The technical and algorithmic strand. AI is treated as a data-driven tool for predicting performance and modelling learner behaviour.
Purple **Acceptance and adoption**	technology adoption, technology acceptance, TAM, intention to use, attitudes, self-efficacy, motivation, skills, university students, education 4.0, public policy	The behavioural and theoretical strand on why stakeholders take up AI, drawing on acceptance

		models.
Orange **Generative AI, ChatGPT and literacy**	GenAI, generative AI tools, AI literacy, AIED, student perceptions, AI tool	The recent surge after generative AI. It centres on perceptions, literacy and use of conversational tools.
Cyan **AI-supported teaching, tutoring and assessment**	teaching and learning, tutoring system, assessment practices, learning support, learning platform, teaching process, flexible learning, life long learning, contrastive learning	The pedagogical strand, covering how AI supports instruction, feedback, assessment and learner support.
Dark blue **Ethics and pedagogical design**	ethical technology, ethical issues, design process, programming education, teaching methodologies, student teachers	Smaller and more peripheral, linking ethics and curriculum or teaching design.
Green **Learner characteristics and contextual factors**	human, female, adolescent, college student, anxiety, self concept, perception, COVID-19, faculty, usability, fairness,	The human and contextual strand, covering psychological factors, the pandemic period and a few national settings.

6 Areas for further research

The overlay visualisation (Figure 2) shows the temporal evolution of the keywords, coloured by average publication year from 2022 to 2025. Earlier work is concentrated in the machine learning and predictive analytics cluster, where algorithms such as random forest, support vector machine and decision tree are associated with the oldest average dates. The adoption cluster occupies an intermediate position, while the most recent keywords, including *ChatGPT*, *GenAI*, *AI literacy*, *student perceptions* and *ethical technology*, appear at the yellow end of the scale. This reveals a shift from predictive and administrative applications towards generative tools and the literacy and ethical questions they raise (Kasneci *et al*., 2023; Strzelecki, 2023). Ethics-related terms are among the newest yet remain small and peripheral, indicating that scholarly attention is emerging but not yet consolidated, in line with critical arguments (Slade and Prinsloo, 2013; Selwyn, 2019). Terms relating to institutional policy, equity and learning outcomes were comparatively weak, which signals priorities for future research.

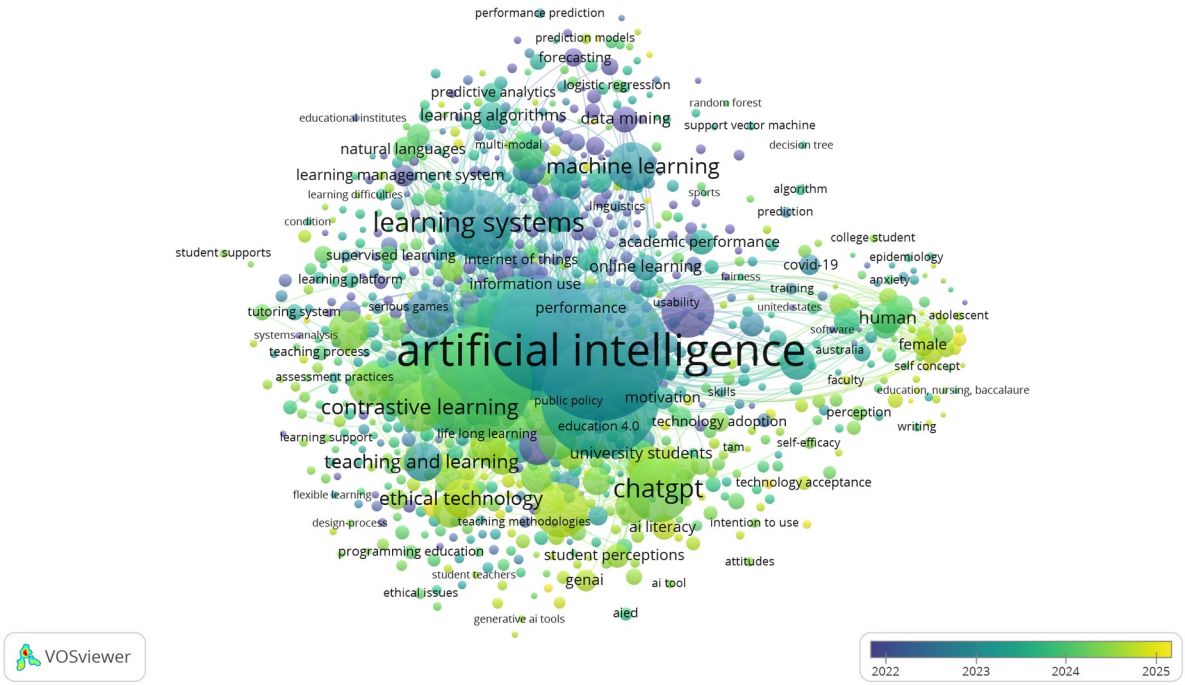


Figure 2: Overlay visualisation

1. Ethics and governance are still thin: *Ethical issues* and *fairness* are small and weakly connected to the central nodes, and *public policy* is small and sits in the teal (2023) range. Governance, privacy and academic integrity are not visible as distinct themes despite their prominence in the critical literature (Slade and Prinsloo, 2013; Cotton, Cotton and Shipway, 2024).
2. Institutional perspectives: Adoption is framed through individual acceptance constructs, while institutional readiness, leadership and policy are barely visible (Tsai *et al.*, 2019).
3. Faculty and educator voices: *Faculty* is small and peripheral, echoing the long-standing concern about the absence of educators (Zawacki-Richter *et al.*, 2019).
4. Assessment redesign. *Assessment practices* are small and older in colour, with little recent link to generative AI. There is a need to review how assessments are administered and examined.
5. Long-term impact and outcomes. Recent keywords are dominated by perceptions rather than measured learning outcomes.

7 Conclusion

This study examined the adoption and impact of artificial intelligence in higher education and identified priorities for future research through a bibliometric analysis of Scopus-indexed literature. The findings reveal a rapidly expanding, technically anchored field in which artificial intelligence links predictive analytics, technology acceptance, generative AI, pedagogy, ethics and learner-related factors.

Adoption scholarship remains dominated by individual-level explanations, particularly intention, attitudes and self-efficacy, while institutional readiness, governance and policy are comparatively underdeveloped. Impact research has shifted from predictive modelling towards teaching, assessment and, most recently, generative AI and AI literacy. Yet much recent work documents perceptions rather than demonstrated effects on learning, equity or academic roles, so claims of transformation appear to outpace the evidence. These patterns carry clear implications. Researchers should integrate ethics and governance into adoption research, complement individual-level models with institutional analysis, and generate longitudinal, comparative evidence on learning outcomes and assessment integrity. University leaders and policymakers should avoid basing decisions on perception-based evidence alone and should pursue context-sensitive strategies, particularly in under-represented regions.

Future research should combine databases, add co-citation and bibliographic coupling analyses, and pair bibliometric mapping with systematic content analysis. Ultimately, the field's next phase will depend less on demonstrating technical possibility than on methodological rigour, ethical integration and contextual breadth.

References

Aria, M. and Cuccurullo, C. (2017) 'bibliometrix: an R-tool for comprehensive science mapping analysis', *Journal of Informetrics*, 11(4), pp. 959-975.

Baidoo-Anu, D. and Owusu Ansah, L. (2023) 'Education in the era of generative artificial intelligence (AI): understanding the potential benefits of ChatGPT in promoting teaching and learning', *Journal of AI*, 7(1), pp. 52-62.

Bates, T., Cobo, C., Mariño, O. and Wheeler, S. (2020) 'Can artificial intelligence transform higher education?', *International Journal of Educational Technology in Higher Education*, 17, Article 42.

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S.W. and Siemens, G. (2024) 'A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour', *International Journal of Educational Technology in Higher Education*, 21, Article 4.

Chatterjee, S. and Bhattacharjee, K.K. (2020) 'Adoption of artificial intelligence in higher education: a quantitative analysis using structural equation modelling', *Education and Information Technologies*, 25(5), pp. 3443-3463.

Chen, L., Chen, P. and Lin, Z. (2020) 'Artificial intelligence in education: a review', *IEEE Access*, 8, pp. 75264-75278.

Cotton, D.R.E., Cotton, P.A. and Shipway, J.R. (2024) 'Chatting and cheating: ensuring academic integrity in the era of ChatGPT', *Innovations in Education and Teaching International*, 61(2), pp. 228-239.

Crompton, H. and Burke, D. (2023) 'Artificial intelligence in higher education: the state of the field', *International Journal of Educational Technology in Higher Education*, 20, Article 22.

Davis, F.D. (1989) 'Perceived usefulness, perceived ease of use, and user acceptance of information technology', *MIS Quarterly*, 13(3), pp. 319-340.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N. and Lim, W.M. (2021) 'How to conduct a bibliometric analysis: an overview and guidelines', *Journal of Business Research*, 133, pp. 285-296.

Dwivedi, Y.K., Kshetri, N., Hughes, L., Slade, E.L., Jeyaraj, A., Kar, A.K., Baabdullah, A.M., Koohang, A., Raghavan, V. and Ahuja, M. (2023) 'So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy', International Journal of Information Management, 71, p. 102642. doi: 10.1016/j.ijinfomgt.2023.102642.

Granić, A. and Marangunić, N. (2019) 'Technology acceptance model in educational context: a systematic literature review', *British Journal of Educational Technology*, 50(5), pp. 2572-2593.

Holmes, W., Bialik, M. and Fadel, C. (2019) *Artificial intelligence in education: promises and implications for teaching and learning*. Boston: Center for Curriculum Redesign.

Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, Stephan Krusche, Gitta Kutyniok, Tilman Michaeli, Claudia Nerdel, Jürgen Pfeffer, Oleksandra Poquet, Michael Sailer, Albrecht Schmidt, Tina Seidel, Matthias Stadler (2023) 'ChatGPT for good? On opportunities and challenges of large language models for education', *Learning and Individual Differences, 103, 1-9.* https://doi.org/10.1016/j.lindif.2023.102274

Luckin, R., Holmes, W., Griffiths, M. and Forcier, L.B. (2016) *Intelligence unleashed: an argument for AI in education*. London: Pearson.

Ouyang, F., Zheng, L. and Jiao, P. (2022) 'Artificial intelligence in online higher education: a systematic review of empirical research from 2011 to 2020', *Education and Information Technologies*, 27, pp. 7893-7925.

Popenici, S.A.D. and Kerr, S. (2017) 'Exploring the impact of artificial intelligence on teaching and learning in higher education', *Research and Practice in Technology Enhanced Learning*, 12, Article 22.

Pritchard, A. (1969) 'Statistical bibliography or bibliometrics?', *Journal of Documentation*, 25(4), pp. 348-349.

Rogers, E.M. (2003) *Diffusion of innovations*. 5th edn. New York: Free Press.

Selwyn, N. (2019) *Should robots replace teachers? AI and the future of education*. Cambridge: Polity Press.

Slade, S. and Prinsloo, P. (2013) 'Learning analytics: ethical issues and dilemmas', *American Behavioral Scientist*, 57(10), pp. 1510-1529.

Strzelecki, A. (2023) 'To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology', *Interactive Learning Environments*. doi:10.1080/10494820.2023.2209881.

Sullivan, M., Kelly, A. and McLaughlan, P. (2023) 'ChatGPT in higher education: considerations for academic integrity and student learning', *Journal of Applied Learning and Teaching*, 6(1), pp. 31-40.

Tornatzky, L.G. and Fleischer, M. (1990) *The processes of technological innovation*. Lexington, MA: Lexington Books.

Tsai, Y.-S., Poquet, O., Gašević, D., Dawson, S. and Pardo, A. (2019) 'Complexity leadership in learning analytics: drivers, challenges and opportunities', *British Journal of Educational Technology*, 50(6), pp. 2839-2854.

UNESCO (2023) *Guidance for generative AI in education and research*. Paris: UNESCO. https://doi.org/10.54675/EWZM9535

van Eck, N.J. and Waltman, L. (2010) 'Software survey: VOSviewer, a computer program for bibliometric mapping', *Scientometrics*, 84(2), pp. 523–538.

van Eck, N.J. and Waltman, L. (2014) 'Visualizing bibliometric networks', in Ding, Y., Rousseau, R. and Wolfram, D. (eds) Measuring scholarly impact: methods and practice. Cham: Springer, pp. 285–320.

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003) 'User acceptance of information technology: toward a unified view', *MIS Quarterly*, 27(3), pp. 425-478.

Waltman, L., van Eck, N.J. and Noyons, E.C.M. (2010) 'A unified approach to mapping and clustering of bibliometric networks', Journal of Informetrics, 4(4), pp. 629-635.

Williamson, B. and Eynon, R. (2020) 'Historical threads, missing links, and future directions in AI in education', *Learning, Media and Technology*, 45(3), pp. 223-235.

Zawacki-Richter, O., Marín, V.I., Bond, M. and Gouverneur, F. (2019) 'Systematic review of research on artificial intelligence applications in higher education- where are the educators?', *International Journal of Educational Technology in Higher Education*, 16, Article 39.

Zhai, X., Chu, X., Chai, C.S., Jong, M.S.Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J. and Li, Y. (2021) ‘A review of artificial intelligence (AI) in education from 2010 to 2020’, Complexity, 2021, Article 8812542

Zupic, I. and Čater, T. (2015) 'Bibliometric methods in management and organization', *Organizational Research Methods*, 18(3), pp. 429-472.

Integrated Urban Wind-Field Assessment and Predictive Control for Autonomous Systems under Multi-Directional Inflow: A Downtown Tulsa Case Study

Ahmed Ghareeb, Jamey Jacob, and Kursat Kara

School of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, OK 74078, USA

Abstract. Urban wind fields are characterized by strong spatial and directional variability caused by interactions between atmospheric flow and building geometry, which can degrade autonomous aerial-system performance in complex urban environments. This study couples multi-directional urban wind-field characterization with wind-aware predictive-control assessment in downtown Tulsa, Oklahoma. Steady incompressible Reynolds-averaged Navier–Stokes (RANS) simulations were performed in OpenFOAM v2606 using the standard k–ε turbulence model for the selected 47-building subset. Four cardinal inflow directions were evaluated using a logarithmic wind profile referenced to 5 m/s at 100 m. Velocity amplification, wake deficit, modeled turbulent kinetic energy, vertical velocity, and horizontal velocity gradients were assessed at 10, 25, 50, and 100 m. Across the evaluated conditions, northward flow (+Y) most frequently ranked highest in directional severity, whereas westward flow (−X) generated the greatest opposing wind and speed exposure along the reference route. The four RANS fields were then used to compare wind-unaware MPC, current-wind persistence, and spatial wind-preview MPC under identical vehicle dynamics, trajectory, constraints, and tuning. Spatial wind-preview MPC reduced route-active three-dimensional tracking RMSE by 63.1–71.3% relative to wind-unaware MPC under the same applied wind. Current-wind persistence achieved comparable tracking performance across the four inflow cases. A bounded disturbance-gain sensitivity analysis retained the wind-information advantage across the tested gains. The framework links urban airflow characteristics to control performance and motivates future LES, wind-aware planning, and obstacle avoidance.

Keywords: Urban wind; computational fluid dynamics; Reynolds-averaged Navier–Stokes; model predictive control; wind-informed control; autonomous systems.

1 Introduction

Urban wind is influenced by buildings through channeling, flow separation, and wakes. Taller buildings can redirect airflow around corners and over rooftops, while streets can accelerate the approaching wind. This means that even with a fixed inflow, wind conditions vary with altitude and location. Computational fluid dynamics (CFD) can capture these patterns and provide local wind fields to assess the performance of

autonomous aerial systems [1–4]. For autonomous flight, this variation changes the disturbance encountered along a route and can influence tracking even when the approaching wind remains steady.

Urban wind has been studied in various aircraft simulation, planning, and control applications. Vuppala et al. [5] and Krawczyk et al. [10] investigated wind models for flight dynamics. Ebert et al. [6], Chan et al. [8], and Wang et al. [9] investigated wind-aware trajectory planning, while Wu et al. [11] used reinforcement learning for navigation through urban wind zones. Kakavitsas et al. [12] combined CFD-generated urban wind fields with quadrotor flight simulation. These studies motivate a more detailed comparison of wind-information strategies in predictive control. A key question is whether the controller gains substantially from future spatial wind information when it already accounts for the wind at its current position.

This study first compares the four cardinal wind directions at the downtown Tulsa analysis region with the same building geometry. We then assess wind exposure on a prescribed reference trajectory and then consider three model predictive control (MPC) strategies: prediction without wind information, constant prediction based on the current local wind, and spatial prediction sampled along the route to be followed. Other tests investigate the impact of prediction horizon, initial velocity and the gain of the disturbance model. Comparing the strategies with the same vehicle, route, and control settings makes the effect of their wind forecasts easier to interpret.

2 Urban wind field assessment

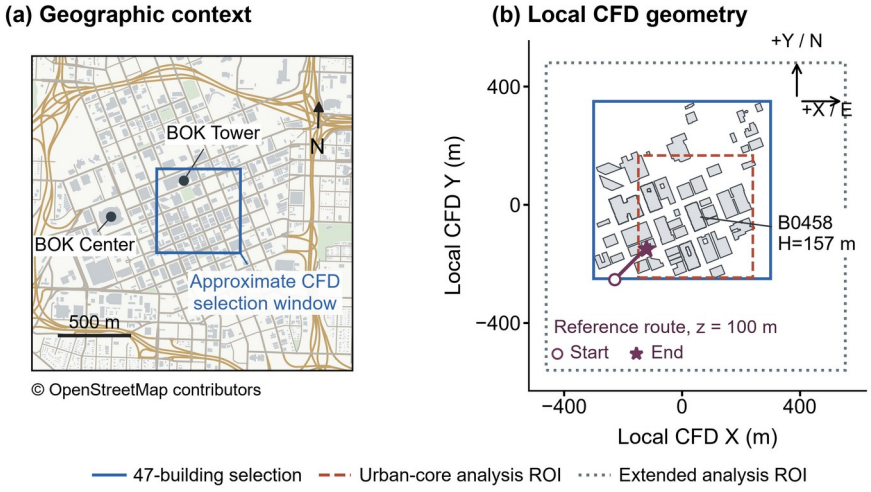


Fig. 1. Downtown Tulsa context and analysis geometry. (a) OpenStreetMap geographic context (© OpenStreetMap contributors, ODbL); blue box: approximate location of the 47-building CFD selection. (b) Selected building footprints and the 149.88-m reference route at z = 100 m. Rectangles mark the whole-building selection window (blue) and analysis ROIs (red/gray). B0458 denotes the tallest included building (H = 157 m).

The 47 buildings were selected using strict whole-building containment within a 600 × 600 m window. The B0168 geometry group, associated with BOK Tower, was excluded because its bounding box extended approximately 0.77 m beyond the northern

selection boundary. Geometry and route coordinates use local east–north–up (ENU) axes.

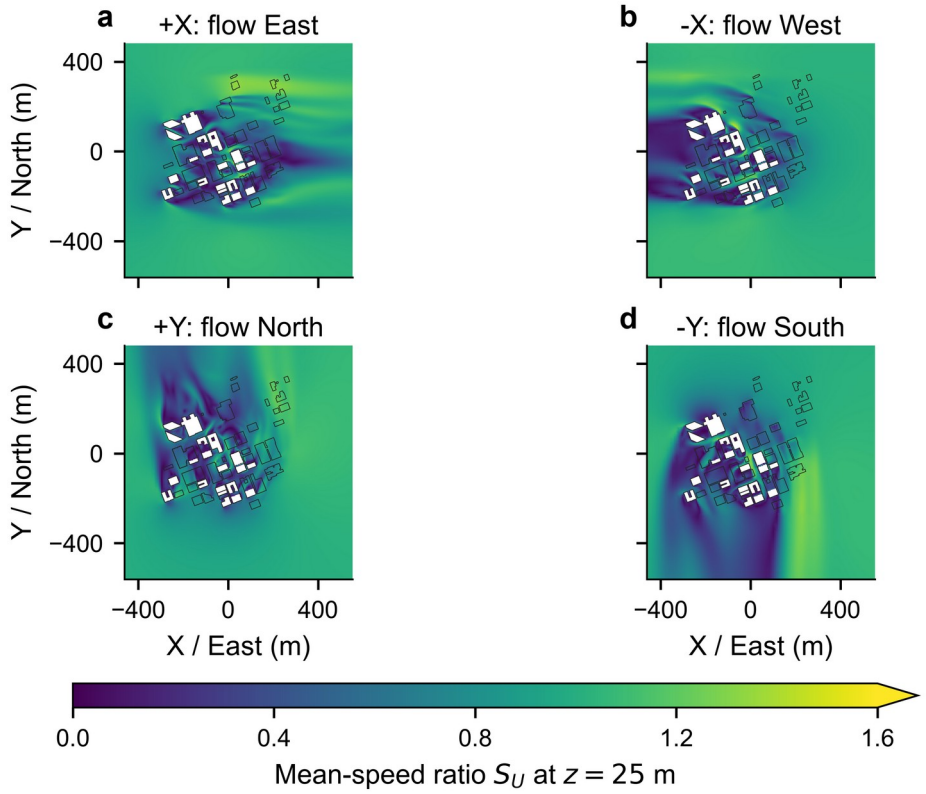


Fig. 2. Normalized mean wind speed at 25 m for four inflow directions. All panels use the same color scale; white regions indicate locations without fluid data.

Steady incompressible RANS with SIMPLE and standard k–ε closure was used in OpenFOAM v2606 [13,14]:

$$\begin{aligned} \nabla \cdot \overline{\mathbf{U}} = 0, \quad \nu_t = C_\mu k^2/\varepsilon, \\ (\overline{\mathbf{U}} \cdot \nabla)\overline{\mathbf{U}} = -\nabla p^\star + \nabla \cdot \tau_{eff}, \\ \tau_{eff} = (\nu + \nu_t)(\nabla \overline{\mathbf{U}} + \nabla \overline{\mathbf{U}}^T). \end{aligned} \tag{1}$$

Here $p^\star = P/\rho + 2k/3$ is modified kinematic pressure and $C_\mu = .09$. Inflow follows Eq. (2), with $U_{\text{ref}} = 5$ m/s, $z_{\text{ref}} = 100$ m and $z_0 = .5$ m [15]. Velocity and turbulence profiles are prescribed at the inlet; outlet pressure is zero. Buildings and ground are no-slip with wall functions; top and lateral boundaries are symmetry planes.

$$U(z) = U_{ref} \frac{\ln\left[(z + z_0)/z_0\right]}{\ln\left[(z_{ref} + z_0)/z_0\right]}. \tag{2}$$

Flows toward $+X$, $-X$, $+Y$, $-Y$ were sampled at 10, 25, 50 and 100 m on a common 2-m grid. Figure 2 uses approaching-flow normalization and a common color scale.

Twelve indicators form three four-indicator groups: speed amplification, wake/turbulence, and vertical motion/horizontal gradients. Directionwise normalization and weighting give

$$\widetilde{f}_{d,j} = \frac{f_{d,j} - \min_{d'} f_{d',j}}{\max_{d'} f_{d',j} - \min_{d'} f_{d',j}},$$
$$g_{d,q} = \frac{1}{4} \sum_{j \in \mathcal{F}_q} \widetilde{f}_{d,j}, \quad C_d = \sum_{q=A,B,C} w_q g_{d,q}. \tag{3}$$

At each height and region, d is direction, j an indicator and F_q a group; constant indicators contribute zero. Equal weights and permutations of (.5,.25,.25) define four schemes: 4 altitudes × 3 regions × 4 schemes = 48 configurations, with four directions compared within each.

The production mesh has 26.90 million cells; a 53.01-million-cell refinement tests selected monitors for southward flow.

3 Predictive control method

The ten-state M100 command-response model [16] couples first-order attitude, yaw-rate and vertical-speed channels to position and velocity through small-angle translation (Fig. 3). The published transfer functions are:

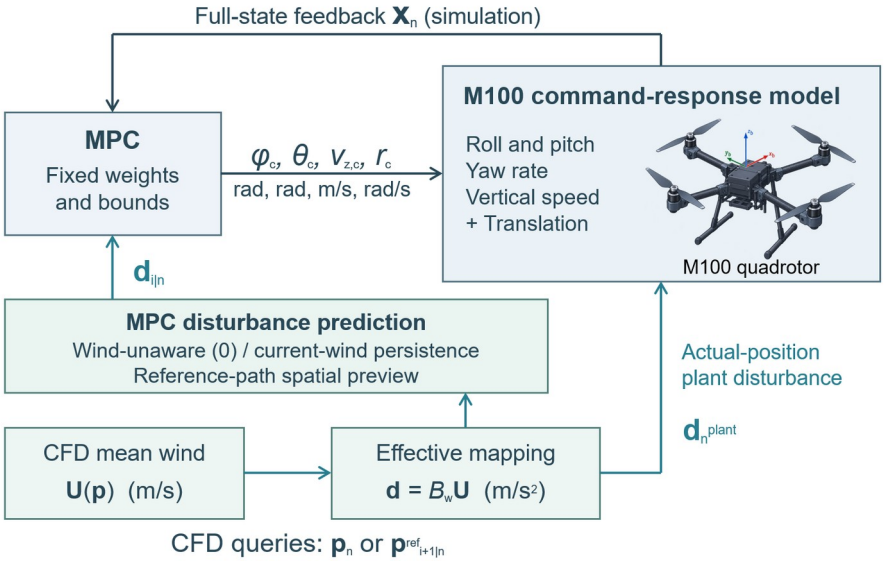


Fig. 3. M100 command-model simulation and MPC wind interfaces.

$$G_\phi(s) = \frac{3.544}{s+2.118}, \quad G_\theta(s) = \frac{3.827}{s+2.430},$$
$$G_r(s) = \frac{5.642}{s+5.268}, \quad G_{v_z}(s) = \frac{3.342}{s+2.990}. \tag{4}$$

The executed roll, pitch and vertical-speed gain/time-constant pairs are (1.673,.472 s), (1.575,.411 s) and (1.118,.334 s); yaw rate uses the published coefficients directly.

Trilinear interpolation samples wind at the actual vehicle position. The effective B_w map in Eq. (5) converts velocity to acceleration; its 18 identification perturbations use a separate 63-g multirotor.

The nominal map is shared by plant and predictor. Sensitivity cases vary only plant disturbance gain by ±25%, keeping the predictor and controller unchanged.

$$\mathbf{d} = B_w \mathbf{U}, \quad B_w = \begin{bmatrix} 0.06815714 & -0.00006228 & 0 \\ 0.00006028 & 0.06854225 & 0 \\ -0.00046850 & 0.00167616 & 0 \end{bmatrix} s^{-1}. \tag{5}$$

Zero-order hold gives the discrete model at T_s = .1 s. Index $i|n$ denotes i prediction steps after update n:

$$\mathbf{x}_{i+1|n} = A_d \mathbf{x}_{i|n} + B_d \mathbf{u}_{i|n} + E_d \mathbf{d}_{i|n}. \tag{6}$$

The matrices A_d, B_d and E_d map the state vector, control inputs, and disturbance accelerations, respectively, to the next predicted state. Table 1 gives the settings for Eq. (7) [17,18], ordered as (x,y,z,v_x,v_y,v_z,φ,θ,ψ,r) and (φ_c,θ_c,$v_{z,c}$,r_c). Angles use radians. Command bounds are applied during optimization; state limits are checked afterward.

$$\begin{aligned} J = & \sum_{i=1}^{N-1} \| \mathbf{x}_i - \mathbf{x}_i^{ref} \|_Q^2 + \| \mathbf{x}_N - \mathbf{x}_N^{ref} \|_{2Q}^2 \\ & + \sum_{i=0}^{N-1} \left(\| \mathbf{u}_i \|_R^2 + \| \Delta \mathbf{u}_i \|_{R_\Delta}^2 \right). \end{aligned} \tag{7}$$

Table 1. Executed MPC settings ($R_{\Delta u} = R_\Delta$). Increments are per 0.1-s step; convergence uses SI/radian command coordinates.

Setting	Executed value
$diagQ$	$(60, 60, 46, 20, 20, 20, 30, 32, 3, 1)$
diag$R$; diag$R_{\Delta u}$	(121.39,121.39,15,3); (.15,.15,.001,.01)
Timing; terminal	T_s = .1 s; N = 20 (2 s); final weight 2Q
Command bounds	$\|\varphi_c\|$,$\|\theta_c\| \le 18.53^\circ$; $\|v_{z,c}\| \le 2$ m/s; $\|r_c\| \le 100^\circ$/s
Increment bounds	$\|\Delta\varphi_c\|$,$\|\Delta\theta_c\| \le 4^\circ$; $\|\Delta v_{z,c}\| \le .35$ m/s; $\|\Delta r_c\| \le 25^\circ$/s per step
Iterations	100 maximum; step .95/$\lambda_{\max}(H)$; stop if $\|\boldsymbol{U}^{j+1} - \boldsymbol{U}^j\|_\infty < 10^{-5}$
State checks after solve	$\|x\|$,$\|y\| \le 1000$ m; $1 \le z \le 300$ m; $\|v_x\|$,$\|v_y\|$,$\|v_z\| \le 15$ m/s; $\|\varphi\|$,$\|\theta\| \le 35^\circ$; $\|r\| \le 100^\circ$/s
Feasibility after solve	Maximum bound violation < 10^{-4} and finite commands

The plant receives actual-position disturbance in every wind-present case. MPC uses zero disturbance, current wind held constant, or future reference-path queries:

$$\begin{aligned} \mathbf{d}_n^{plant} &= B_w \overline{\mathbf{U}}(\mathbf{p}_n), \\ \mathbf{d}_{i|n}^{none} &= \mathbf{0}, \quad \mathbf{d}_{i|n}^{persist} = B_w \overline{\mathbf{U}}(\mathbf{p}_n), \\ \mathbf{d}_{i|n}^{spatial} &= B_w \overline{\mathbf{U}}(\mathbf{p}^{ref}(t_n + (i+1)T_s)). \end{aligned} \tag{8}$$

The future-route and current actual-position wind samples differ by

$$\Delta \mathbf{U}_{i|n} = \overline{\mathbf{U}}(\mathbf{p}_{i+1|n}^{ref}) - \overline{\mathbf{U}}(\mathbf{p}_n), \quad \Delta \mathbf{d}_{i|n} = B_w \Delta \mathbf{U}_{i|n}. \tag{9}$$

All cases follow the same 31-point, 149.88-m route at 100 m and 4 m/s. The vehicle starts from rest; the reference reaches the endpoint at 37.47 s and holds for 4 s. Route-active RMSE (Eq. (10)) uses $2.5 \le t < 37.47$ s. A calm-air run provides the tracking reference.

$$e_{RMSE} = \sqrt{\frac{1}{|\mathscr{J}|}\sum_{j\in\mathscr{J}} \| \mathbf{p}_j - \mathbf{p}_j^{ref} \|_2^2}. \tag{10}$$

4 Results

Across the 48 altitude/region/weight configurations, $+X$, $-X$, $+Y$, $-Y$ lead in 11, 4, 19 and 14 cases. Alternative approaching-flow normalization changes six winners. Thus directional severity depends on the region, altitude, weighting and background reference.

On the reference route, $-X$ yields the highest mean speed (5.231 m/s), strongest mean opposing component (−3.840 m/s) and largest speed-exposure integral (196.18 m). Route exposure (Fig. 4) therefore differs from area-wide severity.

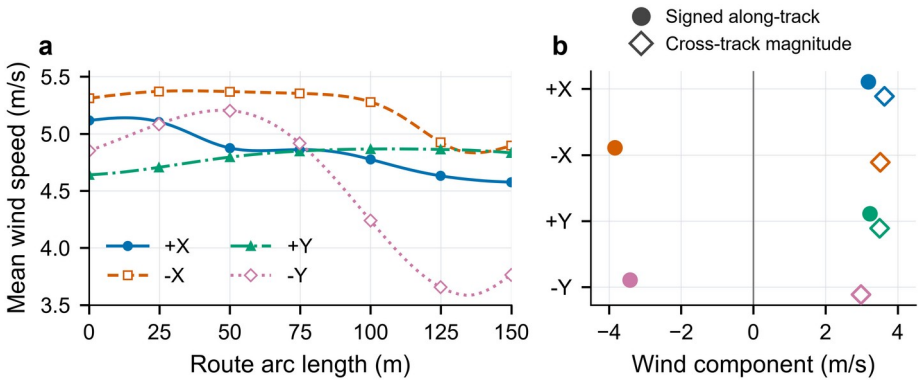


Fig. 4. Reference-route wind at 100 m: (a) unsmoothed speed, with genuine samples near 25-m stations marked; (b) mean signed along-track and unsigned horizontal cross-track components.

Wind information reduces route-active RMSE by 63.1–71.3% (Fig. 5, Table 2). Spatial-minus-persistence differences are only −0.247 to +0.103 mm: compensating current local wind provides most of the improvement.

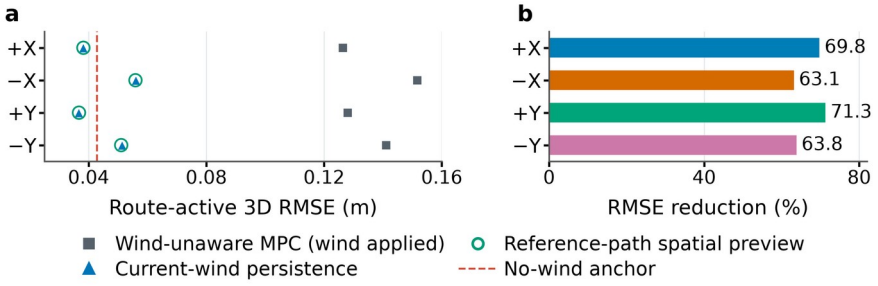


Fig. 5. Route-active tracking for four wind fields. (a) 3D RMSE of the three information modes. (b) Spatial-preview reduction relative to wind-unaware MPC.

Table 2. Route-active 3D RMSE (m). None: wind-unaware; Current: persistence; Spatial: reference-path preview. Gain is relative to None; Δ is Spatial minus Current (mm).

Flow	None	Current	Spatial	Gain	Δ (mm)
+X	0.12638	0.03812	0.03822	69.76%	+0.103 mm
-X	0.15170	0.05600	0.05591	63.14%	-0.088 mm
+Y	0.12805	0.03665	0.03673	71.32%	+0.082 mm
-Y	0.14114	0.05131	0.05107	63.82%	-0.247 mm

5 Discussion and conclusion

The forecast differences in Eq. (9) explain the near-equivalence: mean wind differences are 0.0125–0.0859 m/s (P95: 0.0265–0.2143 m/s), attenuated to 0.00083–0.00503 m/s² by B_w. Spatial variation over the 7.78-m mean horizon-end separation is modest. Persistence and spatial preview remain closely matched across 0.5–3-s horizons. Regional and route-specific rankings highlight different aspects of the wind field: +Y leads the largest number of regional configurations, whereas −X produces the strongest opposing flow along the selected trajectory. This difference shows why route-level exposure is valuable alongside regional CFD metrics when evaluating tracking performance.

Two wake speeds change by 17.4% and 49.4% on refinement, showing separated-flow mesh sensitivity. The separate-vehicle B_w map and hypothetical gain variations do not identify M100 aerodynamics, although wind-informed MPC retains lower error at the tested gains. Future work should address targeted wake LES, vehicle-specific wind response, and longer urban missions with planning and onboard obstacle sensing.

The study links directional CFD, route exposure and predictive control. Current-wind compensation supplies most of the tracking benefit; additional spatial information has little effect over the tested short horizons.

Acknowledgment. The authors recognize the High-Performance Computing Center at Oklahoma State University for providing access to the Pete supercomputer, supported in part by NSF grant OAC-1531128.

References

1. Blocken, B.: 50 years of Computational Wind Engineering: Past, present and future. J. Wind Eng. Ind. Aerodyn. 129, 69–102 (2014). doi:10.1016/j.jweia.2014.03.008
2. Tominaga, Y., et al.: AIJ guidelines for practical applications of CFD to pedestrian wind environment around buildings. J. Wind Eng. Ind. Aerodyn. 96(10-11), 1749–1761 (2008). doi:10.1016/j.jweia.2008.02.058
3. Franke, J., et al. (eds.): Best Practice Guideline for the CFD Simulation of Flows in the Urban Environment. COST Office (2007).
4. Blocken, B.: Computational Fluid Dynamics for urban physics: Importance, scales, possibilities, limitations and ten tips and tricks towards accurate and reliable simulations. Build. Environ. 91, 219–245 (2015). doi:10.1016/j.buildenv.2015.02.015
5. Vuppala, R.K.S.S., et al.: Modeling advanced air mobility aircraft in data-driven reduced order realistic urban winds. Sci. Rep. 14(1), 383 (2024). doi:10.1038/s41598-023-50719-8

6. Ebert, C., et al.: Trajectory Planning in Windy Urban Environment Using Gappy Proper Orthogonal Decomposition for Wind Estimates. AIAA J. 61(6), 2640–2651 (2023). doi:10.2514/1.j062049
7. Tabassum, A., et al.: Variance Reduction of Quadcopter Trajectory Tracking in Turbulent Wind. IFAC-PapersOnLine 54(20), 102–107 (2021). doi:10.1016/j.ifacol.2021.11.160
8. Chan, Y.Y., et al.: Wind dynamic and energy-efficiency path planning for unmanned aerial vehicles in the lower-level airspace and urban air mobility context. Sustain. Energy Technol. Assess. 57, 103202 (2023). doi:10.1016/j.seta.2023.103202
9. Wang, J., et al.: Evaluation of urban wind effects on flight path planning of delivery drones using computational fluid dynamics simulations. Phys. Fluids 37(8), 087124 (2025). doi:10.1063/5.0281373
10. Krawczyk, Z., et al.: Evaluating Reduced-Order Urban Wind Models for Simulating Flight Dynamics of Advanced Aerial Mobility Aircraft. Aerospace 11(10), 830 (2024). doi:10.3390/aerospace11100830
11. Wu, J., Ye, Y., Du, J.: Multi-objective reinforcement learning for autonomous drone navigation in urban areas with wind zones. Autom. Constr. 158, 105253 (2024). doi:10.1016/j.autcon.2023.105253
12. Kakavitsas, N., et al.: Quadrotor Flight Simulation in a CFD-generated Urban Wind Field. Proc. IEEE Aerosp. Conf., 1–8 (2024). doi:10.1109/aero58975.2024.10521032
13. Weller, H.G., et al.: A tensorial approach to computational continuum mechanics using object-oriented techniques. Comput. Phys. 12(6), 620–631 (1998). doi:10.1063/1.168744
14. Launder, B.E., Spalding, D.B.: The numerical computation of turbulent flows. Comput. Methods Appl. Mech. Eng. 3(2), 269–289 (1974). doi:10.1016/0045-7825(74)90029-2
15. Richards, P.J., Hoxey, R.P.: Appropriate boundary conditions for computational wind engineering models using the k–ε turbulence model. J. Wind Eng. Ind. Aerodyn. 46-47, 145–153 (1993). doi:10.1016/0167-6105(93)90124-7
16. Sa, I., Kamel, M., Khanna, R., Popovic, M., Nieto, J., Siegwart, R.: Dynamic System Identification, and Control for a Cost-Effective and Open-Source Multi-rotor MAV. Field and Service Robotics, pp. 605–620 (2018). https://doi.org/10.1007/978-3-319-67361-5_39
17. Mayne, D.Q., et al.: Constrained model predictive control: Stability and optimality. Automatica 36(6), 789–814 (2000). doi:10.1016/s0005-1098(99)00214-9
18. Rawlings, J.B., Mayne, D.Q., Diehl, M.M.: Model Predictive Control: Theory, Computation, and Design. Nob Hill Publishing, 2nd edn. (2017).
19. Tominaga, Y., Kikumoto, H., Okaze, T.: AIJ UWE Benchmark Dataset – Case D (Highrise): Wind Tunnel Experiment for Flow around a High-rise Building in City Blocks. Zenodo (2025). doi:10.5281/zenodo.15402669
20. Mayne, D.Q., Seron, M.M., Rakovic, S.V.: Robust model predictive control of constrained linear systems with bounded disturbances. Automatica 41(2), 219–224 (2005). doi:10.1016/j.automatica.2004.08.019

Trustworthy Geometry-to-Polar Learning for Unseen Airfoils: OOD Generalization, Uncertainty, and Physics-Aware Surrogates

Mohammed Alharbi[1], Ahmed Ghareeb[2], Nuha Alharbi[3]

[1] Department of Mechanical Engineering, College of Engineering and Architecture, Umm Al-Qura University, Makkah, Saudi Arabia
[2]Department of Mechanical Engineering, College of Engineering, University of Kirkuk, Kirkuk, Iraq
[3]Department of Computer Science, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

Abstract. Aerodynamic surrogate models are often evaluated by randomly splitting angle-of-attack samples, although practical deployment requires generalization to airfoils unseen during training. We audit a public fixed-Reynolds-number airfoil dataset and find substantial redundancy and inconsistency: 33,705 aerodynamic states and 2,946 filenames correspond to only 1,928 unique CST representations, while 7,061 identical geometry–angle-of-attack inputs have conflicting labels. These inconsistencies impose RMSE lower bounds of 0.0387 for lift and 0.0067 for drag. We establish six frozen evaluation protocols covering repository splits, matched-row interpolation, unseen filenames, unseen CST representations, family shifts, and geometry-based distribution shifts. A compact query-conditioned geometry-to-polar surrogate predicts the full 13-angle polar with missing-label masking and is compared with the released MLP and XGBoost. On the primary CST-disjoint test, the model achieves a five-seed lift RMSE of 0.0756 (95% interval: 0.0752–0.0760), a 5.0% improvement over XGBoost, while drag RMSE remains 0.0139. Weak structural regularization reduces lift and log-drag polar roughness by 13.0% and 18.2%, respectively, without improving prediction accuracy. Split-conformal intervals achieve 89.9% lift and 88.2% drag coverage at the nominal 90% level, but geometry-level intervals are too wide for practical use and OOD distance is only weakly related to error. Geometry shift increases mean lift RMSE to 0.1012. Despite this degradation, complete-polar L/D ranking reaches Spearman ρ=0.889 with 80.0% top-five overlap. These results highlight the need for grouped evaluation, show improved lift transfer through polar learning, and identify drag prediction and uncertainty estimation as key remaining challenges.

Keywords: airfoil surrogate modeling · out-of-distribution generalization · geometry-to-polar learning · conformal prediction · physics-aware learning

1 Introduction

Aerodynamic surrogate models can make airfoil screening and optimization substantially faster by estimating lift and drag without repeatedly solving the flow equations. Their usefulness depends, however, on what the test data actually measure. A model may perform well on held-out angles of attack yet fail when asked to predict the response of an airfoil absent from training. When different splits contain observations from the same geometry—or distinct filenames that share the same published features—a random row split mainly tests interpolation rather than transfer. Such leakage is a recognized problem in scientific machine learning [1].

A second issue concerns how aerodynamic response is represented. Each geometry has a lift and drag curve over angle of attack, but pointwise regression treats the observations at individual angles as separate prediction problems. Inspired by operator-learning approaches [2, 3], we instead consider a model that predicts an entire fixed-angle polar from a geometry representation. Evaluating this approach requires more than reporting average prediction error: ambiguity in the source data, geometric distribution shift, and the usefulness of uncertainty estimates also matter [4–6].

We make four connected contributions. First, we audit collisions in the released geometry features and quantify conflicting targets. Second, we define six frozen evaluation protocols (S0–S5), with an exact-representation-disjoint split as the principal benchmark. Third, we develop a mask-aware, query-conditioned geometry-to-polar surrogate. Finally, we examine how physics-guided smoothing, uncertainty calibration, and out-of-distribution (OOD) indicators affect the model's reliability, with particular attention to drag. We also test whether its predicted polars preserve rankings relevant to preliminary airfoil selection.

2 Dataset and Representation Audit

Our experiments use the CFD-derived airfoil dataset published by Agarwal et al. [7]. The dataset was generated using the OpenFOAM simpleFoam solver with the Spalart–Allmaras turbulence model at a fixed Reynolds number of 10^5. Each record contains an airfoil filename, angle of attack α (integer values from $-4°$ to $8°$), eight class/shape transformation (CST) coefficients, and lift and drag coefficients (C_l and C_d). Following the source implementation of Kulfan's parameterization [8], four CST coefficients describe each surface. Reconstructions of 100 sampled profiles differed from the released processed profiles by at most 6.61×10^{-9} chord in ordinate. The targets are simulation-derived rather than experimental measurements.

The data contain 33,705 records from 2,946 filenames. Only 1,147 filenames have complete 13-angle polars; the other 1,799 yield 4,593 missing filename–angle states, which we mask rather than impute. Hashing the eight published CST values yields only 1,928 distinct signatures. In all, 1,001 signatures occur under multiple filenames, affecting 2,019 names. We describe these as representation collisions—not necessarily duplicate airfoils—because the available coefficients do not establish that

the underlying coordinate profiles are identical. Fig. 1(a,b) summarizes the polar completeness and feature collisions.

The released mapping is also non-single-valued: 7,061 identical CST–angle inputs are associated with conflicting lift or drag targets (Fig. 1c). For each group g of identical inputs, we compute the mean target value as $\bar{y}_g = n_g^{-1} \sum_{i=1}^{ng} y_{gi}$, where n_g is the number of observations in the group. Summing the squared deviations $(y_{gi} - \bar{y}_g)^2$ across all rows, dividing by the total number of observations, and taking the square root yields deterministic input-conditional RMSE floors of 0.03869 for lift and 0.006659 for drag. These floors reflect ambiguity in the published features and labels rather than universal aerodynamic noise.

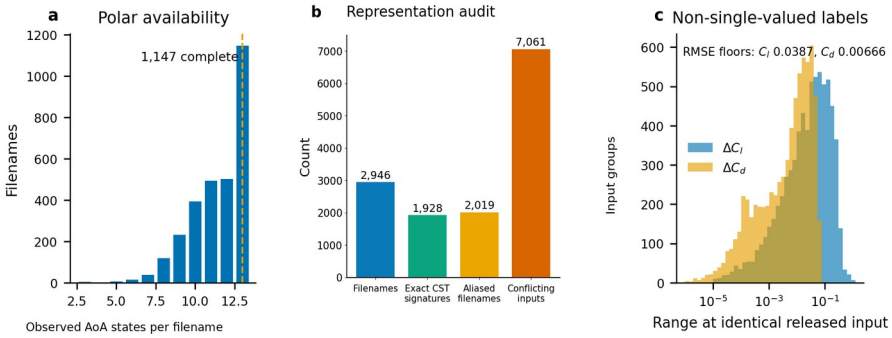


Fig. 1. Dataset audit: (a) number of observed angles per filename, with 1,147 complete 13-angle polars; (b) published identifiers, distinct CST signatures, aliased filenames, and conflicting inputs; (c) target-value ranges for identical CST–angle inputs.

3 Evaluation Protocol and Models

3.1 Grouped generalization splits

The source repository's S0 partition splits rows rather than geometries: all 1,291 test filenames also occur in training, along with 1,082 exact CST signatures and 805 identical CST–angle inputs. We therefore freeze six protocols before final testing. S0 reproduces the original repository split; S1 uses a 70/10/10/10 random-row split for training, validation, calibration, and testing; S2 separates filenames; S3 separates exact eight-CST signatures; S4 holds out five predefined naming families (GOE, Eppler E, NACA, Wortmann FX, and Martin–Hepperle MH); and S5 tests five geometry-descriptor extremes or distant cases. The primary unseen-representation benchmark is S3. Unlike S3–S5, S2 still has 153 exact CST signatures crossing the train–test boundary.

All input and target transformations are fitted to training data alone. We use validation groups to choose models and loss settings, reserve calibration groups for uncertainty quantiles, and evaluate the test groups only after these decisions have been fixed. Experiments use seeds 17, 42, 101, 202, and 314. We report Student-t intervals across seeds to characterize training variability and use 2,000 group-

bootstrap replicates over S3 signatures to assess sampling variability. Paired comparisons use per-signature error, Wilcoxon signed-rank tests, and Holm adjustment.

3.2 Geometry-to-polar prediction

We compare three models. M0 reproduces the source multilayer perceptron (9–12–6–2 units, batch normalization, ReLU, and dropout). M1 uses separate XGBoost regressors [9] for lift and drag, with the eight-component geometry vector g and angle α as inputs; validation selects direct lift and logarithmic drag targets. M2 instead predicts the complete 13-angle response as $G_\theta(g) = \{ [\hat{C}_l(\alpha_j), \hat{C}_d(\alpha_j)] \}$ for j = 1,…,13. An 8–64–64 SiLU geometry encoder and a 1–32–32 angle encoder feed a shared 96–96–2 decoder with learned elementwise interactions. This query-conditioned architecture is inspired by operator learning [2], without claiming a general-purpose neural operator.

A binary observation mask m_{ij} excludes missing filename–angle targets from the normalized smooth-L1 training loss, although M2 still predicts all 13 angles. Based on the prespecified validation score combining normalized lift/drag RMSE and polar-shape error, M2 uses direct drag (0.56328 versus 0.57446 for log drag), whereas M1 favors log drag. Figure 2 illustrates the model architecture and where observation masking and polar regularization enter the workflow.

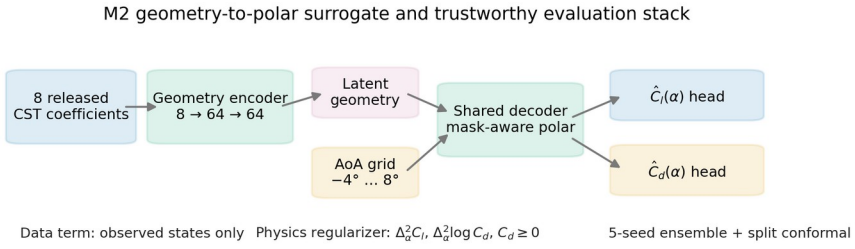


Fig. 2. Query-conditioned M2 geometry-to-polar surrogate. Missing targets are masked in training; all angles are predicted at inference.

3.3 Physics-guided learning and uncertainty

We also examine whether weak response-shape constraints improve the predicted polars. The regularizer penalizes squared second differences of lift and standardized drag across adjacent angles, while a soft positivity penalty discourages drag predictions below 10^{-5}. The total loss is $L = L_{data} + \lambda_s L_{smooth} + \lambda_p L_{pos}$, where the three terms denote masked prediction error, smoothness, and drag positivity. Validation selects $(\lambda_s, \lambda_p) = (0.1, 1)$ for the weakly regularized comparison; $\lambda_s = 1$ tests stronger smoothing. These priors do not evaluate governing-equation residuals, so the approach is physics-aware rather than a Navier–Stokes PINN [10].

To examine predictive uncertainty, we train five independently initialized M2 models and calculate their ensemble mean and sample variance. We construct 90% and 95% marginal split-conformal intervals [6] from S3 calibration residuals. For simultaneous polar intervals, a geometry-level score instead takes the maximum standardized residual over all observed angles and both targets. Finally, we ask whether distance from the training data identifies poor predictions. The tested indicators include CST Mahalanobis distance, nearest-neighbor distance in CST and reconstructed-descriptor spaces, geometry-latent distance, and ensemble spread.

4 Results

4.1 Predictive performance and leakage

On XGBoost (M1), moving from the original row split (S0) to the exact-CST-disjoint test (S3) increases lift RMSE from 0.0730 to 0.0795 (9.0%) and drag RMSE from 0.0126 to 0.0139 (approximately 10%). This is a modest change in error but a crucial change in interpretation: S0 does not test unseen geometry representations. Table 1 reports the S3 comparison. M2 reduces lift RMSE from 0.07953 (M1) to 0.07559 (5.0%); a separate paired test of per-signature lift MAE is significant after Holm adjustment (adjusted $p = 1.45 \times 10^{-7}$). Drag RMSE remains essentially unchanged (0.01390 for M2 versus 0.01387 for M1). Thus, sharing information across angles improves lift transfer but not drag RMSE.

Table 1. Primary S3 test RMSE, mean across five deterministic seeds (95% Student-t intervals).

Model	Lift RMSE	Lift 95% CI	Drag RMSE	Drag 95% CI
M0: source MLP	0.08357	0.08176–0.08537	0.01625	0.01487–0.01763
M1: XGBoost	0.07953	0.07927–0.07979	0.01387	0.01384–0.01390
M2: geometry-to-polar	0.07559	0.07523–0.07595	0.01390	0.01387–0.01394
M2:physics-aware	0.07557	0.07514–0.07599	0.01391	0.01387–0.01394

4.2 Physics-guided ablation

Weak smoothness and positivity penalties reduce lift and log-drag second-difference roughness by 13.0% and 18.2%, respectively, while lift and drag RMSE stay near 0.0756 and 0.0139 (Table 1). Stronger smoothing reduces log-drag roughness by 51.5% but yields no established accuracy advantage. None of the evaluated M2 variants predicts negative drag. The supported conclusion is smoother polars, not

more accurate aerodynamic coefficients. The log-drag and strong-smoothing variants use only one fixed seed and are interpreted descriptively.

4.3 Calibration, shift, and drag

On the S3 test, nominal 90% marginal split-conformal intervals cover 89.9% of observed lift states and 88.2% of drag states, with mean full widths of 0.192 and 0.041, respectively. At nominal 95%, coverage is 94.8% for lift and 94.1% for drag. Simultaneous geometry-level bands, defined to cover every observed angle and both outputs within each geometry, achieve 89.1% coverage at nominal 90%, but their full widths are 3.0606 for lift and 0.1093 for drag. The distinction between state-wise and full-polar intervals is therefore important: coverage near nominal does not necessarily produce useful decision intervals. These S3 findings do not guarantee validity for arbitrary OOD shifts.

Distance-based OOD indicators are weak error predictors. On S3, nearest-neighbor distance in CST space correlates with per-signature error at only Spearman $\rho = 0.104$ (lift) and $\rho = 0.074$ (drag). Mahalanobis, descriptor, latent-distance, and ensemble-spread measures show similarly weak relationships, and rejecting high-ensemble-risk signatures does not consistently reduce error. A reported risk score should therefore not be treated as a dependable failure detector.

Table 2 places these uncertainty and shift results alongside the primary S3 performance. Across five S4 family holdouts, mean lift/drag RMSE is 0.0779/0.0134. Five descriptor-defined S5 cases are harder for lift (mean 0.1012; worst 0.1166), while drag RMSE averages 0.0140 (worst 0.0166). On S3, the M2 ensemble attains $R^2 = 0.9638$ for lift and only 0.1846 for drag. Absolute errors peak at $\alpha = 8°$ for both outputs; drag-relative error also increases with observed drag ($\rho = 0.377$). These observations are consistent with several interacting limitations—label ambiguity, limited drag dynamic range, higher-angle complexity, and the reduced geometry representation—rather than a single identifiable OOD threshold.

Table 2. Summary of S3 results and averages over the five S4 and five S5 holdout cases. S3 coverage and correlations are evaluated on unseen CST signatures.

Evaluation	Lift	Drag / other outcome
S3 M2 RMSE	0.0756	0.0139
S4 family OOD mean RMSE	0.0779	0.0134
S5 descriptor OOD mean RMSE	0.1012	0.0140
S3 marginal 90% coverage	89.9%	88.2%
S3 distance/error Spearman ρ	0.104	0.074

4.4 Downstream polar-ranking utility

To test downstream usefulness, we rank held-out airfoils by the maximum lift-to-drag ratio obtained from their predicted full polars. For the 109 test filenames with all 13 measured angles, the M2 ensemble achieves Spearman $\rho = 0.889$ against the observed

ranking and recovers four of the five highest-ranked airfoils (80%). Normalized selection regret is 0.013, mean operating-angle regret is 0.069, and a dimensionless drag-proportional power proxy has MAE 0.266. The ranking result supports preliminary airfoil screening, but the proxy is not a complete propulsion, battery, or mission-energy model.

5 Discussion and Conclusion

The central methodological lesson is that the unit of evaluation must match the geometry representation used by the learner. Here, filename-based splitting can conceal overlap because different filenames share the same eight-CST vector. Retaining conflicting source labels, rather than averaging them away, also exposes the prediction limit imposed by those published inputs. More informative future datasets would preserve canonical coordinates, richer geometry parameters, and simulation provenance so that representation aliases can be distinguished from genuinely different cases.

Learning an entire polar provides a repeatable improvement in lift over pointwise regression, but the benefits do not extend to drag RMSE. Likewise, response-shape penalties make the curves smoother without showing that they are closer to the source labels. Marginal intervals achieve coverage near their nominal levels on S3, but simultaneous bands are too broad and the tested OOD indicators do not reliably rank errors. In practice, results should be reported on representation-disjoint data, geometric extremes should be flagged explicitly, and predictions should be treated cautiously when drag error or interval width could change an engineering decision.

These conclusions apply only to the available eight-CST inputs, one CFD data source at $Re = 10^5$, and the $-4°$ to $8°$ angle range. Missing polar states may not be random, family assignments inferred from filenames are incomplete, and the experiments cannot support claims about post-stall behavior, different Reynolds or Mach numbers, three-dimensional wings, or experimental measurements. Within the evaluated setting, M2 improves lift RMSE by 5.0% relative to XGBoost and preserves useful lift-to-drag rankings. Accurate drag transfer and informative uncertainty under geometric shift remain unresolved.

References

[1] S. Kapoor and A. Narayanan, "Leakage and the Reproducibility Crisis in Machine-Learning-Based Science," Patterns, vol. 4, no. 9, p. 100804, 2023. doi: 10.1016/j.patter.2023.100804.

[2] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis, "Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators," Nature Machine Intelligence, vol. 3, pp. 218-229, 2021. doi: 10.1038/s42256-021-00302-5.

[3] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, and A. Anandkumar, "Neural Operator: Learning Maps Between Function Spaces with

Applications to PDEs," Journal of Machine Learning Research, vol. 24, no. 89, pp. 1-97, 2023.

[4] F. Bonnet, J. Mazari, P. Cinnella, and P. Gallinari, "AirfRANS: High Fidelity Computational Fluid Dynamics Dataset for Approximating Reynolds-Averaged Navier-Stokes Solutions," in Advances in Neural Information Processing Systems, vol. 35, 2022. doi: 10.52202/068431-1705.

[5] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles," in Advances in Neural Information Processing Systems, vol. 30, 2017.

[6] J. Lei, M. G'Sell, A. Rinaldo, R. J. Tibshirani, and L. Wasserman, "Distribution-Free Predictive Inference for Regression," Journal of the American Statistical Association, vol. 113, no. 523, pp. 1094-1111, 2018. doi: 10.1080/01621459.2017.1307116.

[7] K. Agarwal, V. Vijaykrishnan, D. Mohanty, and M. Murugaiah, "A Comprehensive Dataset of the Aerodynamic and Geometric Coefficients of Airfoils in the Public Domain," Data, vol. 9, no. 5, p. 64, 2024. doi: 10.3390/data9050064.

[8] B. M. Kulfan, "Universal Parametric Geometry Representation Method," Journal of Aircraft, vol. 45, no. 1, pp. 142-158, 2008. doi: 10.2514/1.29958.

[9] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785-794, 2016. doi: 10.1145/2939672.2939785.

[10] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," Nature Reviews Physics, vol. 3, pp. 422-440, 2021. doi: 10.1038/s42254-021-00314-5.

Data-Driven Grid-Aware Coordination of Residential PV–Battery Systems under Intermittent Electricity Supply

Abdulmajeed Mustafa[1], Fatimah Mustafa[1], and Ahmed Ghareeb[2]

[1] Sun Engineering
[2] College of Engineering, University of Kirkuk

Abstract. Residential photovoltaic (PV) systems with battery storage are becoming increasingly important for enhancing energy flexibility and meeting local electricity needs. A significant operating condition occurs when grid supply returns after an interruption: household loads reconnect, and depleted batteries may begin charging simultaneously, resulting in a short-term peak in demand. This study develops a data-driven, grid-aware framework for coordinating residential PV–battery recovery under intermittent electricity supply. Public weather data for Kirkuk, modeled residential demand, and an IEEE 123-node benchmark feeder are combined with short-term forecasting, receding-horizon scheduling, and network-aware correction using voltage sensitivities and AC power-flow feedback. At full adoption with 3 kW chargers, coordination reduces the post-restoration peak from 6.118 to 3.325 MW (45.65%) while preserving 2.917 MWh of recharge energy and restoring all batteries to 90% state of charge. In a 4.5 kW stress case, the network-aware correction reduces the maximum voltage from 1.05123 to 1.04992 pu and removes all three overvoltage intervals without changing the recovery target. Placement, outage-timing, and forecast-ablation experiments also demonstrate that the peak-reduction benefit is robust to spatial placement and that event timing and recovery headroom are more important than forecast sophistication. The findings support coordinated, network-validated battery recovery under Iraq-relevant distribution-network conditions.

Keywords: Residential storage · Grid restoration · Network-aware charging · Photovoltaics · Distribution networks

1 Introduction

Residential photovoltaic (PV)–battery systems provide flexibility and backup. The dispatch paths that can benefit customer savings, self-consumption, and reserve requirements can differ [1, 2], as can the impacts on the network depending on when many systems charge and discharge.

When the power goes off, batteries provide power for the home and discharge. Once supply is restored, normal loads resume operation, and depleted batteries may need to

be charged at similar times. The impacts of interruptions can be reduced with behind-the-meter PV and storage, but there can also be a controllable demand surge during recovery [3]. This can occur in conjunction with, or separately from, thermal cold-load pickup (CLPU). Diversity and temperature dependence during restoration events are highlighted by the results of smart-meter analyses, multi-state models, and measured restoration events [4, 5, 6]. Battery recovery, by contrast, has a finite energy requirement and a controllable recharge time.

The practice of residential storage coordination is well developed. Distributed control supports voltage control, community scheduling is used to accommodate grid constraints and reserves, and domestic schemes are used to coordinate other flexible resources [7, 8, 9]. Household goals do not necessarily imply network benefits, as these studies also show.

Network-aware approaches add active/reactive control, multi-period scheduling, and unbalanced optimization [10, 11, 12, 13]. However, network-constrained optimization itself is not the contribution here; the focus is energy-complete battery recovery after an interruption under temporal and locational constraints. Better forecasts do not necessarily lead to better control, although forecasts can predict recovery demand. Studies of XGBoost and behind-the-meter PV forecasting [14, 15] suggest a straightforward comparison of persistence, XGBoost, and perfect-information inputs.

Published Kirkuk utility analysis [16] and public weather data are used to provide the Iraq-focused context. The case study does not recreate the network of the city of Kirkuk or represent measured Iraqi household operation, but is based on an IEEE benchmark feeder and modeled household operation.

We study synchronized, energy-complete recovery under Iraq-relevant assumptions, focusing on the gap between aggregate peak reduction and locational voltage feasibility.

This study proposes a reproducible recovery case, a voltage-sensitivity and AC-feedback allocation layer for an aggregate scheduler, and robustness tests for the placement of PV and batteries, outage events, and forecast quality. The outage service, initial conditions, recovery energy, and deadlines are the same for conventional, aggregate, and network-aware policies. The overall workflow is summarized in Figure 1.

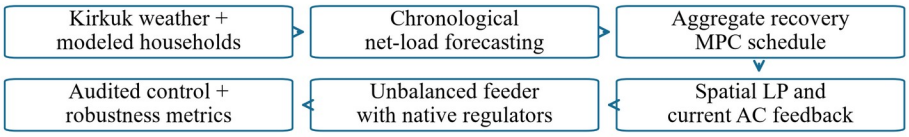


Fig. 1. Temporal scheduling with current-state locational correction.

Table 1. Compact nomenclature; dimensionless quantities use unit 1.

Symbol	Definition	Unit	Symbol	Definition	Unit
$L_{i,t}$	Household demand	kW	η_{dis}	Discharge eff.	1
$G_{i,t}$	PV output	kW	Δt	Interval duration	h
$C_{i,t}$	Charging power	kW	z	Peak epigraph	kW
$D_{i,t}$	Discharging power	kW	P_{lim}	Import limit	kW
$P_{grid,i,t}$	Grid import	kW	R_{loss}	Loss allowance	kW
E_i	Battery capacity	kWh	PRPR	Peak ratio	1
$s_{i,t}$	State of charge	1	H_P	Voltage sensitivity	pu/kW
η_{ch}	Charge efficiency	1	$\hat{B}_t$	Net-load forecast	kW

2 Data and Methodology

2.1 Case Data and Residential Model

NASA POWER supplies 2,208 complete hourly records for June–August 2024 at 35.4681° N, 44.3922° E. Temperature and horizontal, direct, and diffuse irradiation drive PV modeling. Hourly Wh/m² is converted to mean W/m² and held constant over quarter-hours; no measured sub-hourly variability is implied. Times are in Asia/Baghdad (UTC+3).

The IEEE 123-node feeder retains its 4.16/0.48 kV parameters. Its 91 loads represent 875 equivalent households, rounded from cluster kW/4. Modeled demand includes morning/evening peaks, cooling above 28 °C, and seeded variation. PV uses pvlib geometry, a south-facing 30° plane, and temperature derating. Table 2 records provenance; the configuration supplies all coefficients.

Table 2. Main inputs and parameter provenance.

Parameter	Value	Provenance
Weather/feeder	Kirkuk, Jun–Aug 2024 / IEEE 123-node	Public/benchmark
Fleet/PV/BESS	875 equivalent / 5 kWp / 10 kWh	Modeled/assumed
Power/efficiency	3 kW (stress: 4.5) / 0.95 each way	Assumed
SOC/recovery	0.20–0.90 / 0.90 target; six hours	Assumed
Main interruption	15 August, 16:00–19:00 UTC+3	Assumed
Import/reserve	4.5 MW / 0.2 MW	Policy assumptions
Voltage/head	0.95–1.05 pu / 5 MVA	Band/rating

Grid import is positive. With Table 1's symbols and $\Delta t = 0.25$ h,

$$P_{grid,i,t} = L_{i,t} + C_{i,t} - D_{i,t} - G_{i,t} \quad (1)$$

$$s_{i,t+1} = s_{i,t} + (\eta_{ch} C_{i,t} - D_{i,t}/\eta_{dis}) \Delta t/E_i \quad (2)$$

Capacity/power scale with adopters; $0.20 \le s \le 0.90$ and $0 \le C,D \le P_{max}$. Outage backup serves adopters' 55% critical demand, retaining unserved load and PV curtailment. Nonadopters have no backup; initial SOC is 0.90. Thermal CLPU, diesel, switching and degradation are excluded.

2.2 Temporal Recovery and Forecast Inputs

Conventional recovery charges at the power limit to 90% SOC. Coordinated policies preserve load, PV, outage support, initial conditions, recovery energy, and deadline; only recovery charging changes. Recovery discharge is disabled.

Persistence uses yesterday's values. XGBoost uses calendar features and 24/48-hour load/PV/temperature lags: 180 trees, depth 3, learning rate 0.05. Training precedes 1 August; holdout ends 1 September 02:45 local. Six-hour forecast inputs predate each origin. Future realized weather is excluded except in the labeled perfect-information ablation.

At each step τ, ideal current metering replaces the first net-demand forecast $\hat{B}$. The remaining-horizon LP minimizes peak and aggregate charge variation:

$$\min z + \lambda \Sigma_t u_t, \quad \lambda = 0.001 \tag{3}$$

$$\widehat{B}_t + \Sigma_i C_{i,t} \le z, \quad \widehat{B}_t + \Sigma_i C_{i,t} \le P_{\text{lim}} - R_{loss} \tag{4}$$

$$\Sigma_t \eta_{ch} C_{i,t} \Delta t = (s_{\text{target}} - s_{i,\tau}) E_i, \quad 0 \le C_{i,t} \le P_{\text{max},i} \tag{5}$$

Ramp bounds impose $u_t \ge |\Sigma_i C_{i,t} - \Sigma_i C_{i,t-1}|$. The assumed policy is Plim = 4.5 MW with Rloss = 0.2 MW. Positive charging and exact per-cluster terminal energy preserve SOC bounds. SciPy/HiGHS executes the first action and replans; solver status/objective/runtime are logged.

2.3 Locational Allocation with AC Feedback

At the current AC point (starred quantities), +1 kW perturbations at adopting clusters estimate the voltage sensitivity with regulator taps frozen:

$$V_\tau \approx V_{\tau^*} + H_{P,\tau}(C_\tau - C_{\tau^*}) \tag{6}$$

If voltage or head-rating limits are violated, the allocation LP retains $\Sigma_i C_{i,t}(\text{loc}) = \Sigma_i C_{i,t}(\text{temporal})$ for every remaining slot and each cluster's remaining-energy equality and charging limits. Nonnegative charging and the terminal target preserve SOC bounds. Voltage limits are 0.95–1.05 pu with an initial 8×10^{-5} pu margin; sixteen P/Q inner-polygon constraints conservatively approximate 5 MVA.

Each retry recomputes voltage and head P/Q sensitivities at the updated candidate, settling taps before freezing them for perturbations. Native controls are restored for AC checks; LP equalities and bounds are retained. Only an AC-feasible first action is executed, then SOC is updated and MPC replans. LP failure or six-iteration exhaustion raises an error without execution or recovery relaxation. Current measurements, not future realized states, are used; independent AC replay checks execution.

2.4 Experiments, Metrics and Audits

The principal interruption is from 16:00 to 19:00 on 15 August, with a recovery period of 6 hours. Nested 0/25/50/75/100% adoption gives 0/219/438/656/875 households. Chargers are rated at 3 kW (4.5 kW is used as a voltage-stress case). At each level of partial adoption, 10 random rankings are assessed; full adoption is deterministic.

Nine event cases combine 1/3/5-hour outages and 17:00/19:00/21:00 restoration times, with a 6-hour recovery period. Forecast ablation is performed using persistence/XGBoost/perfect information at both charger ratings. OpenDSS solves settled unbalanced snapshots with native controls and constant-power setpoints, while island service remains an energy model.

At Transformer.reg1a, Phead and Qhead sum first-terminal phase powers. Head loading (%) is $100\sqrt{(Phead^2 + Qhead^2)}/Srated$, with Srated = 5 MVA. Reactive power explains why source MW peaks differ from apparent-power loading. Final voltage and loading metrics use nonlinear OpenDSS AC results with native regulators restored. PRPR uses the same case without recharge; reduction is $100 \times (Pbase - Pcoord)/Pbase$.

3 Results and Discussion

3.1 Recovery Peaks and Energy Fairness

The modeled IEEE 123-node case uses modeled demand and public NASA POWER Kirkuk weather for a specified synchronized outage/restoration. Placement and nine event tests assess simulation robustness, not measured Iraqi utility performance. The primary interruption reduces the mean SOC of the clusters from 90% to 58.33%. With full adoption and 3 kW chargers, conventional peak demand is 6.118 MW, while it is 3.325 MW under coordination, a reduction of 45.65%. In both coordinated policies, the same 2.917 MWh of AC charging energy is recovered, and all clusters reach 90% SOC. Head loading falls from 125.86% to 70.26%; minimum voltage improves from 0.9651 to 0.9784 pu.

Conventional peaks grow approximately with the number of adopters because charger ratings result in coincident charging requests (Figs. 2–3). Coordination moves charging into periods of declining household demand. The main deficit requires an average of 0.486 MW over a 6-hour period, fitting within later headroom without exceeding the no-recharge peak; PRPR falls from 1.840 to 1.000. As the charging current becomes less concentrated, recovery losses fall from 459.24 to 382.13 kWh.

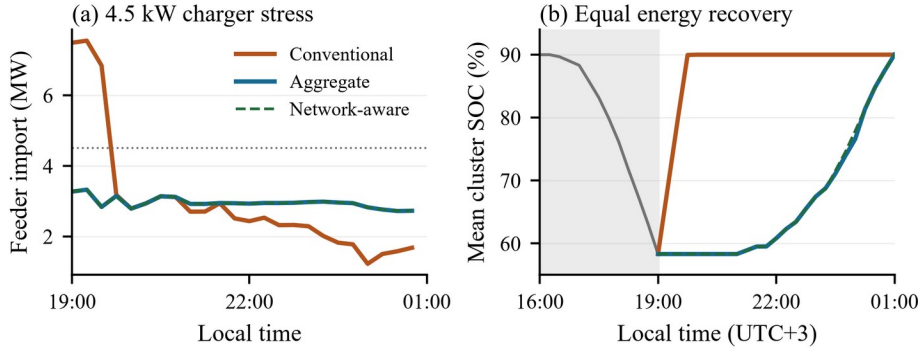


Fig. 2. 4.5 kW recovery: feeder import and mean SOC; shading denotes the outage.

3.2 Why Locational Correction Matters

The main 3 kW aggregate schedules already meet the voltage band, so locational correction leaves them essentially unchanged (Table 3). With 4.5 kW chargers, conventional charging reaches 7.550 MW and a maximum voltage of 1.06249 pu. With aggregate scheduling, the peak drops to 3.325 MW, but three overvoltage intervals remain, with a maximum voltage of 1.05123 pu.

Table 3. Conventional / aggregate / network-aware (C/A/N); all reach 90% SOC.

Adoption / charger kW	Peak (MW) C / A / N	Maximum voltage (pu) C / A / N
25% / 3	4.015 / 3.325 / 3.325	1.0507 / 1.0495 / 1.0495
50% / 3	4.712 / 3.325 / 3.325	1.0511 / 1.0495 / 1.0495
75% / 3	5.411 / 3.325 / 3.325	1.0512 / 1.0499 / 1.0499
100% / 3	6.118 / 3.325 / 3.325	1.0500 / 1.0498 / 1.0498
100% / 4.5	7.550 / 3.325 / 3.325	1.0625 / 1.0512 / 1.0499

One corrective LP removes all three violations and results in a maximum voltage of 1.04992 pu, with unchanged peak, recovery energy, and terminal SOC. The phase/bus allocation and the response of the regulators are therefore relevant, in addition to the total power. Recovery losses increase slightly from 384.41 to 385.48 kWh.

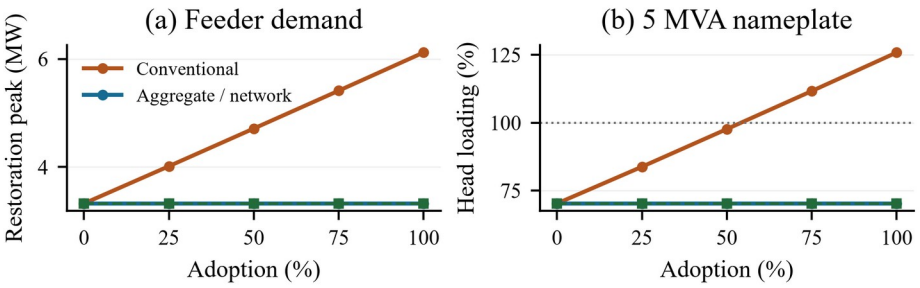


Fig. 3. 3 kW adoption study; coordinated curves overlap. Dotted line: head-device rating.

3.3 Placement and Restoration Robustness

Placement tests maintain the peak-reduction conclusion. At 75% adoption, conventional peaks span 5.410–5.412 MW and minimum voltage ranges from 0.9649 to 0.9692 pu. The coordinated policies remain close to 3.325 MW and 0.9784 pu, with feasible terminal SOC. Under common temporal demand and a fixed adopter count, location has more impact on voltage and losses than on peak demand.

The event matrix is shown in combination with voltage ranges in Figure 4. At 17:00, solar-supported restoration leaves only 0.154 MWh to be replenished and produces no peak reduction. At 21:00, a 5-hour outage requires 6.440 MWh, or an average of 1.073 MW during recovery. The coordinated peak reaches 3.123 MW compared to 2.940 MW without recharge (PRPR 1.0623), while still recovering all batteries. Thus, this 6-hour flexibility does not make the no-recharge peak attainable in all cases. The timing of

outages matters alongside their duration because daylight PV can refill storage before restoration.

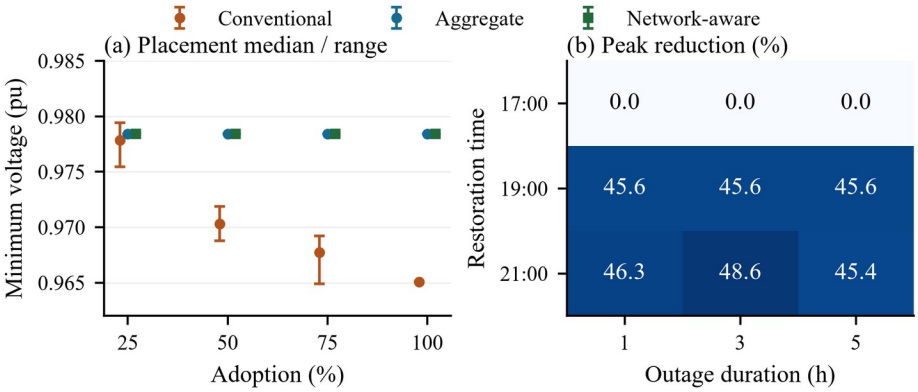


Fig. 4. Placement and event robustness.

3.4 Forecast Ablation and Scope

Held-out modeled-demand MAE/RMSE improve from persistence's 217.12/272.02 kW to XGBoost's 168.24/211.95 kW. PV MAE worsens from 0.05673 to 0.05975 kW despite lower RMSE. All six forecast-to-control ablations produce the same 3.325 MW peak and 2.917 MWh recovery. At 3 kW, loss energy ranges from 381.18 to 382.35 kWh; voltage excursions also depend on allocation. The peak benefit therefore comes mainly from recovery coordination, not forecast sophistication.

Simulink reproduces the Python SOC trajectory at all 45 logged instants, providing an energy-level software-consistency check rather than switching/electrochemical validation. Generalization remains limited by one summer-day base event, modeled households, synchronous assumed outages, fixed efficiencies, ideal metering, and quasi-static controls. Communication delay and utility-specific uncertainty require further assessment.

4 Conclusion

Simultaneous residential battery charging after grid restoration can create high demand peaks and voltage violations. To address this problem, this study develops a grid-aware recovery framework combining receding-horizon scheduling with voltage-sensitivity and AC-feedback correction.

The modeled full-adoption peak is reduced by 45.65%, while preserving 2.917 MWh of recharge energy and restoring all batteries to 90% SOC. The network-aware layer eliminates all three overvoltage intervals with fast chargers without sacrificing recovery. The peak-reduction result is supported by placement tests, while a recovery-headroom limit is revealed by late restoration. Forecast ablation indicates that the main peak-reduction benefit is due to scheduling flexibility rather than increased machine-learning complexity.

The study evaluates an IEEE 123-node benchmark feeder under Iraq-focused operating assumptions. Utility feeder and household PV–battery measurements are the next steps for operational validation.

References

1. Ratnam, E.L., Weller, S.R., Kellett, C.M.: An optimization-based approach to scheduling residential battery storage with solar PV: Assessing customer benefit. Renew. Energy 75, 123-134 (2015). doi:10.1016/j.renene.2014.09.008
2. Gorman, W., et al.: Backup power or bill savings? How electricity tariffs impact residential solar-plus-storage usage in the United States. Utilities Policy 96, 102035 (2025). doi:10.1016/j.jup.2025.102035
3. Gorman, W., et al.: County-level assessment of behind-the-meter solar and storage to mitigate long duration power interruptions for residential customers. Applied Energy 342, 121166 (2023). doi:10.1016/j.apenergy.2023.121166
4. Bu, F., et al.: A Data-Driven Framework for Assessing Cold Load Pick-Up Demand in Service Restoration. IEEE Trans. Power Syst. 34(6), 4739-4750 (2019). doi:10.1109/tpwrs.2019.2922333
5. Schneider, K.P., et al.: Evaluating the Magnitude and Duration of Cold Load Pick-up on Residential Distribution Feeders Using Multi-State Load Models. IEEE Trans. Power Syst. 31(5), 3765-3774 (2016). doi:10.1109/tpwrs.2015.2494882
6. Hachmann, C., Becker, H., Braun, M.: Cold Load Pickup Model Adequacy for Power System Restoration Studies. Energies 15(20), 7675 (2022). doi:10.3390/en15207675
7. Ranaweera, I., Midtgård, O., Korpås, M.: Distributed control scheme for residential battery energy storage units coupled with PV systems. Renew. Energy 113, 1099-1110 (2017). doi:10.1016/j.renene.2017.06.084
8. Alnaser, S.W., et al.: Residential community with PV and batteries: Reserve provision under grid constraints. Int. J. Electr. Power Energy Syst. 119, 105856 (2020). doi:10.1016/j.ijepes.2020.105856
9. Das, N., et al.: Domestic Load Management With Coordinated Photovoltaics, Battery Storage and Electric Vehicle Operation. IEEE Access 11, 12075-12087 (2023). doi:10.1109/access.2023.3241244
10. Kabir, M.N., et al.: Coordinated Control of Grid-Connected Photovoltaic Reactive Power and Battery Energy Storage Systems to Improve the Voltage Profile of a Residential Distribution Feeder. IEEE Trans. Ind. Inform. 10(2), 967-977 (2014). doi:10.1109/tii.2014.2299336
11. Fortenbacher, P., Mathieu, J.L., Andersson, G.: Modeling and Optimal Operation of Distributed Battery Storage in Low Voltage Grids. IEEE Trans. Power Syst. 32(6), 4340-4350 (2017). doi:10.1109/tpwrs.2017.2682339
12. Hu, R., et al.: Coordinated active and reactive power control for distribution networks with high penetrations of photovoltaic systems. Solar Energy 231, 809-827 (2022). doi:10.1016/j.solener.2021.12.025
13. Makhalfih, A., Ibrahim, I.A.: A Two Stage Optimal Power Flow Coordination for PVs, Batteries, and EVs in Unbalanced Low Voltage Distribution Networks. IEEE Open Access Journal of Power and Energy 13, 641-650 (2026). doi:10.1109/oajpe.2026.3721623
14. Bae, D., Kwon, B., Song, K.: XGBoost-Based Day-Ahead Load Forecasting Algorithm Considering Behind-the-Meter Solar PV Generation. Energies 15(1), 128 (2022). doi:10.3390/en15010128

15. Erdener, B.C., et al.: A review of behind-the-meter solar forecasting. Renew. Sustain. Energy Rev. 160, 112224 (2022). doi:10.1016/j.rser.2022.112224
16. Al-bayaty, H., et al.: Electrical distribution grid of Kirkuk City: A case study of load flow and short circuit valuation using ETAP. Period. Eng. Nat. Sci. 10(3), 311-322 (2022). doi:10.21533/pen.v10.i3.665

Mapping the Evolution of Optimization Research on Seismic Isolation Systems Through Bibliometric Analysis

Fatih Tarık Sarı[1][0000-0001-7728-736X]

[1] Department of Civil Engineering, Faculty of Engineering Architecture, Istanbul Beykent University, Istanbul, Türkiye
tariksari@beykent.edu.tr

Abstract. Seismic isolation is recognized as one of the effective approaches for protecting structures against earthquakes. Its performance depends largely on the properties of the isolation layer, and the determination of these properties is mostly treated as an optimization problem. A quantitative overview of the optimization approaches used in the field has not yet been presented. This study presents a bibliometric analysis of research on the optimization of seismic isolation systems. Records published between 1990 and 2025 were retrieved from the Web of Science Core Collection and Scopus databases. The resulting 1,325 documents were analyzed with Bibliometrix and VOSviewer. Annual output rose from a single study in 1990 to 220 studies in 2025, and more than half of the documents were published in the last five years. The genetic algorithm was the most frequently used optimization technique. It was followed by multi-objective formulations and by surrogate and machine learning models. Over time, the share of surrogate and machine learning models doubled while the share of genetic algorithms declined. Keyword co-occurrence and bibliographic coupling analyses revealed a clear shift in research focus. Earlier studies concentrated on semi-active control and on parametric investigations of lead rubber and friction pendulum bearings. Recent studies focus on hybrid passive systems that combine isolation with inerters, tuned mass dampers and negative stiffness devices. Machine learning and seismic resilience emerged as the newest topics. The findings provide a structured map of the field and offer guidance for future studies on isolated structures.

Keywords: Seismic Isolation, Base Isolation, Optimization, Bibliometric Analysis, Bibliometrix, VOSviewer.

1 Introduction

Seismic isolation systems are generally installed as a separate layer between the superstructure and the foundation. This layer shifts the fundamental period of the structure and reduces the seismic energy transmitted to the superstructure [1, 2]. Seismic isolation has a long history [3] and is widely applied to buildings and bridges [4].

The effectiveness of an isolation system depends on the mechanical properties of its isolators. A flexible isolation layer reduces floor accelerations and base shear, but it increases the isolator displacement. This conflict becomes critical under near-fault ground motions containing long-period velocity pulses [5–7]. For this reason, isolator parameters have commonly been determined by means of optimization methods [8].

Early studies relied mainly on parametric and analytical approaches. Subsequent studies adopted genetic algorithms [9, 10], teaching-learning-based optimization [11] and multi-objective methods such as NSGA-II [12]. Optimization has also been used to determine the elevation of the isolation layer [13] and to tune supplementary devices at the isolation level [14–16]. More recently, surrogate and machine learning models have entered the field, and the body of literature has expanded considerably.

Several review studies have summarized the development of seismic isolation [4, 17]. These reviews, however, did not quantify which optimization methods have been used. Bibliometric analysis addresses such questions through objective indicators [18]. Previous review and bibliometric studies in this area have focused on specific isolator types [19], metamaterial-based protection systems [20] or retrofitting techniques [21]. However, none of these studies has examined the optimization methods used for seismic isolation systems.

This study analyzes Web of Science and Scopus records published between 1990 and 2025 with Bibliometrix [22] and VOSviewer [23] and examines the evolution of publication output, optimization methods and research themes over time.

2 Methodology

2.1 Data Collection

Data was retrieved from the Web of Science Core Collection and Scopus on 28 September 2026. Both databases were used because each indexes journals that the other does not cover [24]. Titles, abstracts and keywords were searched with two groups of terms. The first group described the isolation system, such as “base isolation”, “lead rubber bearing” and “friction pendulum”. The second group described optimization, such as “optim*”, “genetic algorithm” and “multi-objective”. English-language articles and reviews published between 1990 and 2025 were included, yielding 1,254 Web of Science and 1,125 Scopus records.

2.2 Data Merging and Cleaning

The datasets were merged in R with a script based on the bibliometrix package, following the approach of Caputo and Kargina [25]. After 31 records published in 2026 and two duplicate records were removed, 987 duplicates between the two databases were identified by DOI and title matching and removed. Missing fields and citation counts were taken from Scopus. Cited references were converted into a common key consisting of the first author, year, volume and first page. Finally, 34 records without a seismic context were excluded. The final dataset contained 1,325 documents, 971 of which were indexed in both databases.

2.3 Analysis

Performance indicators, the thematic map and thematic evolution were obtained with Biblioshiny. The thematic map positions keyword clusters according to their centrality and density [26]. For the thematic evolution, the study period was divided into 1990 to 2017, 2018 to 2020 and 2021 to 2025, since annual output increased sharply in 2018 and 2021.

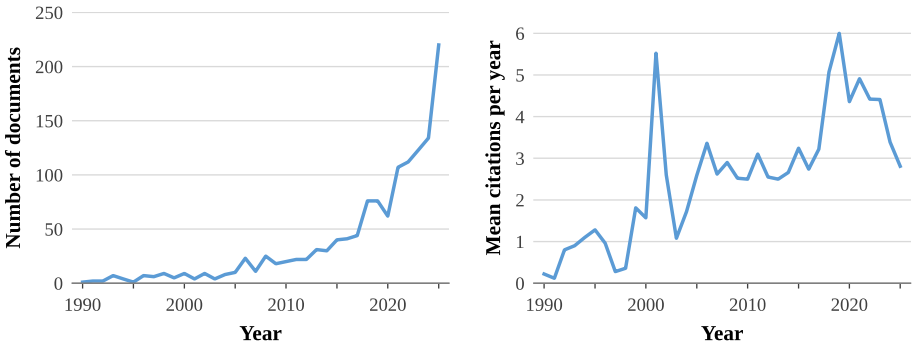
Keyword co-occurrence and bibliographic coupling maps were generated with VOSviewer. After synonyms were merged and four generic terms were removed, author keywords with at least seven occurrences were used. For bibliographic coupling [27], a low threshold of ten citations was adopted so that recent studies would also be represented.

The frequency of optimization methods was determined by searching titles, abstracts and keywords with a dictionary of method names and their abbreviations grouped into predefined categories.

3 Results

3.1 Publication Output and Citation Impact

The 1,325 documents were published in 302 sources. Journal articles accounted for 97.6% of the dataset, and each document received 27.6 citations on average. Annual output did not exceed ten documents until 2005, increased gradually until 2017 and rose sharply in 2018 and 2021 (Fig. 1a). Output reached 220 documents in 2025, and 52.5% of all documents were published after 2020. The annual growth rate was 16.7%. Since 1990 contains a single document, the growth rate for 2000 to 2025 was also calculated as 13.6%.



(a) (b)

Fig. 1. (a) Annual scientific production, (b) average citations per year, 1990–2025.

Average citations per year peaked in 2001 and in 2018 and 2019 (Fig. 1b). The first peak reflects the study of Jangid and Kelly [5], which is the most cited document with

530 citations. The second peak coincides with the first studies that combined isolation with inerter-based devices [15, 16]. The next most cited documents were those of De Domenico and Ricciardi [15] with 508 citations, Zhang and Huo [28] with 501 citations and Fenz and Constantinou [29] with 450 citations.

3.2 Optimization Methods

Genetic algorithms were the most frequently used technique with 140 documents (10.6%), followed by multi-objective formulations (115 documents, 8.7%) and surrogate and machine learning models (90 documents, 6.8%) (Table 1). Other metaheuristics, such as harmony search and grey wolf optimizer, appeared in fewer than ten documents each. Since optimum parameters are frequently obtained through parametric or analytical studies, most documents did not mention any specific algorithm.

Table 1. Number and share of documents mentioning each optimization method by period.

Method category	1990–2017 (n = 415)	2018–2020 (n = 214)	2021–2025 (n = 696)	Total (n = 1,325)
Genetic algorithm (incl. NSGA)	51 (12.3%)	25 (11.7%)	64 (9.2%)	140 (10.6%)
Multi-objective / Pareto	35 (8.4%)	15 (7.0%)	65 (9.3%)	115 (8.7%)
Surrogate and machine learning	18 (4.3%)	12 (5.6%)	60 (8.6%)	90 (6.8%)
Particle swarm optimization	4 (1.0%)	9 (4.2%)	21 (3.0%)	34 (2.6%)
Ant, bee and firefly algorithms	4 (1.0%)	3 (1.4%)	7 (1.0%)	14 (1.1%)
Reliability-based design opt.	4 (1.0%)	1 (0.5%)	5 (0.7%)	10 (0.8%)

Note: A document can mention more than one method. Of the 90 surrogate and machine learning documents, 59 mention only neural networks or machine learning.

Although the absolute number of genetic algorithm studies continued to increase, their share decreased from 12.3% before 2018 to 9.2% after 2020. The share of surrogate and machine learning models doubled from 4.3% to 8.6%. Particle swarm optimization increased from 1.0% to 3.0%, while multi-objective formulations maintained a stable share of approximately 8 to 9%.

3.3 Keyword Structure and Thematic Map

The co-occurrence network comprises 71 author keywords in eight clusters, with seismic isolation (463 occurrences) at its center. In the overlay map (Fig. 2), with average publication years given in parentheses, optimal control (2008.1), active control (2013.1), MR damper (2015.4) and semi-active control (2015.8) are the oldest keywords. Genetic algorithm (2016.9) and friction pendulum (2018.4) occupy an intermediate position. The newest keywords are multi-objective optimization (2021.7), inerter (2022.0), H_2 optimization (2022.3), tuned mass damper inerter (2022.9), negative stiffness (2023.3), seismic resilience (2024.0) and machine learning (2024.1).

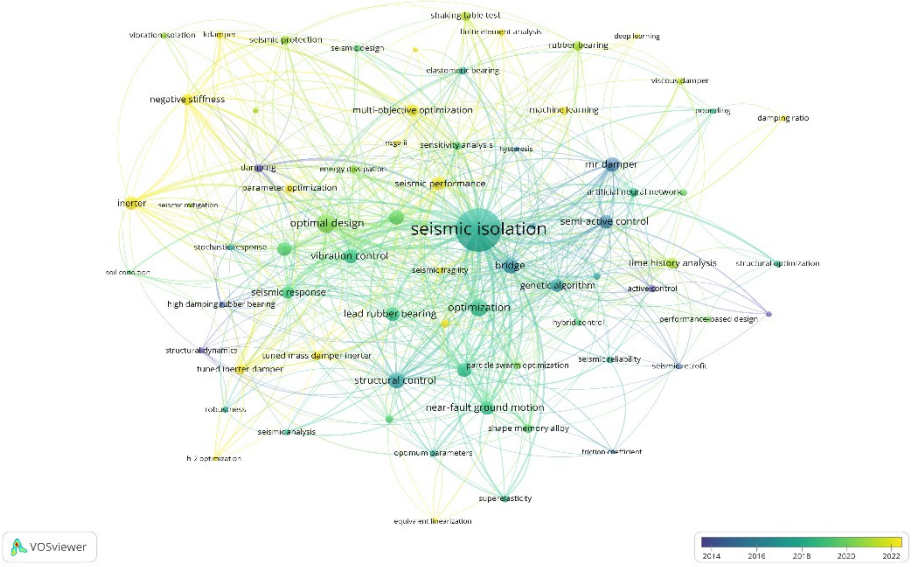


Fig. 1. Overlay map of author keywords colored by average publication year.

The thematic map (Fig. 3) contains two motor themes, one formed by MR damper, semi-active control and genetic algorithm and the other by bridge, near-fault ground motion and friction pendulum. Seismic isolation, optimization and lead rubber bearing constitute a basic theme. Multi-objective optimization and negative stiffness fall into the emerging or declining quadrant, but their recent average publication year mark them as emerging. In the thematic evolution, the MR damper theme flowed mainly into the genetic algorithm theme, while tuned mass damper inerter emerged after 2020.

3.4 Research Fronts

The largest connected coupling network includes 764 documents in nine clusters, and its five main clusters (732 documents, 95.8%) constitute the research fronts in Table 2. Semi-active isolation with MR dampers is the oldest front, whereas hybrid isolation with inerters and negative stiffness devices is the newest.

Table 1. Research fronts from bibliographic coupling (Year = mean publication year).

Front	Docs.	Year	Research focus	Key studies
1	348	2017.0	Parameter optimization of passive isolators (LRB, FPS)	[7, 28, 29]
2	175	2012.5	Semi-active and smart isolation with MR dampers	[30]
3	119	2021.8	Isolation with inerters and negative stiffness devices	[15, 16]
4	55	2016.8	Isolation combined with tuned mass dampers	[14, 31]
5	35	2021.5	Low-cost isolation with rubber–soil mixtures	[32, 33]

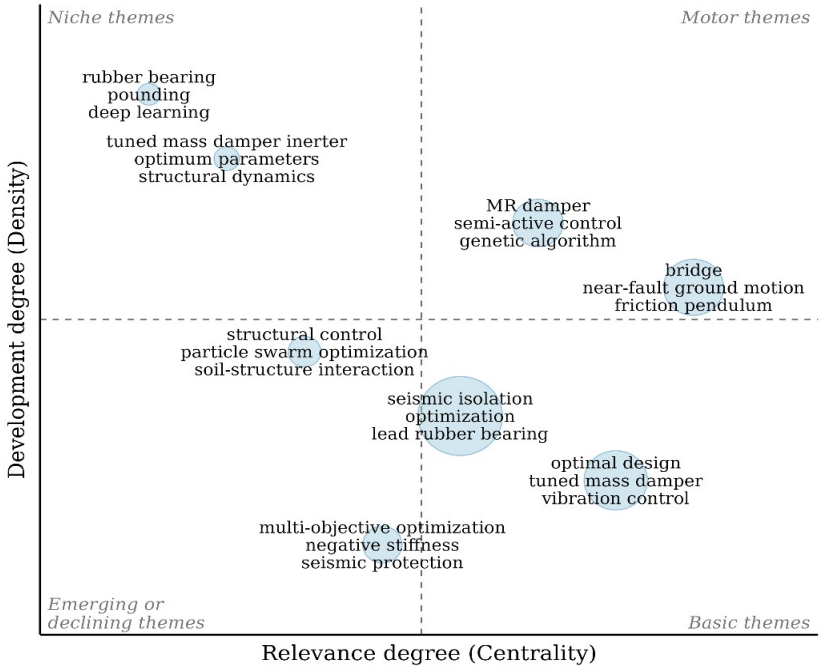


Fig. 1. Thematic map of author keywords (Relevance degree versus development degree).

4 Discussion

The field has grown rapidly and changed direction. Since more than half of the documents were published in the last five years, it appears that the topic is still far from saturation.

Genetic algorithms remain one of the most widely used techniques. Objective functions for isolated structures are generally evaluated through nonlinear time-history analysis and are often nonsmooth and discontinuous, which favors population-based algorithms over gradient-based methods. The decreasing share of genetic algorithms indicates a diversification of the field rather than a loss of preference. Surrogate and machine learning models have become particularly important, since they reduce the computational cost of evaluating many candidate designs.

The maps also reveal a shift in the systems being optimized. In the oldest research front, optimization is mostly concerned with controllers for semi-active devices. In the most recent front, isolation is combined with inerters, tuned mass dampers and negative stiffness elements, which require no external power. Most of these studies employed linear models that yield closed-form H_2 solutions. This may partly explain the declining share of genetic algorithms.

Several gaps remain in the literature. Systematic comparisons of algorithms on common benchmark problems are scarce. Machine learning and seismic resilience have appeared only in recent years, which leaves room for resilience-based objectives.

Although uncertainty is a central issue for isolated structures, reliability-based design optimization appears in only ten documents. The number of documents linking optimization results with design codes is also limited.

This study has some limitations. Records were screened at title level only, and methods not named in titles, abstracts or keywords could not be detected. Citation counts also differ between databases.

5 Conclusions

This study mapped research on the optimization of seismic isolation systems published between 1990 and 2025 using 1,325 documents. Output has grown steadily, with sharp increases in 2018 and 2021. Genetic algorithms are the most frequently used technique, followed by multi-objective formulations and by surrogate and machine learning models, whose share has doubled. The research focus has shifted from semi-active control and single isolators towards hybrid passive systems incorporating inerters, tuned mass dampers and negative stiffness devices. Future studies would benefit from systematic comparisons of algorithms, surrogate-assisted and reliability-based design, and stronger links with design codes.

References

1. Kelly, J.M.: Aseismic base isolation: review and bibliography. Soil Dynamics and Earthquake Engineering. 5, 202–216 (1986). https://doi.org/10.1016/0267-7261(86)90006-0.
2. Naeim, F., Kelly, J.M.: Design of seismic isolated structures: from theory to practice. John Wiley, New York (1999).
3. Makris, N.: Seismic isolation: Early history. Earthquake Engineering & Structural Dynamics. 48, 269–283 (2019). https://doi.org/10.1002/eqe.3124.
4. Warn, G.P., Ryan, K.L.: A Review of Seismic Isolation for Buildings: Historical Development and Research Needs. Buildings. 2, 300–325 (2012). https://doi.org/10.3390/buildings2030300.
5. Jangid, R.S., Kelly, J.M.: Base isolation for near-fault motions. Earthquake Engineering & Structural Dynamics. 30, 691–707 (2001). https://doi.org/10.1002/eqe.31.
6. Jangid, R.S.: Optimum friction pendulum system for near-fault motions. Engineering Structures. 27, 349–359 (2005). https://doi.org/10.1016/j.engstruct.2004.09.013.
7. Jangid, R.S.: Optimum lead–rubber isolation bearings for near-fault motions. Engineering Structures. 29, 2503–2513 (2007). https://doi.org/10.1016/j.engstruct.2006.12.010.
8. Castaldo, P., Ripani, M.: Optimal design of friction pendulum system properties for isolated structures considering different soil conditions. Soil Dynamics and Earthquake Engineering. 90, 74–87 (2016). https://doi.org/10.1016/j.soildyn.2016.08.025.
9. Pourzeynali, S., Zarif, M.: Multi-objective optimization of seismically isolated high-rise building structures using genetic algorithms. Journal of Sound and Vibration. 311, 1141–1160 (2008). https://doi.org/10.1016/j.jsv.2007.10.008.
10. Fallah, N., Zamiri, G.: Multi-objective optimal design of sliding base isolation using genetic algorithm. Scientia Iranica. 20, 87–96 (2013). https://doi.org/10.1016/j.scient.2012.11.004.

11. Öncü-davas, S., Temür, R., Alhan, C.: Teaching-Learning Based Optimization of Nonlinear Isolation Systems under Far Fault Earthquakes. Teknik Dergi. 33, 11487–11505 (2022). https://doi.org/10.18400/tekderg.633636.
12. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation. 6, 182–197 (2002). https://doi.org/10.1109/4235.996017.
13. Charmpis, D.C., Komodromos, P., Phocas, M.C.: Optimized earthquake response of multi-storey buildings with seismic isolation at various elevations. Earthquake Engineering & Structural Dynamics. 41, 2289–2310 (2012). https://doi.org/10.1002/eqe.2187.
14. Taniguchi, T., Der Kiureghian, A., Melkumyan, M.: Effect of tuned mass damper on displacement demand of base-isolated structures. Engineering Structures. 30, 3478–3488 (2008). https://doi.org/10.1016/j.engstruct.2008.05.027.
15. De Domenico, D., Ricciardi, G.: An enhanced base isolation system equipped with optimal tuned mass damper inerter (TMDI). Earthquake Engineering & Structural Dynamics. 47, 1169–1192 (2018). https://doi.org/10.1002/eqe.3011.
16. De Domenico, D., Ricciardi, G.: Optimal design and seismic performance of tuned mass damper inerter (TMDI) for structures with nonlinear base isolation systems. Earthquake Engineering & Structural Dynamics. 47, 2539–2560 (2018). https://doi.org/10.1002/eqe.3098.
17. Bermany, T.H.R., Osman, S.A., Yatim, M.Y.Md.: A state-of-the-art analysis of base isolation systems and future directions for developing a novel multi-directional smart-hybrid isolation system integrated with earthquake early warning system for building structures. Results in Engineering. 25, 104501 (2025). https://doi.org/10.1016/j.rineng.2025.104501.
18. Zupic, I., Čater, T.: Bibliometric Methods in Management and Organization. Organizational Research Methods. 18, 429–472 (2015). https://doi.org/10.1177/1094428114562629.
19. Habib, A., AL Houri, A., Al-Sadoon, Z.A., Barakat, S., Saatcioglu, M.: Seismic isolation of buildings with multi-stage friction pendulum bearings: design, analysis, and energy dissipation perspectives. International Journal of Dynamics and Control. 13, 48 (2025). https://doi.org/10.1007/s40435-024-01555-9.
20. Guevara-Corzo, J.J., Begambre-Carrillo, O.J., García-Sanchez, J.A., Sanchez-Acevedo, H.G.: Passive seismic protection systems with mechanical metamaterials: A current review. Structural Engineering and Mechanics. 82, 417–434 (2022). https://doi.org/10.12989/sem.2022.82.4.417.
21. Li, J., Fomin, N.I., Xiao, S., Yang, K., Zhao, S., Yang, H.: Seismic Enhancement Techniques for Reinforced Concrete Frame Buildings: A Contemporary Review. Buildings. 15, 984 (2025). https://doi.org/10.3390/buildings15060984.
22. Aria, M., Cuccurullo, C.: *bibliometrix*: An R-tool for comprehensive science mapping analysis. Journal of Informetrics. 11, 959–975 (2017). https://doi.org/10.1016/j.joi.2017.08.007.
23. van Eck, N.J., Waltman, L.: Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. 84, 523–538 (2010). https://doi.org/10.1007/s11192-009-0146-3.
24. Mongeon, P., Paul-Hus, A.: The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics. 106, 213–228 (2016). https://doi.org/10.1007/s11192-015-1765-5.

25. Caputo, A., Kargina, M.: A user-friendly method to merge Scopus and Web of Science data during bibliometric analysis. Journal of Marketing Analytics. 10, 82–88 (2022). https://doi.org/10.1057/s41270-021-00142-7.
26. Cobo, M.J., López-Herrera, A.G., Herrera-Viedma, E., Herrera, F.: An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. Journal of Informetrics. 5, 146–166 (2011). https://doi.org/10.1016/j.joi.2010.10.002.
27. Kessler, M.M.: Bibliographic coupling between scientific papers. American Documentation. 14, 10–25 (1963). https://doi.org/10.1002/asi.5090140103.
28. Zhang, J., Huo, Y.: Evaluating effectiveness and optimum design of isolation devices for highway bridges using the fragility function method. Engineering Structures. 31, 1648–1660 (2009). https://doi.org/10.1016/j.engstruct.2009.02.017.
29. Fenz, D.M., Constantinou, M.C.: Behaviour of the double concave Friction Pendulum bearing. Earthquake Engineering & Structural Dynamics. 35, 1403–1424 (2006). https://doi.org/10.1002/eqe.589.
30. Yoshioka, H., Ramallo, J.C., Spencer, B.F.: "Smart" Base Isolation Strategies Employing Magnetorheological Dampers. Journal of Engineering Mechanics. 128, 540–551 (2002). https://doi.org/10.1061/(ASCE)0733-9399(2002)128:5(540).
31. De Angelis, M., Perno, S., Reggio, A.: Dynamic response and optimal design of structures with large mass ratio TMD. Earthquake Engineering & Structural Dynamics. 41, 41–60 (2012). https://doi.org/10.1002/eqe.1117.
32. Tsiavos, A., Alexander, N.A., Diambra, A., Ibraim, E., Vardanega, P.J., Gonzalez-Buelga, A., Sextos, A.: A sand-rubber deformable granular layer as a low-cost seismic isolation strategy in developing countries: Experimental investigation. Soil Dynamics and Earthquake Engineering. 125, 105731 (2019). https://doi.org/10.1016/j.soildyn.2019.105731.
33. Pitilakis, D., Anastasiadis, A., Vratsikidis, A., Kapouniaris, A., Massimino, M.R., Abate, G., Corsico, S.: Large-scale field testing of geotechnical seismic isolation of structures using gravel-rubber mixtures. Earthquake Engineering & Structural Dynamics. 50, 2712–2731 (2021). https://doi.org/10.1002/eqe.3468.

Evaluation of the Performance and Generalizability of the TabPFN Under Different Data Splitting Strategies for Bond Strength Prediction

Yaren Aydın[1], Sinan Melih Nigdeli[1], Gebrail Bekdaş[1] and Ümit Işıkdağ[2]

[1]Department of Civil Engineering, Istanbul University-Cerrahpaşa,
34320 Avcılar/Istanbul/Turkey
yaren.aydin@iuc.edu.tr, melihnig@iuc.edu.tr, bekdas@iuc.edu.tr
[2] Department of Architecture, Mimar Sinan Fine Arts University, 34427 Istanbul, Turkey
umit.isikdag@msgsu.edu.tr

Abstract. Bond strength is one of the critical factors determining the collapse behavior of reinforced concrete structures. This study investigates the performance of the Tabular Prior-Data Fitted Network (TabPFN) model, which has yielded successful results in predicting bond strength between concrete and fiber-reinforced polymers, under various data-splitting strategies. The model was evaluated using Random Hold-out, k-fold cross-validation, repeated cross-validation, Shuffle Split, and group-based splitting strategies that account for duplicate input combinations. Performance was assessed using R^2, RMSE, MAE, and MAPE metrics. The results demonstrated that the TabPFN model achieved high predictive accuracy across all data-splitting strategies. The highest performance was obtained with the Random Hold-out 90/10 strategy, yielding values of $R^2 = 0.967$, RMSE = 1.899 MPa, and MAE = 1.522 MPa. Variations in R^2 values were observed in the duplicate-aware group-based splits. These findings indicate that, in bond strength prediction, not only model selection but also the data-splitting strategy employed plays a significant role in performance evaluation.

Keywords: Bond strength, Machine learning, TabPFN, Data splitting, Prediction.

1 Introduction

Ensuring bond between concrete and reinforcement is of great importance for the composite action of reinforced concrete structural elements [1]. Bond strength is crucial for load transfer, composite behavior, and strain compatibility in structural elements, as it ensures the adhesion between concrete and reinforcing bars [2]. Insufficient bond strength can cause reinforced concrete elements to deviate from their intended design behavior and increase the risk of local failure [1]. Bond performance is influenced by numerous parameters, such as the tensile strength of the concrete, reinforcement diameter and surface characteristics, concrete workability and compactability,

reinforcement position and yield strength, bond length, and concrete cover [3]. Therefore, machine learning-based approaches are increasingly being used to elucidate the complex and non-linear relationships between the parameters in question. Machine learning methods are widely used in civil engineering [4]. These methods enable the identification of the key parameters governing bond behavior, in addition to allowing for the reliable prediction of bond strength.

Prediction of CFRP-steel bond strength by implementing ensemble learning algorithms [5], prediction of interfacial bond strength between FRPs and concrete [6], bond strength prediction of reinforced concrete with corrosion-induced cracks and feature sensitivity analysis [7], prediction of bond strength between UHPC and reinforcing bars [8], bond strength prediction of concrete-encased steel structures [9], prediction of bond strength of FRP bars [10], prediction of FRP-concrete interfacial bond strength [11], prediction of interfacial bond strength of fiber-reinforced polymer-concrete [12], FRP-concrete interfacial bond strength prediction [13], prediction of flexural bond strength between FRP/steel bars and concrete [14], prediction of FRP-concrete interfacial bond strength [15], bond strength prediction at ultra-high-performance-to-normal-strength-concrete interfaces [16] can be cited as examples of research conducted on this topic in the literature.

The success of machine learning models demonstrating high performance in bond strength prediction can depend not only on the model itself but also on how the dataset is partitioned. This study aims to demonstrate this effect by evaluating the TabPFN model under various data-splitting strategies. In particular, incorporating a group-based splitting approach enables a more comprehensive examination of the model's generalization capability across different data groups. The findings are expected to provide a more reliable evaluation framework for selecting data-splitting strategies in bond strength prediction studies.

2 Research Methodology

2.1 Tabular Prior-data Fitted Network (TabPFN)

Tabular Prior-data Fitted Network (TabPFN) [17] is a model developed for tabular data. Thanks to its in-context learning approach, it can generate fast and accurate predictions on small datasets and demonstrates competitive performance across various machine learning benchmarks without requiring fine-tuning. TabPFN addresses fundamental challenges associated with tabular data -such as the lack of sequential structure, feature heterogeneity, and data quality issues- through a one-time pre-training approach conducted on billions of synthetic tabular datasets [18].

2.2 Dataset

The dataset used in the study was obtained from Zhang et al. [19]. A total of 14 independent variables representing the properties of concrete, fibers, and reinforcement were used to predict bond strength. All independent variables are numerical and served as model inputs. These variables include: volume fraction of steel fibers, volume

fraction of polypropylene/polyethylene fibers, concrete type, compressive strength of concrete (MPa), water-to-binder ratio, sand-to-binder ratio, reinforcement bar diameter (mm), ratio of concrete cover diameter, ratio of adhesion length to diameter, ratio of rib spacing to diameter, ratio of rib height to diameter, reinforcement bar surface treatment, bar fiber type, and tensile modulus of elasticity (GPa) of the bar. Fig. 1 shows the distribution of the dataset's target variable, bond strength.

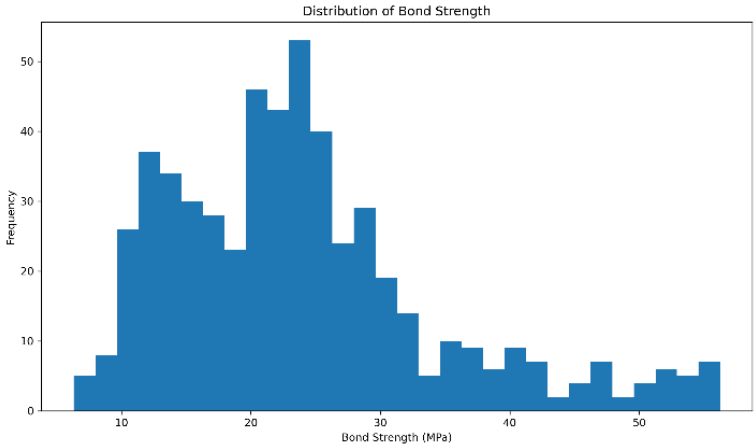


Fig. 1. Distribution of bond strength.

2.3 Data Splitting Strategies

In the holdout method, the dataset is split into two mutually exclusive subsets (training and test). In the k-fold cross-validation method, the dataset is first split into training and testing sets, and the training data is then divided into k equal parts. This process is repeated for all groups, and the resulting error values are compared to determine the optimal model and parameters [20]. To make the selection of a single optimal model structure more stable, the repeated cross-validation (repeated CV) method -which reduces the variability associated with a single run of k-fold CV- is employed. In this method, the performance metric is calculated by aggregating the prediction results across different CV folds and repetitions for each model [21]. Using the duplicate-aware Group K-Fold validation method, inputs are grouped together, and these groups are kept entirely separate between the training and test sets during validation [22]. In the Shuffle Split approach, training and test sets are generated randomly [23].

2.4 Performance Evaluation Criteria

The performance of the TabPFN method in bond strength prediction under different data-splitting strategies was examined using various evaluation criteria (Coefficient of Determination (R^2), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). R^2 indicates the extent to which the

variation in the dependent variable is explained by the model. RMSE measures the error between actual and predicted values. MAE is the average of the absolute differences between predicted values and actual values. MAPE represents the mean absolute percentage error between predicted and actual values.

3 Results and Discussion

The performance of the TabPFN model in predicting bond strength was evaluated under various data-splitting strategies using R^2, RMSE, MAE, and MAPE metrics (Table 1). Results from these different strategies demonstrated that TabPFN's performance can vary depending on the data-splitting approach. Differences were observed between results obtained via random hold-out methods and those from duplicate-aware, group-based methods that account for observation dependencies. In random hold-out strategies, average R^2 values of 0.951, 0.9532, and 0.9667 were obtained for 70/30, 80/20, and 90/10 train-test splits, respectively. Notably, the 90/10 split strategy demonstrated the highest performance, with an average R^2 of 0.9667 ± 0.006. The results indicate that the TabPFN model can predict bond strength with high accuracy using the random split approach. Cross-validation results also revealed high performance metrics. Average R^2 values of 0.961 ± 0.005 and 0.962 ± 0.010 were found for 5-fold and 10-fold cross-validation, respectively. For repeated 5-fold and repeated 10-fold strategies, average R^2 values of 0.9539 ± 0.021 and 0.9582 ± 0.017 were obtained, respectively. With the Shuffle Split strategy, an average R^2 of 0.947 ± 0.023 was obtained; this value was lower than those observed with the random hold-out and standard k-fold methods. This result demonstrates that varying random train-test combinations can lead to significant variability in model performance. The results from Duplicate-aware Group K-Fold demonstrated that the presence of identical independent variable combinations across multiple observations in the dataset is a significant factor in evaluating model performance.

In the study by Zhang et al. [24], which utilized the same dataset, the R^2, MAE, and RMSE values for the test dataset were 0.956, 0.016, and 0.075, respectively. The study conducted by Aydın et al. [25] also employed the same dataset, yielding R^2, MAE, and MAPE values of 0.96, 1.57, and 7.85%, respectively. These results demonstrate that model performance across different studies using the same dataset can vary depending on the employed method and the data-splitting strategy.

Table 1. Cross-validation (CV) and test metrics of the models used.

Strategy	Partition type	Explanation	R^2 Mean	R^2 Std	RMSE Mean	RMSE Stf	MAE Min	MAE Std	MAPE (%) Mean	MAPE (%) Std
1	Random Hold-out	90/10	0.9667	0.0064	1.8991	0.163	1.5221	0.1122	7.621	0.7506
2	k-fold CV	k=10	0.962	0.0097	1.9893	0.1952	1.5771	0.1555	7.8352	1.531
	k-fold CV	k=5	0.961	0.0053	2.0572	0.1762	1.6235	0.1417	8.0822	1.1162
3	Repeated CV	k=10	0.9582	0.0171	2.0652	0.2858	1.6045	0.1526	7.996	0.7853
	Repeated CV	k=5	0.9539	0.0206	2.1928	0.3938	1.6377	0.1361	8.0796	0.7026

1	Random Hold-out	80/20	0.9532	0.0215	2.1804	0.4755	1.6419	0.1887	8.0916	0.9863
	Random Hold-out	70/30	0.951	0.0144	2.3007	0.3058	1.7411	0.1198	8.5941	0.7049
4	Shuffle Split		0.9467	0.0228	2.3376	0.4808	1.7168	0.1863	8.3349	0.8161
5	Duplicate-aware Group	k=5	0.8782	0.0833	3.3974	1.0374	2.3285	0.5746	11.5276	2.9621
	Duplicate-aware Group	k=10	0.865	0.1107	3.4159	1.3034	2.6263	0.7028	11.3038	3.3288
	Duplicate-aware Group	k=3	0.8231	0.1185	4.0435	1.4442	2.5975	0.6427	13.2871	2.9122

In the overall performance comparison, the highest R^2 value was achieved with the Random Hold-out 90/10 strategy ($R^2 = 0.967 \pm 0.006$). However, this result should not be considered in isolation as an indicator of the model's true generalization capacity. The presence of identical or highly similar input combinations in both the training and test sets under the Random Hold-out approach may have enabled test samples to be predicted based on similar instances within the training data. Fig. 2 compares the performance metric values across different data-splitting strategies.

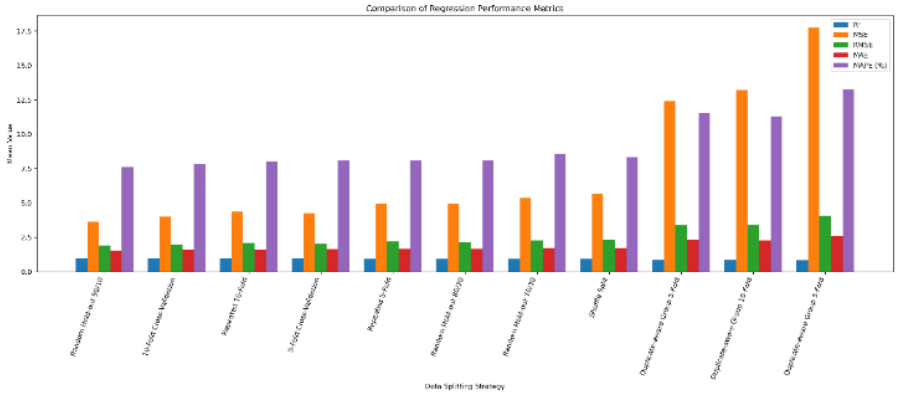


Fig. 2. Comparison of regression performance metrics.

The box plot in Fig. 3 illustrates the distribution of performance across different splits. Notably, the performance distribution widens in duplicate-aware, group-based strategies.

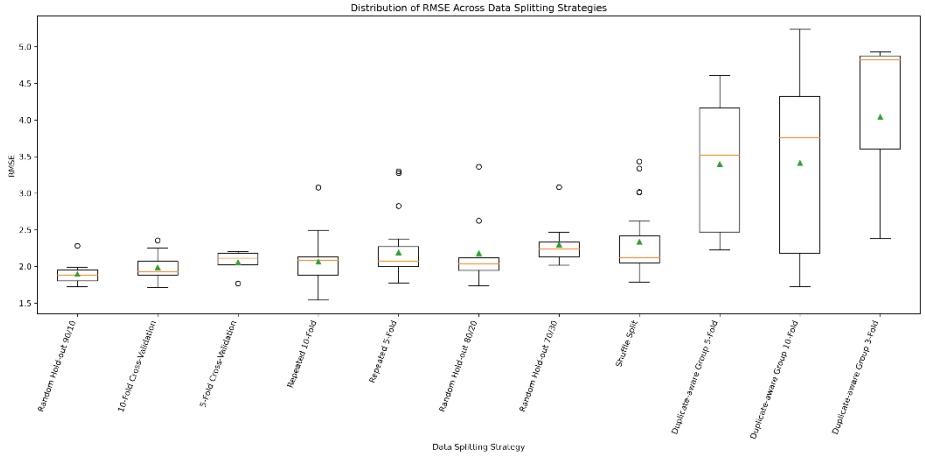


Fig. 3. Distribution of RMSE across data splitting strategies.

A comparison of actual and predicted bond strength values visually demonstrates the performance difference between the data-splitting strategies (Fig. 4). In the Random Hold-out 90/10 strategy, the predictions exhibit a tighter distribution around the 1:1 line, and a high R^2 value is achieved. In contrast, the Duplicate-aware Group 5-Fold strategy shows a prediction distribution that deviates more significantly from the linear reference line.

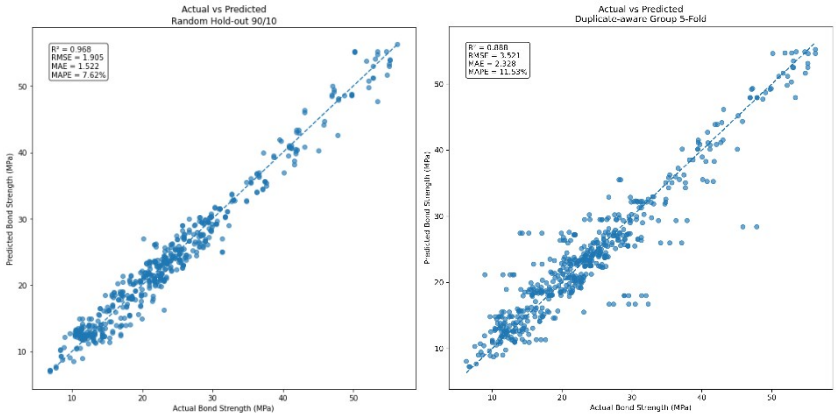


Fig. 4. a) Actual vs Predicted for Random Hold-out 90/10 b) Actual vs Predicted for Duplicate-aware Group 5-Fold.

In conclusion, Random Hold-out 90/10 yielded the highest predictive performance. The results from Duplicate-aware GroupKFold, however, indicate that this high performance is sensitive to the data splitting strategy.

4 Conclusions

In this study, the performance of the TabPFN model -which has proven successful in bond strength prediction- was evaluated under various data-splitting strategies. The results revealed that while the model generally demonstrated high predictive accuracy across all splitting strategies, its performance level varied depending on the specific approach employed. An R^2 value of 0.967 was achieved with the Random Hold-out 90/10 strategy, whereas lower R^2 values were observed with duplicate-aware, group-based splitting methods. These findings indicate that the data-splitting strategy is a crucial factor in assessing model performance and that diverse splitting approaches should be considered to obtain more realistic evaluations of generalization capability.

References

[1] M. W. T. Mak and J. M. Lees, “Bond strength and confinement in reinforced concrete,” *Constr. Build. Mater.*, vol. 355, p. 129012, Nov. 2022, doi: 10.1016/j.conbuildmat.2022.129012.

[2] M. John Robert Prince and B. Singh, "Bond behaviour of deformed steel bars embedded in recycled aggregate concrete," *Constr. Build. Mater.*, vol. 49, pp. 852–862, Dec. 2013, doi: 10.1016/j.conbuildmat.2013.08.031.

[3] T. H. Tusher, K. S. Ahmed, and M. S. B. I. Sarker, "Bond behavior of deformed rebar in concrete containing recycled brick aggregate and waste rubber," *Journal of Building Engineering*, vol. 92, p. 109729, Sep. 2024, doi: 10.1016/j.jobe.2024.109729.

[4] Y. Aydın, G. Bekdaş, and S. M. Nigdeli, "Review of Recent Advances on AI Applications in Civil Engineering," 2025, pp. 107–130. doi: 10.1007/978-981-97-5979-8_6.

[5] S.-Z. Chen, D.-C. Feng, W.-S. Han, and G. Wu, "Development of data-driven prediction model for CFRP-steel bond strength by implementing ensemble learning algorithms," *Constr. Build. Mater.*, vol. 303, p. 124470, Oct. 2021, doi: 10.1016/j.conbuildmat.2021.124470.

[6] M. Su, Q. Zhong, H. Peng, and S. Li, "Selected machine learning approaches for predicting the interfacial bond strength between FRPs and concrete," *Constr. Build. Mater.*, vol. 270, p. 121456, Feb. 2021, doi: 10.1016/j.conbuildmat.2020.121456.

[7] T. Huang, C. Wan, T. Liu, S. Gao, G. Zhang, and C. Miao, "Machine learning-driven interface bond strength prediction of reinforced concrete with corrosion-induced cracks and feature sensitivity analysis," *Constr. Build. Mater.*, vol. 514, p. 145573, Mar. 2026, doi: 10.1016/j.conbuildmat.2026.145573.

[8] Z. Li, J. Qi, Y. Hu, and J. Wang, "Estimation of bond strength between UHPC and reinforcing bars using machine learning approaches," *Eng. Struct.*, vol. 262, p. 114311, Jul. 2022, doi: 10.1016/j.engstruct.2022.114311.

[9] X. Wang, Y. Liu, and H. Xin, "Bond strength prediction of concrete-encased steel structures using hybrid machine learning method," *Structures*, vol. 32, pp. 2279–2292, Aug. 2021, doi: 10.1016/j.istruc.2021.04.018.

[10] R. Li, L. Liu, and M. Cheng, "Estimating the Bond Strength of FRP Bars Using a Hybrid Machine Learning Model," *Buildings*, vol. 12, no. 10, p. 1654, Oct. 2022, doi: 10.3390/buildings12101654.

[11] F. Zhang *et al.*, "Prediction of FRP-concrete interfacial bond strength based on machine learning," *Eng. Struct.*, vol. 274, p. 115156, Jan. 2023, doi: 10.1016/j.engstruct.2022.115156.

[12] M. Kovačević, M. Hadzima-Nyarko, P. Petronijević, T. Vasiljević, and M. Radomirović, "Comparative Analysis of Machine Learning Models for Predicting Interfacial Bond Strength of Fiber-Reinforced Polymer-Concrete," *Computation*, vol. 13, no. 1, p. 17, Jan. 2025, doi: 10.3390/computation13010017.

[13] P. Ge, O. Yang, J. He, Z. Liu, and H. Chen, "Metaheuristic algorithms-optimized machine learning models for FRP-concrete interfacial bond strength prediction," *Advances in Engineering Software*, vol. 208, p. 103971, Oct. 2025, doi: 10.1016/j.advengsoft.2025.103971.

[14] M. Ebrahimzadeh *et al.*, "Interpretable machine learning models for predicting flexural bond strength between FRP/steel bars and concrete," *Structures*, vol. 74, p. 108587, Apr. 2025, doi: 10.1016/j.istruc.2025.108587.

[15] Y. Chen, Y. Zhang, C. Li, and J. Zhou, "Application of XGBoost Model Optimized by Multi-Algorithm Ensemble in Predicting FRP-Concrete Interfacial Bond Strength," *Materials*, vol. 18, no. 12, p. 2868, Jun. 2025, doi: 10.3390/ma18122868.

[16] S. C. Sapkota *et al.*, "Explainable Ensemble Machine Learning Models for Bond Strength Prediction at Ultra-High-Performance-to-Normal-Strength-Concrete Interfaces," *Buildings*, vol. 16, no. 15, p. 3081, Aug. 2026, doi: 10.3390/buildings16153081.

[17] N. Hollmann, S. Müller, K. Eggensperger, and F. Hutter, “TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second,” *arXiv preprint* , 2022.

[18] N. Hollmann *et al.*, “Accurate predictions on small data with a tabular foundation model,” *Nature*, vol. 637, no. 8045, pp. 319–326, Jan. 2025, doi: 10.1038/s41586-024-08328-6.

[19] P. F. Zhang, X. L. Zhao, D. Zhang, M. Iqbal, and X. Zhao, “Dataset of bond strength and failure mode between FRP bars and UHPC or UHPSSC,” Mendeley Data.

[20] M. Stone, “Cross-Validatory Choice and Assessment of Statistical Predictions,” *J. R. Stat. Soc. Series B Stat. Methodol.*, vol. 36, no. 2, pp. 111–133, Jan. 1974, doi: 10.1111/j.2517-6161.1974.tb00994.x.

[21] R. Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” *Int. Jt. Conf. Artif. Intell.*, vol. 14, pp. 1137–1145, 1995.

[22] Y. Zhang *et al.*, “Physics-Consistent Monotonic Machine Learning for RMR Prediction from Q-System Parameters Under Duplicate-Aware Validation,” *Rock Mech. Rock Eng.*, Sep. 2026, doi: 10.1007/s00603-026-05949-3.

[23] D. Stefas, N. Gyftokostas, E. Nanou, P. Kourelias, and S. Couris, “Laser-Induced Breakdown Spectroscopy: An Efficient Tool for Food Science and Technology (from the Analysis of Martian Rocks to the Analysis of Olive Oil, Honey, Milk, and Other Natural Earth Products),” *Molecules*, vol. 26, no. 16, p. 4981, Aug. 2021, doi: 10.3390/molecules26164981.

[24] P.-F. Zhang *et al.*, “Prediction of bond strength and failure mode of FRP bars embedded in UHPC or UHPSSC utilising extreme gradient boosting technique,” *Compos. Struct.*, vol. 346, p. 118437, Oct. 2024, doi: 10.1016/j.compstruct.2024.118437.

[25] Y. Aydın, G. Bekdaş, S. M. Nigdeli, and Ü. Işıkdağ, “Comparative Analysis of Machine Learning Models for Predicting Bond Strength Between FRP Bars and UHPC/UHPSSC,” 2026, pp. 153–161. doi: 10.1007/978-3-032-30310-3_13.

Performance of the Competition-Amensalism Optimization (CAO) Algorithm in RC Beam Cost Optimization Under Different Cost Scenarios and Comparison with Other Metaheuristic Algorithms

Yaren Aydın[1], Sinan Melih Nigdeli[1] and Gebrail Bekdaş[1]

[1]Department of Civil Engineering, Istanbul University-Cerrahpaşa,
34320 Avcılar/Istanbul/Turkey
yaren.aydin@iuc.edu.tr, melihnig@iuc.edu.tr, bekdas@iuc.edu.tr

Abstract. In this study, the performance of Competition-Amensalism Optimization (CAO), Genetic Algorithm (GA), Bat Algorithm (BA), and Particle Swarm Optimization (PSO) algorithms was compared for the cost optimization of RC beams. The algorithms were evaluated based on various criteria. The comparison results revealed that while all algorithms achieved the same minimum cost value, they differed in terms of computational efficiency. The CAO algorithm, which demonstrated lower average number of iterations, was selected for the second stage of the study. In this stage, a total of 20 cost scenarios were generated by utilizing four different unit cost values for concrete and five different unit cost values for reinforcement steel. For each scenario, the effects of variations in material costs on the optimum cost, beam cross-section dimensions, and reinforcement area were examined while keeping other problem parameters constant. The findings have revealed that, in the cost optimization of reinforced concrete beams, the effects of variations in material costs on the optimum design must be considered alongside the computational performance of the algorithms. Different material cost scenarios have demonstrated that the optimum beam cross-section and reinforcement layout can vary depending on the interaction between cost components and design constraints.

Keywords: Cost optimization, RC beam, material cost scenarios, Competition-Amensalism Optimization, metaheuristic algorithms.

1 Introduction

Reinforced concrete (RC) beams play a significant role in transferring vertical loads from slabs to columns and in resisting lateral loads through frame action. The safe and economical design of RC beams requires not only meeting load-bearing capacity requirements but also appropriately determining cross-sectional dimensions and the amount of reinforcement [1]. Optimization is the process of minimizing or maximizing the objective function of a problem subject to specific conditions and constraints [2]. In

engineering problems, optimization aims to determine the optimal solution based on the mathematical relationships between objectives, decision variables, and constraints.

Various techniques are employed to solve optimization problems. While classical methods rely primarily on mathematical and differential calculations, artificial intelligence and heuristic approaches offer effective alternatives for solving complex problems [3], [4]. Particularly in engineering design problems, the simultaneous evaluation of criteria such as cost and reliability enables the system to be designed in a more suitable and sustainable manner.

Optimization is frequently used in structural engineering. Selected studies (design optimization of steel coupling beams [5], optimization of RC circular columns [6], optimization of Tuned Liquid Damper [7], topology optimization for nonlinear materials [8], optimization for eco-friendly RC frames [9], optimization of real-size complex steel space frame structures [10], Optimization of Tuned Mass Damper Inerter [11], optimization and performance assessment of Multi-Tuned Mass Dampers [12]) show the diverse applications of optimization methods in structural engineering.

In this study, the problem involves the cost-optimal design of an RC beam. To achieve

2 Materials and Method

The study consists of two stages. In the first stage, the CAO, GA, BA, and PSO algorithms were compared regarding the cost optimization problem of RC beams. The performance of the algorithms was evaluated based on the optimal cost values obtained, and the number of iterations required for convergence. Based on this evaluation, the most successful algorithm was identified, and its performance under varying material costs was examined in the second stage of the study. Thus, the study investigated how the optimization performance of the selected algorithm and the resulting solutions changed in response to variations in cost parameters.

2.1 Problem Definition and Constraints

A simply supported RC beam with an 8 m span and a uniformly distributed load of 30 kN/m was used to identify the best-performing algorithm among those employed. Calculations were based on the TS500 [13] standard. In Fig. 1, b represents the beam width, h represents the beam height, and A_s represents the reinforcement area. In the cost calculations, current unit prices for concrete and reinforcing steel for the year 2026-the year the study was conducted -were used as the basis to ensure that the optimization results reflect the material costs of the relevant period.

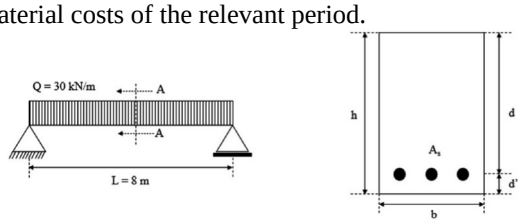


Fig. 1. Simply supported RC beam and cross-section.

In the optimization problem, the mathematical expression of the objective function is given in Equation 1. C_c is the cost of concrete per m³ ($); C_s is the cost of steel per ton ($); and γ_s is the specific weight of steel (t/m³).

$$f(x) = C_c \times b \times d \times L + C_s \times A_s \times L \times \gamma_s \quad (1)$$

The lower and upper limits for the design variables of the RC beam cross-section have been selected as 250-500 (mm) for beam width and 300-1000 (mm) for beam height. To ensure the load-bearing capacity of the beam cross-section under the effect of the bending moment, the condition that the discriminant expression must not be negative is defined as a design constraint. This condition ensures that the moment-carrying capacity is sufficient given the cross-sectional dimensions and material strengths.

2.2 Optimization Algorithms

No single metaheuristic algorithm performs optimally across all optimization problems; consequently, researchers continue to develop a wide range of such algorithms. One of the most recent algorithms in this field is the Competition-Amensalism Optimization (CAO) algorithm, which draws inspiration from ecological interaction theory and competitive survival behaviors. It integrates exploration and exploitation mechanisms through competitive interactions [14]. Genetic Algorithm (GA) is a method inspired by the principles of natural selection and the survival of the fittest. The algorithm consists of selection, crossover, and mutation stages. This process aims to pass successful individuals on to subsequent generations and improve the quality of the solution over successive generations [15]. Particle Swarm Optimization (PSO) is an optimization algorithm inspired by the movements of flocks and schools, such as those of birds and fish. Particles approach the optimal solution in each iteration by leveraging individual and global best solutions [16] The Bat Algorithm (BA) is a bio-inspired metaheuristic optimization algorithm inspired by the echolocation behavior of bats. The behavior of bats -using sound echoes to locate prey and avoid obstacles- has been adapted for the optimization process [17].

2.3 Creation of Different Cost Scenarios

To evaluate the performance of the best-performing algorithm -selected specifically for the problem addressed in this study- under varying material costs, 20 distinct cost scenarios were generated. These scenarios incorporated different values for the unit cost of concrete (C_c) and the unit cost of reinforcing steel (C_s). The values used were 85, 90, 95, and 100 for concrete unit cost, and 900, 1000, 1100, 1200, and 1300 for reinforcing steel unit cost. A total of 20 scenarios were derived by evaluating all possible combinations of these values. By keeping other problem parameters constant, the study examined the impact of fluctuations in material costs on the optimal RC beam design and the resulting minimum cost.

3 Results

In this section, the performance of metaheuristic algorithms used for the cost optimization problem of RC beams was compared, and the most successful algorithm was identified. The CAO algorithm, selected based on the comparison results, was evaluated under 20 different material cost scenarios. Within this scope, the effects of variations in the unit costs of concrete and reinforcing steel on the optimum beam cross-section and the minimum total cost were investigated.

It is observed that all the algorithms CAO, GA, BA, and PSO achieved the same minimum cost value (Table 1). However, the algorithms differed in terms of number of iterations required to reach the best solution. The CAO algorithm stood out by requiring fewer average iterations compared to the other algorithms. These results demonstrate that the CAO algorithm can find the optimal cost value with fewer iterations. Consequently, given that it offers advantages in computational efficiency and convergence speed among algorithms that yield identical results in terms of solution quality, the CAO algorithm was selected for the second stage of the study, which involved analyzing 20 different material cost scenarios.

All four algorithms used in the study reached the same minimum cost value of 213.4968. Across 20 independent runs, the CAO, GA, and PSO algorithms consistently achieved this cost value in every run; consequently, their mean costs were equal to the minimum value, and their standard deviations remained at a negligible level. In contrast, BA reached this minimum value in only one of the 20 runs, while yielding higher costs in the remaining runs. As a result, the BA produced a mean cost of 220.6926 and a standard deviation of 6.7279. These findings indicate that, although the BA is capable of identifying the optimal solution, it exhibits lower consistency under the selected parameter settings, leading to greater variability in solution quality.

Table 1. Performance comparison of algorithms used.

Algorithm	Cost (Min)	Cost (Mean)	Cost (Std)	Best Iteration (Mean)
CAO	213.49681257	213.5	0	138.95
GA	213.49681257	213.5	4.072e-14	2522.1
BA	213.49681257	220.69	6.7279	9707.4
PSO	213.49681257	213.5	0	3047.3

Fig. 2(a) compares the convergence behaviors of the algorithms during the optimization process, while Fig. 2(b) illustrates the general convergence trend observed across different runs. These graphs enable a joint evaluation of the algorithms' solution quality and convergence speeds.

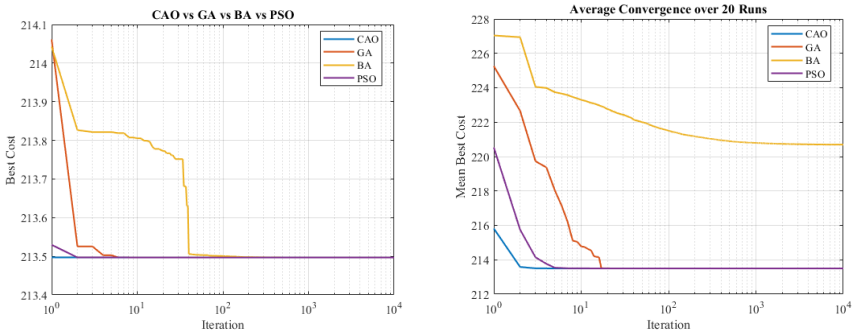


Fig. 2. Comparison of the convergence performance of the algorithms (a) Convergence curves of the CAO, GA, BA, and PSO algorithms. (b) Average convergence curves of the algorithms over 20 independent runs.

Fig. 3 presents the convergence curves obtained for the CAO algorithm across 20 different cost scenarios. Thus, the impact of varying cost conditions on the optimization process and the resulting objective function values can be observed.

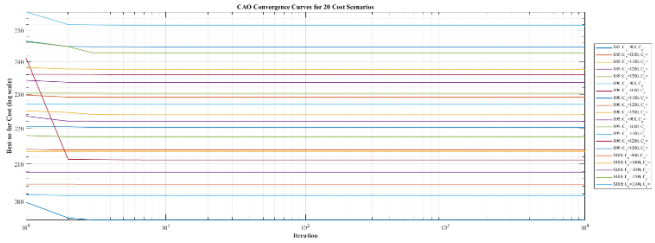


Fig. 3. CAO convergence curves for 20 cost scenarios.

Fig. 4 compares the minimum cost values obtained for 20 different cost scenarios using the CAO algorithm. Variations in the unit costs of concrete and reinforcing steel across the scenarios influenced the total cost of the optimal RC beam design.

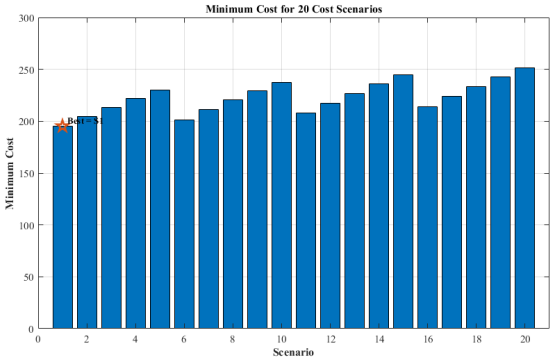


Fig. 4. Minimum cost values obtained with the CAO under different cost scenarios.

Fig. 5 illustrates, via a heat map, the minimum cost values obtained using CAO for various unit costs of concrete and reinforcing steel. Thus, the impact of variations in the unit costs of concrete and reinforcing steel on the total cost of the optimum RC beam design has been visually evaluated. Lower costs are represented by the dark blue color range, while higher costs are represented by yellow.

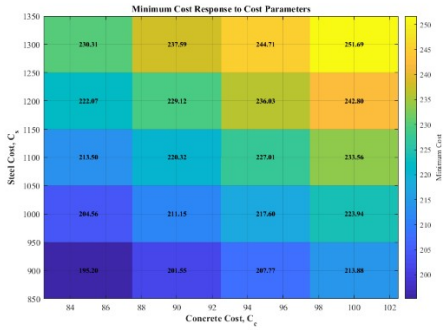


Fig. 5. Minimum cost response to cost parameters.

Fig. 6 illustrates the optimum beam cross-section dimensions and reinforcement areas obtained for 20 different cost scenarios. Figure 6(a) presents the variation of the optimum beam width and height across the scenarios, while Figure 6(b) shows the variation in the optimum reinforcement area. These results enable an assessment of whether changes in the unit costs of concrete and reinforcement steel affect not only the total cost but also the optimum design variables.

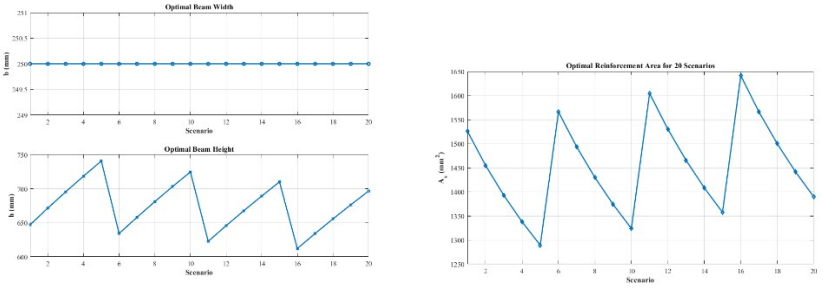


Fig. 6. (a) Optimal beam width – Optimal beam height. (b) Optimal reinforcement area for 20 scenarios.

Variations in the unit costs of concrete and reinforcing steel across different scenarios can influence the combination of design variables that minimizes total cost. Consequently, the fact that optimum beam depth and reinforcement area do not exhibit a monotonic or consistent trend across scenarios can be attributed to the differences between the optimum solutions obtained under varying cost conditions.

The relationship between beam depth and reinforcement area depends not only on material costs but also on the conditions governing the section's flexural moment capacity. In rectangular RC beams, the internal force couple formed by the tension

reinforcement and the concrete compression zone determines the moment capacity. While other factors remain constant, increasing the effective depth can lengthen the lever arm and reduce the reinforcement area required for a given moment capacity. However, since the value of the compression block depth also depends on the reinforcement area and design constraints must be satisfied simultaneously, this relationship does not necessarily manifest in the same way across every scenario.

In conclusion, the optimum beam depth and reinforcement area values obtained in different scenarios should be interpreted by considering the relative impact of concrete and reinforcing steel costs alongside the section's load-bearing capacity requirements. The solution determined by the CAO algorithm for each scenario represents a design configuration aimed at minimizing total cost within the framework of the defined objective function and design constraints.

4 Conclusions

In this study, the performance of CAO, GA, BA, and PSO algorithms was compared for the cost optimization of RC beams, and the CAO algorithm was evaluated across 20 different material cost scenarios. All algorithms achieved the same minimum cost value; however, differences were observed regarding the average number of iterations. Due to its shorter lower average iteration count, the CAO algorithm was selected for the analysis of the various cost scenarios.

In the second stage, a total of 20 scenarios were generated by setting the unit cost of concrete to 85, 90, 95, and 100, and the unit cost of reinforcement steel to 900, 1000, 1100, 1200, and 1300. The results indicated that variations in material costs should be evaluated in conjunction with beam cross-sectional dimensions and reinforcement area, alongside the optimum total cost. Differences in optimum design variables were interpreted within the context of the relative impact of cost parameters and the requirements regarding the cross-section's flexural moment capacity.

In conclusion, this study highlights the importance of comparing algorithms for RC beam cost optimization not only based on the minimum cost achieved but also in terms of computational efficiency. Furthermore, the examination of different material cost combinations offers a systematic approach to assessing the sensitivity of the optimum design to economic conditions.

References

[1] Z. Celep, *Betonarme Yapılar*. Istanbul: Beta Basım Yayın, 2005.

[2] A. M. Nassef, M. A. Abdelkareem, H. M. Maghrabie, and A. Baroutaji, “Review of Metaheuristic Optimization Algorithms for Power Systems Problems,” *Sustainability*, vol. 15, no. 12, p. 9434, Jun. 2023, doi: 10.3390/su15129434.

[3] Z. A. Arfeen *et al.*, “A comprehensive review of modern trends in optimization techniques applied to hybrid microgrid systems,” *Concurr. Comput.*, vol. 33, no. 10, May 2021, doi: 10.1002/cpe.6165.

[4] S. Ni *et al.*, “Review on Data- and Mathematics-Driven Optimization for Metamaterial Lattice Structures,” *Adv. Eng. Mater.*, vol. 28, no. 3, Feb. 2026, doi: 10.1002/adem.202501701.

[5] O. Tunca and S. Carbas, “Performance-based design optimization of steel coupling beams: Evaluating trade-offs between seismic energy dissipation and material efficiency under geometric imperfections and nonlinear behavior,” *Structures*, vol. 89, p. 112078, Jul. 2026, doi: 10.1016/j.istruc.2026.112078.

[6] G. Bekdaş, C. Cakiroglu, S. Kim, and Z. W. Geem, “Optimization and Predictive Modeling of Reinforced Concrete Circular Columns,” *Materials*, vol. 15, no. 19, p. 6624, Sep. 2022, doi: 10.3390/ma15196624.

[7] A. Ocak, G. Bekdaş, S. M. Nigdeli, S. Kim, and Z. W. Geem, “Optimization of Tuned Liquid Damper Including Different Liquids for Lateral Displacement Control of Single and Multi-Story Structures,” *Buildings*, vol. 12, no. 3, p. 377, Mar. 2022, doi: 10.3390/buildings12030377.

[8] D. Wei, F. Zeng, J. Cui, and Y. Wang, “Machine learning-based multiscale topology optimization framework for nonlinear materials,” *Int. J. Mech. Sci.*, vol. 314, p. 111380, Mar. 2026, doi: 10.1016/j.ijmecsci.2026.111380.

[9] G. Bekdaş, S. M. Nigdeli, S. Kim, and Z. W. Geem, “Modified Harmony Search Algorithm-Based Optimization for Eco-Friendly Reinforced Concrete Frames,” *Sustainability*, vol. 14, no. 6, p. 3361, Mar. 2022, doi: 10.3390/su14063361.

[10] S. Carbas and O. Tunca, “Numerical design optimization of real size complex steel space frame structures using a novel adaptive cuckoo search method,” *Int. J. Numer. Methods Eng.*, vol. 125, no. 4, Feb. 2024, doi: 10.1002/nme.7386.

[11] Y. Aydın, G. Bekdaş, S. M. Nigdeli, and Z. W. Geem, “Parameter Optimization of Tuned Mass Damper Inerter via Adaptive Harmony Search,” *Computer Modeling in Engineering & Sciences*, vol. 141, no. 3, pp. 2471–2499, 2024, doi: 10.32604/cmes.2024.056693.

[12] S. Djerouni, G. Bekdaş, and S. M. Nigdeli, “Optimization and performance assessment of Multi-Tuned Mass Dampers (MTMD) to mitigate seismic pounding of adjacent buildings via a novel hybrid algorithm,” *Journal of Building Engineering*, vol. 103, p. 112168, Jun. 2025, doi: 10.1016/j.jobe.2025.112168.

[13] Turkish Standards Institution, “TS500 Requirements for Design and Construction of Reinforced Concrete Structures,” ANKARA, 2000.

[14] A. A. Yilma and B. H. Woldegiorgis, “Competition-amensalism optimization: A novel metaheuristic optimization algorithm,” *Results in Engineering*, vol. 30, p. 111048, Jun. 2026, doi: 10.1016/j.rineng.2026.111048.

[15] J. Holland, *Adaption in natural and artificial systems: an introductory analysis with application to Biology*. University of Michigan Press, 1975.

[16] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proceedings of ICNN’95 - International Conference on Neural Networks*, IEEE, pp. 1942–1948. doi: 10.1109/ICNN.1995.488968.

[17] X.-S. Yang, “A New Metaheuristic Bat-Inspired Algorithm,” 2010, pp. 65–74. doi: 10.1007/978-3-642-12538-6_6.

A Photo-to-3D Workflow Using TRELLIS.2 for Generative Architectural Representation

Handan Aş Çemrek[1][0009-0007-7369-8576], Ümit Işıkdağ[2][0000-0002-2660-0106], Gebrail Bekdaş[3][0000-0002-7327-9810] and Sinan Melih Niğdeli[4][0000-0002-9517-7313]

[1] Istanbul Health and Technology University, Istanbul, 34275, Türkiye
[2] Mimar Sinan Fine Arts University, Istanbul, 34427, Türkiye
[3,4] Istanbul University, Cerrahpasa, Istanbul, 34098, Türkiye

[1] handan.as@istun.edu.tr; [2] umit.isikdag@msgsu.edu.tr
[3] bekdas@iuc.edu.tr; [4] melihnig@iuc.edu.tr

Abstract. Single-image 3D generation offers new possibilities for representing buildings documented by limited visual records. However, a textured model's realistic appearance does not demonstrate the geometric preservation of openings, roof connections, or architectural details. This study presents a TRELLIS.2-based photo-to-3D workflow using the same 16-building dataset employed in previous research and an evaluation method that separates geometry from surface appearance. The material comprises photographs of buildings in Turkey and approximately isometric daytime images prepared with Gemini Flash 2.5. Keeping these inputs constant reduces variation caused by building selection and image reprocessing. The three-dimensional generation stage is redefined with TRELLIS.2, while evaluation examines main mass, facade layout, openings, roof shape, architectural details, and the consistency of unseen surfaces. Observed surfaces are assessed against fixed input images, whereas unseen regions are examined for internal architectural consistency. This distinction prevents plausible visual completion from being treated as verified information about real buildings. Components are classified as geometric, textural, combined, incomplete, or inconsistent representations. Comparing geometric and textured views clarifies whether windows, doors, balconies, and eaves are represented through surface geometry or texture alone. Previous experimental findings indicate preservation of overall form alongside simplified details and generative completion of unseen surfaces. Two pilot cases illustrate the workflow; no model superiority is claimed. The contribution is an evaluation approach for testing a new generative model on a fixed dataset, distinguishing visual similarity from geometric fidelity and clarifying architectural usability. The approach supports assessing generated assets for architectural visualization, education, and digital heritage presentation.

Keywords: Single-image 3D generation; TRELLIS.2; Generative artificial intelligence; Architectural representation; Digital cultural heritage; Geometry and texture.

1 Introduction

Architectural heritage models support documentation, education, and public access, but photographs alone cannot establish dimensional accuracy. Studies combined image editing and 3D generation on the same dataset [1–3]. Münster et al. identify research, access, and virtual experiences as applications of AI in digital heritage [4]. Single-image generation combines visible cues with learned priors. RealFusion reconstructs observed and plausible unseen surfaces [5], while Zero-1-to-3 generates alternative views using diffusion-based priors [6]. Such completion does not establish correspondence to the actual building. Architectural fidelity concerns massing, openings, floor alignment, and roof relationships. A textured window may resemble a geometric opening but offer different possibilities for editing and analysis.

TRELLIS uses structured latent representations [7]. TRELLIS.2 jointly encodes geometry and appearance through O-Voxel, supporting open and non-manifold surfaces and PBR properties [8,9]. Architectural detail preservation nevertheless requires evaluation on building images.

1.1 Related Work and Research Positioning

ICASE 2025 combined Gemini Flash 2.5 isometric preprocessing with Hun-yuan3D 2.1 for 16 buildings [1]. ICICC compared Gemini and Qwen-Image-Edit preprocessing on the same buildings [2], while ABC2 examined architectural usability and LOD3-oriented representation [3]. Hunyuan3D 2.1 separates shape and texture generation [10,11]; Step1X-3D com-bines VAE-DiT geometry with diffusion-based texturing [12,13]. Hi3DGen uses normal maps [14], whereas SAM 3D Objects generates geometry, texture, and placement from masked inputs in complex scenes [15,16]. ArchiSet evaluates window-to-wall ratios, window consistency, and editability [17]. ArchiDiff combines single-image architectural generation with interactive editing [18]. Both establish architectural criteria beyond visual plausibility.

1.2 Research Aim

This study examines how recognizable architectural details correspond to mesh geometry and how generated surfaces extend the visible arrangement. Its scope is external architectural representation.

2 Materials and Methods

2.1 Dataset and Visual Inputs

The 16 buildings [1–3] cover varied massing, roofs, and facades: Ankara Palace, Turkish Grand National Assembly Museum, Ankara Republic Museum, Gazi Education Institute, Aynalıkavak Palace, Baruthane, Baruthane Tower, Bruno Taut House, Doğan Apartment, Florya Atatürk Sea Pavilion, Frej Apartment, Arif Paşa Apartment, Narmanlı Han, Pera Palace, Safranbolu House, and Second Evkaf Apartment.

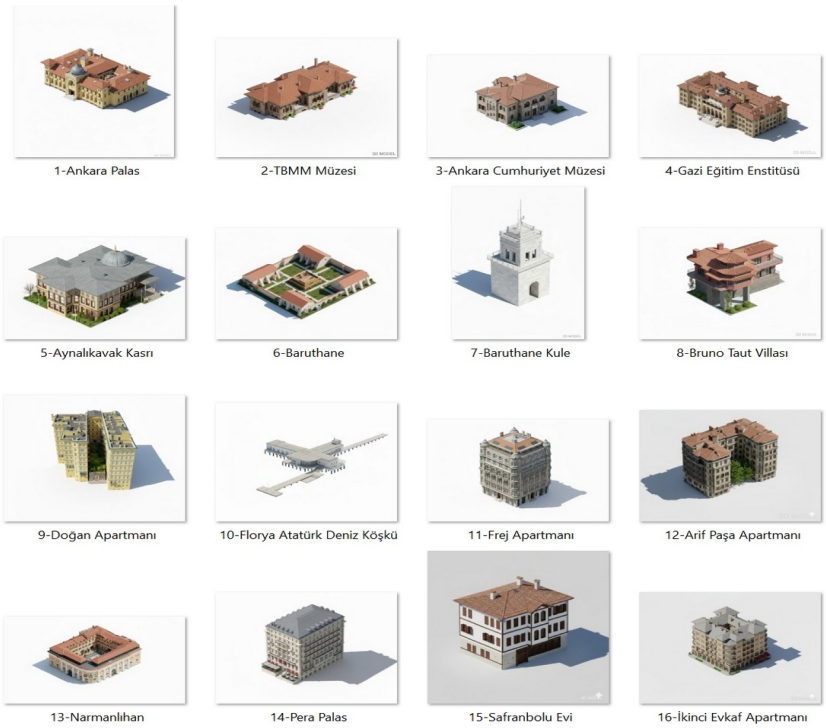


Fig. 1. Approximate isometric images of the same 16-structure dataset prepared with Gemini Flash 2.5. Image set from previous study [3]; these images are not TRELLIS.2 outputs.

The photographs provide preprocessing references. The same approximately isometric daytime Gemini Flash 2.5 images used in earlier work [1,3] serve as fixed TRELLIS.2 inputs. Qwen images are excluded.

The prompt was “make image daytime and isometric (building only)” [1,3]. Images show multiple faces from an approximate isometric viewpoint. Inferred side surfaces require separate assessment of photograph-to-image and image-to-mesh fidelity.

2.2 TRELLIS.2 Generation Workflow

TRELLIS.2-4B replaces Hunyuan3D in the 3D generation stage. Its O-Voxel representation encodes geometry and appearance using a VAE and flow matching [8]. Figure 2 shows the three-stage workflow for Ankara Palas and the TBMM Museum.

The Microsoft TRELLIS.2 demo used resolution 1024, fixed seed 42, and default advanced settings. GLB export targeted 300,000 faces and 2048 × 2048 pixel textures.

The exported GLBs required EXT_texture_webp, which the desktop viewer did not support. The embedded WebP textures were decoded and losslessly re-encoded as PNG in separate GLB copies. Texture resolution and decoded pixels, mesh geometry, UV coordinates, and material parameters were verified unchanged.

Mesh simplification and remeshing occur during GLB export [9]. The later PNG conversion served only to make the exported assets viewable.

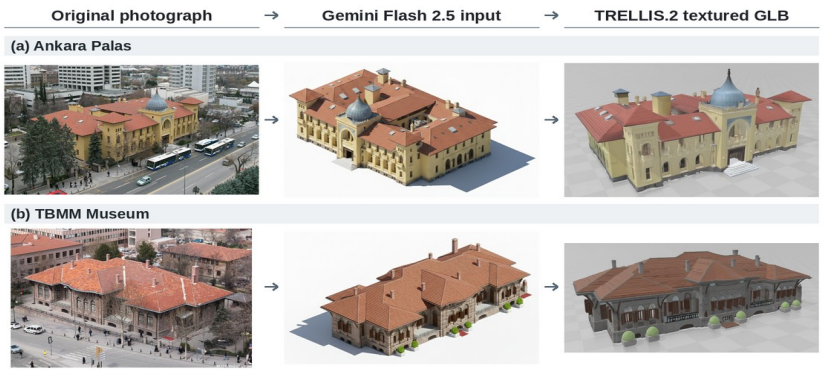


Fig. 2. Photo-to-3D workflow for two pilot cases: source photographs [1–3], fixed Gemini inputs, and textured TRELLIS.2 GLBs shown from illustrative viewpoints.

2.3 Geometry and Surface Appearance

Matched geometry and PBR views retain identical camera positions, framing, and lighting. Geometry views use a uniform opaque material with generated maps disabled; PBR views retain generated surface properties. Material information is thus the principal variable.

The comparison distinguishes textured window effects from mesh recesses, frames, and openings. Neutral renders are supplemented by alternative views, sections, or mesh inspection when topology remains uncertain. Transparency does not establish topological openness, and a recess differs from a through-opening. Opening assessment therefore includes geometric construction.

2.4 Architectural Evaluation Criteria

Table 1 defines six evaluation dimensions. The first five use the processed input as reference; original photographs assess overall workflow fidelity. Unseen surfaces require internal consistency assessment unless independent references exist.

Table 1. Dimensions of evaluation and scope of examination of architectural representation.

Dimension	Feature examined	Reference basis
Massing	Silhouette, volumes, proportions	Input image
Facade	Floor alignment, divisions, rhythm	Input image
Openings	Position, repetition, geometry	Visible windows and doors
Roof	Form, ridges, parapets, eaves, junctions	Visible roof surfaces
Details	Balconies, projections, columns, arches, mouldings, railings	Visible architectural elements
Unseen surfaces	Massing, floor, and roof continuity	Internal consistency; independent data if available

Obscured or unresolved details are excluded from scoring, as are inapplicable components. ArchiSet supports component-level evaluation of window arrangements [17]. Each dimension is rated separately on five levels: clear preservation, minor simplification, substantial alteration, partial preservation, and loss or major inconsistency.

2.5 Architectural Component Classes

Classification concerns observable components rather than entire buildings (Table 2); multiple types may coexist within one model.

Table 2. Architectural component representation classes.

Code	Representation type	Classification criteria
G	Geometry	Geometry present; no distinct texture counterpart
T	Texture	Component depicted through texture; corresponding geometry absent
G+T	Combined	Both geometry and texture/material information present
M	Missing	Visible input component lacks a meaningful output counterpart
H	Inconsistent	Contradicts visible reference or creates architectural inconsistency

Representation type and fidelity are scored separately. Classification prioritizes H, then M, then G, T, or G+T. Single-image generation necessarily completes unseen areas [5,6]; H requires a visible contradiction or internal inconsistency.

2.6 Technical Assessment

Technical assessment records valid outputs, vertices, triangles, file size, texture resolution, and generation time. Inference and export are timed separately from queue and network delays. Mesh complexity and file size require interpretation alongside architectural quality. Without validated 3D references, image-based assessment cannot establish dimensional accuracy [3]. Scores from different protocols are reported separately.

3 Findings

TRELLIS.2 produced textured GLBs for Ankara Palas and the TBMM Museum (Fig. 2). The entrance dome and roof volumes are recognizable at Ankara Palas, as are the roof profile and repeated facade openings at the TBMM Museum. These textured views show visual appearance alone and cannot establish which details are represented in the mesh.

Figure 3 compares the Hunyuan3D 2.1 and TRELLIS.2 outputs for the Bruno Taut House using the same Gemini input. The tiered roof and projecting balconies are recognizable in both outputs. Separate inspections of the untextured meshes showed that the balcony railings appeared thicker and more continuous in the Hunyuan3D output, and thinner and locally discontinuous in the TRELLIS.2 output.



Fig. 3. Bruno Taut House: (a) Gemini input [1], (b) Hunyuan3D 2.1, and (c) TRELLIS.2. Both outputs use matched scale, camera, and lighting.

3.1 Prior Evidence and Evaluation Implications

Gemini–Hunyuan3D produced textured assets for all 16 buildings. ABC2 reported preserved massing, roof form, and architectural character, with simplified details and unseen surfaces [3]. Challenges concerned ornamentation and courtyard surfaces at Ankara Palas; repetitive modules at Gazi Education Institute; courtyard depth at Doğan Apartment, Arif Paşa Apartment, and Narmanlı Han; thin supports and railings at Florya Atatürk Sea Pavilion; and ornamentation and roofs at Frej Apartment and Pera Palas [3]. These are Hunyuan3D observations.

ICICC found better color and texture retention with Qwen and simpler, stylized inputs with Gemini [2], demonstrating that preprocessing contributes to final representation. Existing Gemini inputs are therefore retained.

Earlier detail simplification [1,3] motivates separating mesh presence from textured visibility. The component classes describe representation type; fidelity ratings describe reference correspondence. O-Voxel supports geometry and PBR appearance [8,9], but the pilot does not establish performance across the dataset.

4 Discussion

4.1 Intended Uses of Architectural Representation

Recognizable assets support heritage communication, education, and visualization; surveying and conservation require verified geometry. The workflow complements metric acquisition given limited imagery [1–3]. Textured windows support recognition without ensuring area calculation or opening modification. ArchiSet's component criteria [17] and ArchiDiff's editability focus [18] reinforce this distinction. PBR appearance does not establish measured material properties and detailed meshes are not semantically structured BIM. Suitability depends on the intended operation.

4.2 Dataset Continuity and Uncertainty

The 16 buildings span several architectural forms but do not constitute a typologically comprehensive benchmark. Reusing identical inputs controls image-generation variation; direct ranking against earlier Hunyuan3D results still requires compatible production and evaluation conditions.

Uncertainty enters when Gemini completes unseen surfaces and TRELLIS.2 infers missing geometry. Agreement with processed inputs cannot validate Gemini's additions against the building. A rear facade may preserve floor rhythms without matching reality; generated surfaces remain hypotheses unless independently verified [5,6].

4.3 Limitations and Future Comparative Research

The evaluation includes two TRELLIS.2 pilot cases and one qualitative comparison with Hunyuan3D. Comparative accuracy and processing time were not measured. Incomplete 3D references constrain absolute accuracy, while independent raters and agreement analysis could address interpretive variability.

Future benchmarking could include Hunyuan3D 2.1, Step1X-3D, Hi3DGen, and SAM 3D Objects. Protocols must accommodate geometry–texture separation [10–13], normal-based generation [14], and masked inputs [15,16]. Shared geometry criteria and capability-specific texture assessment would support comparison; this experiment remains future work.

5 Conclusion

This study defines a TRELLIS.2 photo-to-3D workflow using the same 16-building dataset and Gemini Flash 2.5 isometric inputs retained from previous research. Evaluation addresses massing, facades, openings, roofs, secondary details, and unseen surfaces. The framework distinguishes visual recognizability from geometric representation through component classes G, T, G+T, M, and H, alongside separate fidelity ratings. Visible texture does not establish mesh geometry, and unseen-surface consistency does not establish correspondence to the actual building. Previous experiments support overall building recognizability but highlight limitations in fine details and unseen surfaces. The two TRELLIS.2 pilot outputs illustrate the workflow. The study does not report comparative accuracy or processing time measurements, so it cannot establish superiority over Hunyuan3D. The framework provides a basis for controlled assessment and subsequent comparison across generative models.

References

1. Aş Çemrek, H., Işıkdağ, Ü., Bekdaş, G., Niğdeli, S.M.: A Photo to 3d Workflow for Generation of LOD3 Digital Building Representations. In: International Conference on Applied Sciences and Engineering ICASE 2025, pp. 201–210, İstanbul, Türkiye (2025).
2. Aş Çemrek, H., Işıkdağ, Ü., Bekdaş, G., & Niğdeli, S. M. (2026, February). A Photo to Generative AI Workflow for Rapid 3D Heritage Representations. In International

Conference On Innovative Computing And Communication (pp. 296-305). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-28310-8_21
3. Aş Çemrek, H., Işıkdağ, Ü., Bekdaş, G., Niğdeli, S.M.: A Photo to 3D Workflow for Generation of LOD3 Digital Building Representation. ABC2 Journal of Architecture, Building, Construction, and Cities 2026(04), 26–42 (2026). https://doi.org/10.66408/abc2.2026.20
4. Münster, S., Maiwald, F., di Lenardo, I., Henriksson, J., Isaac, A., Graf, M.M., Beck, C., Oomen, J.: Artificial Intelligence for Digital Heritage Innovation: Setting up a R&D Agenda for Europe. Heritage 7(2), 794–816 (2024). https://doi.org/10.3390/heritage7020038
5. Melas-Kyriazi, L., Laina, I., Rupprecht, C., Vedaldi, A.: RealFusion: 360° Reconstruction of Any Object From a Single Image. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8446–8455 (2023). https://doi.org/10.48550/arXiv.2302.10663
6. Liu, R., Wu, R., Van Hoorick, B., Tokmakov, P., Zakharov, S., Vondrick, C.: Zero-1-to-3: Zero-shot One Image to 3D Object. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9298–9309 (2023). https://doi.org/10.48550/arXiv.2303.11328
7. Xiang, J., Lv, Z., Xu, S., Deng, Y., Wang, R., Zhang, B., Chen, D., Tong, X., Yang, J.: Structured 3D Latents for Scalable and Versatile 3D Generation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 21469–21480 (2025). https://doi.org/10.48550/arXiv.2412.01506
8. Xiang, J., Chen, X., Xu, S., Wang, R., Lv, Z., Deng, Y., Zhu, H., Dong, Y., Zhao, H., Yuan, N.J., Yang, J.: Native and Compact Structured Latents for 3D Generation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 14419–14429 (2026). https://doi.org/10.48550/arXiv.2512.14692
9. Microsoft: TRELLIS.2 Native and Compact Structured Latents for 3D Generation. Official GitHub repository. https://github.com/microsoft/TRELLIS.2 last accessed 2026/09/30.
10. Team Hunyuan3D, Yang, S., Yang, M., Feng, Y., Huang, X., Zhang, S., et al.: Hunyuan3D 2.1: From Images to High-Fidelity 3D Assets with Production-Ready PBR Material. arXiv (2025). https://doi.org/10.48550/arXiv.2506.15442
11. Tencent Hunyuan3D Team: Hunyuan3D-2.1. Official GitHub repository. https://github.com/Tencent-Hunyuan/Hunyuan3D-2.1 last accessed 2026/09/30.
12. Li, W., Zhang, X., Sun, Z., Qi, D., Li, H., Cheng, W., et al.: Step1X-3D: Towards High-Fidelity and Controllable Generation of Textured 3D Assets. arXiv (2025). https://doi.org/10.48550/arXiv.2505.07747
13. StepFun AI: Step1X-3D. Official GitHub repository. https://github.com/stepfun-ai/Step1X-3D last accessed 2026/09/30.
14. Ye, C., Wu, Y., Lu, Z., Chang, J., Guo, X., Zhou, J., Zhao, H., Han, X.: Hi3DGen: High-fidelity 3D Geometry Generation from Images via Normal Bridging. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 25050–25061 (2025). https://doi.org/10.48550/arXiv.2503.22236
15. SAM 3D Team, Chen, X., Chu, F.-J., Gleize, P., Liang, K.J., Sax, A., et al.: SAM 3D: 3Dfy Anything in Images. arXiv preprint arXiv:2511.16624v2 (2026). https://doi.org/10.48550/arXiv.2511.16624
16. Meta Facebook Research: SAM 3D Objects. Official GitHub repository. https://github.com/facebookresearch/sam-3d-objects last accessed 2026/09/30.
17. Yin, J., Zeng, P., Shen, L., Zhang, M., Zhong, J., Han, Y., Lu, S.: ArchiSet: Benchmarking Editable and Consistent Single-View 3D Reconstruction of Buildings with Specific

Window-to-Wall Ratios. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 26004–26014 (2025).
18. Yin, J., Gao, W., Li, J., Xu, P., Wu, C., Lin, B., Lu, S.: ArchiDiff: Interactive design of 3D architectural forms generated from a single image. Computers in Industry 168, 104275 (2025). https://doi.org/10.1016/j.compind.2025.104275

Risk Management Maturity in Public Transportation Digitalisation

Kudzai Mbowa[1*] and Clinton Aigbavboa[1]

[1] Sustainable Human Settlement and Construction Research Centre, Faculty of Engineering and the Built Environment, University of Johannesburg, Johannesburg, South Africa
* Corresponding author: Kudzai Mbowa; sammykmbowa@gmail.com

Abstract. Digitalisation is changing how public transportation providers collect fares, coordinate services and communicate with passengers. Dependable services require organisational capabilities that manage digital disruption. This study developed a risk management maturity framework through a structured integrative literature review. Thematic synthesis connected risks, responsibilities and practices across transportation research, cybersecurity guidance and maturity instruments. Seven dimensions were derived: governance and accountability; assets and dependencies; workforce capability; data quality and privacy; fare and supplier assurance; continuity and incident response; and passenger accessibility. Four stages distinguished reactive, documented, demonstrated and improving practice. The synthesis linked technical preparedness with commercial arrangements, staff competence and passengers' access to alternatives. Evidence from road and rail settings informed common management requirements, with mode-specific adaptation retained. The framework provides a basis for identifying management gaps and prioritising improvements, including in Southern African public transportation. The analysis distinguished shared management responsibilities from technical controls requiring specialist assessment. Operator-based testing should refine the criteria.

Keywords: public transportation; digitalisation; risk management; maturity framework; integrative review; Southern Africa.

1 Introduction

Digital technologies increasingly influence the organisation and experience of public transportation. Electronic fares, vehicle tracking, intelligent vehicle systems and real-time information connect operational decisions with digital records and external platforms [18]. Smart-card records support planning, while passenger information helps travellers manage waiting and journey uncertainty [21, 27]. These benefits depend on reliable systems and the people, agreements and procedures supporting them. A payment failure or unavailable supplier can interrupt a journey even when vehicles remain operational. This paper considers shared passenger services, including road and rail transportation, through the management functions supporting them.

Digitalisation creates continuing management responsibilities: identifying dependencies, protecting records, maintaining competence and organising alternatives. Fare-application vulnerabilities and agency-vendor research show that these responsibilities cross organisational boundaries [10, 12], while security guidance reinforces accountability [2, 8]. Risk management maturity concerns their consistent implementation and improvement. Technology adoption establishes whether a system is present; digital maturity more broadly concerns transformation capability.

Existing approaches offer different foundations. APTA's operational technology framework and the FTA's Cybersecurity Assessment Tool for Transit address preparedness; NIST organises cybersecurity outcomes around governance and related functions [3, 4, 15, 20]. DMM-TRAM assesses digital capabilities, while Tubis connects maturity with transformation risks [25, 26]. The gap addressed here is the integration of technical readiness, payment relationships, usable passenger alternatives and shared responsibilities into a service-focused assessment profile. Common management requirements can be examined across modes, although technical indicators and operational consequences require local adaptation.

This integration is relevant to Southern Africa, where public transportation can distribute decisions across authorities, operators, owners and platform providers. Research comparing Zimbabwe and South Africa links reform with institutional coordination and proposes integrated payment and GPS applications [30]. A multicity fare study includes Maputo alongside Cape Town, Kigali, Lagos and Nairobi, illustrating differing implementation contexts [31]. These sources complement passenger-information and payment studies [5, 17, 22, 24]; they inform regional interpretation rather than define a single-country population. The review asks which capabilities should be assessed, what evidence demonstrates progression, and how the resulting framework can guide improvement in public transportation.

2 Methodology

A structured integrative review combined transportation research and institutional guidance to develop a conceptual framework [23]. Search reporting followed Xiao and Watson [29], thematic synthesis drew on Braun and Clarke [11], and framework construction followed explicit maturity-model reasoning [7].

2.1 Search strategy

Targeted searches covered publicly accessible scholarly records, publisher websites, university repositories and APTA, ENISA, NIST and FTA collections, supplemented by reference checking. Search concepts combined ("public transport" OR "public transportation" OR "transit" OR "rail" OR "bus") with ("digital maturity" OR "risk management" OR "cybersecurity"). Supplementary terms included fare payment, suppliers, passenger information, digital inequality and Southern Africa. Titles, abstracts and substantive accounts established relevance; exact-title and DOI checks

established publication identity. Searches were updated on 22 October 2026. Retrieval supported conceptual integration rather than exhaustive database coverage.

2.2 Inclusion and exclusion criteria

Eligible English-language sources addressed digital passenger transportation, organisational risks, accessibility or transferable maturity principles. Transport publications covered 2011–2026, with earlier methodological works retained. Peer-reviewed research supported findings and identifiable institutional instruments supplied guidance. Sources lacking verifiable provenance or adequate supporting accounts were excluded, together with freight, private-car and construction material without a transferable passenger-service principle. Road, rail and integrated-mobility evidence was interpreted within its original context. Related reports were not treated as independent corroboration. The reference base comprises 31 substantive and four methodological publications.

2.3 Data analysis and framework development

Thematic synthesis followed familiarisation, coding, grouping and theme review [11]. Distinct evidence requirements separated the themes: assigned recovery duties indicate accountability, whereas an exercised response demonstrates continuity. Literature-supported requirements were distinguished from the proposed assessment rules [7].

Sources were selected for their relevance to digital transportation risks and maturity assessment. Reported risks, responsibilities and management practices were compared and grouped into seven dimensions according to their assessment purpose [11, 23, 29]. Table 1 links each dimension to supporting sources and observable evidence. Four maturity stages distinguish reactive, documented, demonstrated and improving practice [7].

3 Synthesis of digitalisation risks

Connected transportation services depend on communications, data and external organisations. Transit cybersecurity and mobility-as-a-service research link these dependencies with governance across actors [1, 6, 14]. ENISA’s railway guidance recommends selecting and adapting risk-management methods to organisational needs, reinforcing the importance of operational context [32]. DMM-TRAM demonstrates assessment across personnel, information, passenger flow, infrastructure and operational risk; its application found uneven maturity between functions [26]. These sources support a capability profile rather than an assumption that installing connected systems establishes readiness. The transferable principle is to identify dependencies and accountable actors, while preserving mode-specific technical assessment.

Workforce capability determines whether formal arrangements translate into action. Transit preparedness research identifies implementation gaps [9], and Ali et al. report information-technology and operational-technology competency shortages, staffing

constraints and the importance of experiential learning across transportation subsectors [34]. This strengthens the case for assessing role-specific demonstrations and continuity of expertise. Supplier alignment research adds responsibility for maintenance, escalation and access to operational records [10]. Neither training attendance nor a signed contract demonstrates that staff and partners can restore a service; their coordinated response must be evidenced.

Passenger information connects operational reliability with journey decisions. Smart-card research establishes the analytical value of transaction records, while South African information-system research highlights integration and stakeholder requirements [17, 21]. Bus information research reports waiting-time benefits [27], whereas a Boston commuter-rail study found shorter self-reported usual waits among information users but modest wider benefits [35]. This comparison cautions against assuming uniform effects across modes. Physical network connections do not establish digital interoperability [16]; Nairobi's Digital Matatu illustrates adaptation of data standards to actual services [28]. Assessment should therefore examine record quality, update responsibility and usability, rather than channel availability alone.

Digital access is also uneven. Durand et al.'s review explains why engagement with technology affects access to transportation services [33]; subsequent interview research shows how self-service arrangements can increase reliance on assistance from others [13]. Commuter application research provides complementary evidence about actual use [19]. These findings support passenger-tested alternatives for information and payment, particularly during disruption. An available application or functioning server cannot establish that a traveller can complete the required task independently.

Fare systems combine these technical, organisational and accessibility issues. African studies identify acceptability, resources, settlement and stakeholder incentives as implementation concerns [5, 22, 24, 31]. Application-security research demonstrates a distinct protection requirement [12]. Payment assurance should consequently examine reconciliation and supplier arrangements alongside security and usable alternatives. Tubis further shows why transformation risks must be interpreted in relation to maturity [25]: postponing digital change can retain operational weaknesses. The synthesis therefore supports evaluating the management of both existing and proposed digital services.

4 Risk management maturity framework

The framework assesses a defined passenger service, including shared responsibilities, to identify management capability and improvement priorities. Application to different modes requires service-specific evidence and complements specialist technical assessment.

4.1 Assessment dimensions

Table 1 connects the seven dimensions to their evidential basis and observable assessment material. Accurate records do not establish privacy controls; assigned

recovery duties do not demonstrate successful recovery. Data quality and privacy, and fare and supplier assurance, therefore retain explicit subdomains within their respective dimensions. The framework judges management practice rather than the number of technologies installed.

Table 1. Assessment dimensions and supporting evidence

Dimension and sources	Assessment question and observable evidence
Governance and accountability [2, 10]	Who can decide and act? Named service owner, assigned supplier responsibilities and recorded risk decisions.
Assets and dependencies [6, 14, 32]	What must remain available? Service-to-asset inventory, communication links and external support dependencies.
Workforce capability [6, 9, 34]	Can relevant staff apply the arrangements? Role-specific demonstrations, induction and cover for key personnel.
Data quality and privacy [12, 17, 21]	Are records reliable and appropriately handled? Correction logs, authorised access and access reviews.
Fare and supplier assurance [10, 24, 31]	Are payments and support arrangements dependable? Reconciliation, settlement, recurring costs, support and data-access terms.
Continuity and incident response [14, 15]	Can the service continue and recover? Exercised fallback procedures, escalation and reconciliation evidence.
Passenger accessibility [13, 19, 33]	Can passengers use an alternative? Passenger-tested information and payment channels, with recorded access problems.

4.2 Maturity stages and assessment rules

Table 2 distinguishes four stages through evidence of established, applied and improving practice. Existing instruments demonstrate staged assessment but serve different purposes [3, 4, 26]. The proposed structure is a framework design choice, supported by procedures, records and practical demonstrations; it does not convert existing cybersecurity or digital maturity scores.

Table 2. Maturity stages and illustrative fare-outage evidence

Stage	Required practice	Illustrative evidence
1 Reactive	Response relies on individual knowledge; consistent arrangements are not demonstrated.	Staff improvise boarding and contact support informally.
2 Documented	Responsibilities, dependencies and procedures are recorded and communicated.	An approved boarding alternative identifies activation authority.
3 Demonstrated	Records and exercises show that relevant staff and partners apply the procedures.	An exercise demonstrates activation, passenger updates and reconciliation.
4 Improving	Findings lead to documented changes whose effectiveness is subsequently checked.	An outage review produces changes verified through a further exercise.

Dimensions are assessed separately for a defined service and period. Assign the highest evidenced stage: stages 3–4 require documented arrangements, and stage 4 requires

verified improvement. Record unavailable evidence as “not assessed”, with reasons. Results form an ordinal capability profile, without arithmetic averaging or certification. The fare-outage examples are hypothetical.

5 Discussion and implications

The framework’s contribution is the connection between organisational risk practices and the continuity of the passenger journey. Cybersecurity guidance organises protection and preparedness, while information and fare studies establish why usability, incentives and coordination also matter [5, 15, 17]. Combining these perspectives prevents a technically secure platform from being treated as sufficient evidence of dependable service. It also makes improvement actionable: an operator can identify whether a gap concerns authority, staff response, supplier arrangements or passengers’ alternatives.

For Southern African application, assessment should begin with one critical service and identify where decisions and evidence are held. Responsibilities may span authorities, operators, associations and providers [22, 24]. An outage exercise can test agreed arrangements within available resources. Zimbabwe–South Africa research and the multicity fare study, including Maputo, support examining institutional fit [30, 31]. Road, rail and cross-border services require different dependency maps and fallback procedures; shared framework dimensions do not imply identical technical requirements.

The framework integrates secondary evidence and proposes assessment criteria; it does not measure operator performance. Its selective, geographically uneven evidence requires contextual adaptation. Stakeholder review and field pilots should test clarity, evidence availability and assessor agreement, followed by relationships with service recovery and passenger access [7, 23].

6 Conclusion

Public transportation digitalisation requires capabilities that sustain dependable and accessible services. The integrative review connects seven dimensions with four descriptions of practice, from reactive response to demonstrated improvement. The framework offers a basis for identifying management gaps and selecting feasible actions across providers and their partners. Application in Southern Africa requires locally appropriate evidence and passenger alternatives. Operator-based evaluation should refine the criteria and establish their practical assessment value.

References

1. Alderete Peralta, A., Balta-Ozkan, N., & Li, S.: The road not taken yet: A review of cyber security risks in mobility-as-a-service (MaaS) ecosystems and a research agenda. Res. Transp. Bus. Manag., 56, 101162. doi:10.1016/j.rtbm.2024.101162 (2024).

2. American Public Transportation Association: Security risk assessment methodology for public transit (APTA SS-SIS-S-017-21). (2021).
3. American Public Transportation Association: Operational Technology Cybersecurity Maturity Framework (OT-CMF) overview (APTA SS-CCS-RP-006-23). (2023).
4. American Public Transportation Association: Operational Technology Cybersecurity Maturity Framework (APTA SS-TCS-RP-003-26). (2026).
5. Aruho, A. T., Behrens, R., & Zuidgeest, M.: Cashless fare collection systems acceptability in the paratransit sector in Cape Town. Transp. Res. Procedia, 89, 650–668. doi:10.1016/j.trpro.2025.05.088 (2025).
6. Barbeau, S. J., et al.: Enhancing cybersecurity in public transportation (Project BDV25-977-51). University of South Florida, Center for Urban Transportation Research. (2019).
7. Becker, J., Knackstedt, R., & Pöppelbuß, J.: Developing maturity models for IT management—A procedure model and its application. Bus. Inf. Syst. Eng., 1(3), 213–222. doi:10.1007/s12599-009-0044-5 (2009).
8. Beecroft, M.: The future security of travel by public transport: A review of evidence. Res. Transp. Bus. Manag., 32, 100388. doi:10.1016/j.rtbm.2019.100388 (2019).
9. Belcher, S., et al.: Does the transit industry understand the risks of cybersecurity and are the risks being appropriately prioritized? (Report 25-04). Mineta Transportation Institute. doi:10.31979/mti.2025.2405 (2025).
10. Belcher, S., et al.: Aligning the transit industry and their vendors in the face of increasing cyber risk: Recommendations for identifying and addressing cybersecurity challenges. Mineta Transportation Institute. doi:10.31979/mti.2022.2113 (2022).
11. Braun, V., & Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol., 3(2), 77–101. doi:10.1191/1478088706qp063oa (2006).
12. Dennis, K., et al.: Cybersecurity vulnerabilities in mobile fare payment applications: A case study. Transp. Res. Rec., 2674(11), 616–624. doi:10.1177/0361198120945982 (2020).
13. Durand, A., et al.: “Who can I ask for help?”: Mechanisms behind digital inequality in public transport. Cities, 137, 104335. doi:10.1016/j.cities.2023.104335 (2023).
14. European Union Agency for Network and Information Security: Cyber security and resilience of intelligent public transport: Good practices and recommendations. doi:10.2824/778225 (2015).
15. Federal Transit Administration: Cybersecurity resilience assessment tool to enhance public confidence in transit (FTA Report No. 0250). United States Department of Transportation. (2023).
16. Gumbo, T., & Moyo, T.: Exploring the interoperability of public transport systems for sustainable mobility in developing cities: Lessons from Johannesburg Metropolitan City, South Africa. Sustainability, 12(15), 5875. doi:10.3390/su12155875 (2020).
17. Hlophe, A., Afolayan, A., & Daramola, O.: Integrated real-time information system for public commuting: Perspectives of stakeholders in South Africa. Int. J. Transp. Dev. Integr., 8(1), 31–48. doi:10.18280/ijtdi.080104 (2024).
18. Mbowa, K., et al.: An overview of key emerging technologies transforming public transportation in the Fourth Industrial Revolution era. IOP Conf. Ser. Mater. Sci. Eng., 1107, 012169. doi:10.1088/1757-899X/1107/1/012169 (2021).
19. Niemand, C. J. P., & Chauke, H.: The use of mobile applications by public transport commuters in Gauteng, South Africa. S. Afr. J. Inf. Manag., 19(1), a782. doi:10.4102/sajim.v19i1.782 (2017).
20. Pascoe, C., Quinn, S., & Scarfone, K.: The NIST Cybersecurity Framework (CSF) 2.0 (NIST CSWP 29). National Institute of Standards and Technology. doi:10.6028/NIST.CSWP.29 (2024).

21. Pelletier, M.-P., Trépanier, M., & Morency, C.: Smart card data use in public transit: A literature review. Transp. Res. Part C, 19(4), 557–568. doi:10.1016/j.trc.2010.12.003 (2011).
22. Schalekamp, H., McLaren, M., & Behrens, R.: Exploring cashless fare collection in the context of urban public transport reform in South Africa. Proceedings of the 36th Southern African Transport Conference, 610–625. (2017).
23. Snyder, H.: Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333-339. doi:10.1016/j.jbusres.2019.07.039 (2019).
24. Tinka, A. A., & Behrens, R.: Cashless fare collection in Sub-Saharan African paratransit: A review of experiences. Proceedings of the 38th Southern African Transport Conference. (2019).
25. Tubis, A. A.: A framework for managing digital transformation risks in transport systems: Linking digital maturity and risk categories. Applied Sciences, 16(1), 206. doi:10.3390/app16010206 (2026).
26. Tubis, A. A., Rydlewski, M., & Skiba, M.: Digital maturity in tram systems: Model design and practical application. Bull. Pol. Acad. Sci. Tech. Sci., 73(6), e155034. doi:10.24425/bpasts.2025.155034 (2025).
27. Watkins, K. E., et al.: Where Is My Bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders. Transp. Res. Part A, 45(8), 839–848. doi:10.1016/j.tra.2011.06.010 (2011).
28. Williams, S., et al.: The digital matatu project: Using cell phones to create an open source data for Nairobi's semi-formal bus system. J. Transp. Geogr., 49, 39–51. doi:10.1016/j.jtrangeo.2015.10.005 (2015).
29. Xiao, Y., & Watson, M.: Guidance on conducting a systematic literature review. J. Plan. Educ. Res., 39(1), 93–112. doi:10.1177/0739456X17723971 (2019).
30. Cirella, G. T., Mtizi, C., & Iyalomhe, F. O.: Public transportation solutions in Southern Africa: case study Zimbabwe and South Africa. Transport Economics and Logistics, 79, 77–86. doi:10.26881/etil.2018.79.06 (2018).
31. Arroyo Arroyo, F., et al.: Innovation in Fare Payment Systems for Public Transport in African Cities. Report 168374. World Bank, Washington, DC. (2022).
32. European Union Agency for Cybersecurity: Railway cybersecurity—Good practices in cyber risk management. ENISA (2021).
33. Durand, A., et al.: Access denied? Digital inequality in transport services. Transp. Rev., 42(1), 32–57. doi:10.1080/01441647.2021.1923584 (2022).
34. Ali, A., et al.: Strengthening the transportation cybersecurity workforce: A mixed-methods analysis of workforce development, organizational preparedness and cross-sector collaboration. Future Transp., 6(4), 140. doi:10.3390/futuretransp6040140 (2026).
35. Brakewood, C., et al.: An analysis of commuter rail real-time information in Boston. J. Public Transp., 18(1), 1–20. doi:10.5038/2375-0901.18.1.1 (2015).

Performance and robustness assessment of optimized tuned mass dampers in soft-story structures

Farah Alnayhoum[1], Sinan Melih Nigdeli[1] and Gebrail Bekdaş[1]

[1]Department of Civil Engineering, Istanbul University - Cerrahpaşa, 34320 Avcılar, Istanbul, Turkey

Email: f.alnayhoum@iuc.ogr.edu.tr, melihnig@iuc.edu.tr, bekdas@iuc.edu.tr

Abstract. TMDs are passive control systems that are widely used to mitigate seismic responses and vibrations in structures by absorbing lateral excitations. Structures with a soft story usually suffer damage in that story, while the upper stories remain more stable. Optimal damper parameters are typically derived using analytical or metaheuristic optimization algorithms, both of which assume perfectly tuned devices. In practice, manufacturing tolerances, material degradation, and environmental changes cause damper parameters to deviate from their theoretical optimal values; this phenomenon is known as "Mistuning". Therefore, this study investigates the robustness of single and multiple tuned mass damper systems optimized using both the Den-Hartog technique and the Jaya metaheuristic algorithm against parameter mistuning in a five-story building model. The analysis is conducted using 44 far-field ground motion records from FEMA P-695, and intentional perturbations of ±5%, ±10%, ±20%, and ±30% relative to their optimal values in order to evaluate relative displacement and acceleration responses through numerical time-history analysis.

Keywords: Tuned mass damper, mistuning, robustness, metaheuristic algorithms, soft-story structures.

1 Introduction

Structures containing soft stories are considered as one of the structures that may suffer damage under earthquake excitations because of the concentration of lateral deformations at these stories. According to ASCE, stories with stiffness less than 70% of the structure’s upper stories' stiffness are considered as soft stories[1]. Typically, due to the placement of open ground floors, parking areas, or other architectural reasons, the soft story is located on the first floor. Field studies following the Kahramanmaraş earthquake in 2023 showed that the presence of the soft story was one of the primary reasons for building collapse [2].

Tuned mass dampers are passive control devices that mitigate the response of structures under lateral excitations by tuning extra mass attached by a damper and a spring close to

the structure's natural frequency. Den Hartog proposed equations for optimum frequency and optimum damping ratio for an undamped system under harmonic excitation[3].

Sadek et al. [4] studied the parameters of damped structures under seismic excitation, and Rana and Soong [5] examined the effect of detuning. Kareem and Kline proposed using MTMD in parallel configuration to widen the frequency band caused by the STMD[6]. Joshi and Jangid also derived MTMD parameters for base-excited damped systems [7].

Because a closed-form solution is not available for most realistic cases, metaheuristic algorithms, such as Harmony Search (HS) [8], Bat algorithm (BA) [9], and Jaya Algorithm (JA)[10], that needs no specific parameters make it easy to implement, have become standard tools for TMD design. Jaya algorithm was used to tune TMDs in torsionally irregular structures[11].

In practice, manufacturing tolerances, material degradation, and environmental changes move the damper parameters away from their optimum values; this is known as mistuning. Pellizzari et al. proposed a robust design of MTMDs in order to show that deterministic optimization is inadequate for systems as sensitive as TMDs[12]. Matta found that TMD effectiveness drops when the structural stiffness changes, but that it is preserved when stiffness uncertainty is included in the design[13].

In this study, a five-story shear building subjected to 44 far-field earthquake excitations with three first-story stiffness ratios with perturbations of ±5%, ±10%, ±20%, and ±30% relative to STMD and MTMD optimal values is studied and analysed in order to evaluate relative displacement and acceleration responses through numerical time-history analysis.

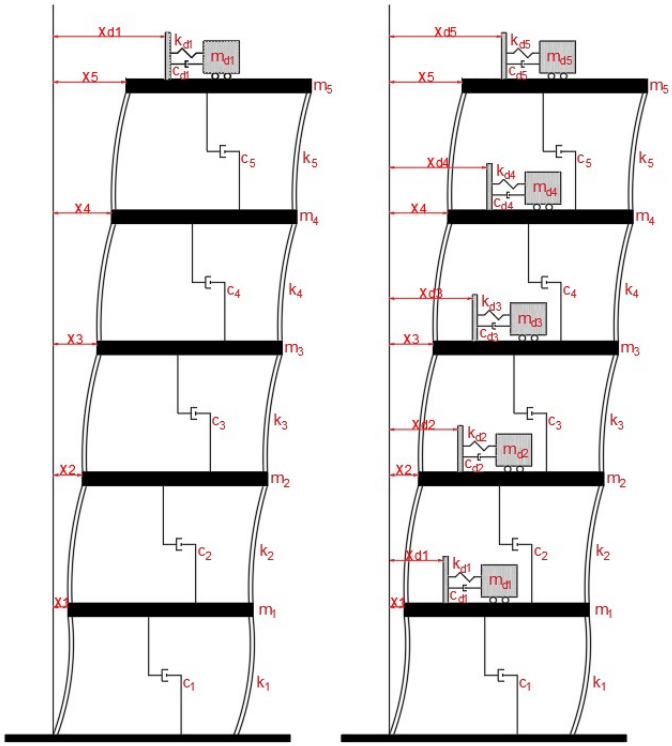
2 Methodology

This study proposes a five-story shear building that is exposed to a set of 44 far-field earthquake excitation records from FEMA P-695 [14]. Two main cases are studied: the first case consists of a single TMD placed on the top roof of the structure, while the second case consists of multiple TMDs placed at each floor, as shown in Figure 1.

m, k, and c represent the structure parameters as mass, stiffness, and damping coefficient, where m_d, k_d, and c_d represent the TMD parameters. x,$x_d$represent the structure response and the TMD response, respectively.

The structure has a mass of 2924 kg lumped at each story and a typical story stiffness of $5.7x10^6$ N/m. The soft story is modelled by reducing the stiffness by a certain ratio. Three different ratios of stiffness are used and are illustrated in Table 1.

TMD parameters were established using both Den Hartog and Jaya metaheuristic algorithm techniques. The governing equation of motion for both STMD and MTMD cases is given in Equation 1. The two TMD layouts are compared with the same total damper mass using μ of 5%.



(a) Case 1 (b) Case 2

Fig.1. Shear structure model (a) STMS at the roof; (b) MTMD with one damper on each floor

$$M\ddot{x} + C\dot{x} + Kx = -M\{1\}\ddot{x}_g \quad (1)$$

Where *M*,*C*, and *K* represent the mass, damping, and stiffness matrices. $-M\{1\}\ddot{x}_g$ represents the inertial forces due to ground motion. *x* represents the displacement of the floors and dampers relative to the ground, $\ddot{x}_g$ represents the ground acceleration.

Rayleigh technique was also used to determine the damping coefficients according to Equation 2.

$$[c] = \alpha[m] + \beta[k] \quad (2)$$

Where α and β represent Rayleigh damping coefficients.

Table 1. Dynamic characteristics of the building models

$\frac{K_1}{K}$	T_1 s	T_2 s	α $1/s$	β s
1.0	0.500	0.171	0.936	0.0020
0.7	0.538	0.181	0.875	0.0022
0.5	0.585	0.190	0.810	0.0023

2.1 Den Hartog process:

Each damper is tuned to the first mode using Den Hartog expressions[15]. Damping ratio, frequency, mass, and stiffness were obtained as the TMD parameters using the following equations:

$$f_{opt} = \frac{1}{1+\mu} \tag{3}$$

$$\zeta_{d,opt} = \sqrt{\frac{3\mu}{8(1+\mu)}} \tag{4}$$

Whereas, f_{opt}, μ, and $\zeta_{d,opt}$ represent the optimal frequency, the mass ratio, and the optimal damping ratio, respectively.

2.2 Jaya algorithm process:

Jaya algorithm moves each candidate solution toward the best solution and away from the worst one [10]. The design variables were specified as the frequency and damping ratio of each damper with specified limits, while the damper masses are fixed. The objective function was chosen to be the mean of peak inter-story drift over 44 records, normalized by the corresponding value of the uncontrolled building. The main equation of Jaya algorithm is shown in equation 5.

$$X_{i,new} = X_{i,j} + r_1 \left(X_{i,g_{best}} - \left| X_{i,j} \right| \right) - r_2 \left(X_{i,g_{worst}} - \left| X_{i,j} \right| \right) \tag{5}$$

The population size was 20, with 100 iterations for STMD and 400 iterations for MTMD.

2.3 Mistuning scheme:

A sensitivity analysis is conducted by perturbing the parameters of the dampers. The damper stiffness k_d, damping coefficient c_d, and mass m_d are each varied individually, with the other parameters kept at their optimum values by ratios of ±5%, ±10%, ±20%, and ±30%. For the MTMD system, identical perturbations are applied across all five dampers. Using these parameters and percentages, the peak response values are computed and averaged over the 44 records.

3 Results:

3.1 Uncontrolled response

The presence of the soft story results in an increase in the peak first-story drift from 19.5 mm to 27.2 mm with a softening ratio of 0.7 and to 35.1 mm with a softening ratio of 0.5.

3.2 Optimum damper parameters

For STMD cases, Jaya algorithm tunes the STMD close to the Den Hartog frequency but with lower damping. Whereas in MTMD cases, it spreads the five dampers over a band of frequencies instead of tuning them all to only one value. The optimum values are illustrated in Table 2. And Figure 2 shows the convergence of Jaya algorithm, where all runs converge within about 10 iterations for STMD , and about 40 to 120 iterations for MTMD.

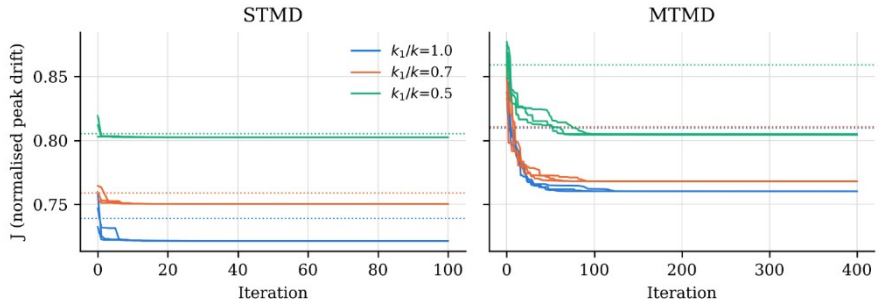


Fig. 2. Convergence of Jaya algorithm

Table 2. Optimum damper parameters

k_1/k	TMD	f	ζ_d	T_d s	K_d N/m	C_d N.s/m
1.0	S-DH	0.918	0.175	0.545	97325.63	2954.27
	S-JA	0.958	0.097	0.522	105990.2	1711.25
	M-DH	0.952	0.134	0.525	20940.57	467.631
	M-JA	0.658	0.05	0.760	9984.935	120.822
		0.884		0.565	18051.98	162.456
		0.915		0.546	19330.41	168.110
		0.930		0.537	19961.43	170.832
		1.108		0.451	28318.89	203.475
0.7	S-DH	0.923	0.170	0.582	85092.62	2679.98
	S-JA	0.988	0.134	0.544	97544.84	2269.85
	M-DH	0.952	0.134	0.564	18118.34	434.979
	M-JA	0.712	0.05	0.755	10116.46	121.615
		0.730		0.736	10651.43	124.789
		0.742		0.724	10989.07	126.751
		0.973		0.552	18916.76	166.301
		1.029		0.522	21155.84	175.868
0.5	S-DH	0.928	0.164	0.631	72524.71	2394.20
	S-JA	0.882	0.123	0.664	65493.58	1708.11
	M-DH	0.952	0.134	0.615	15279.82	399.455
	M-JA	0.792	0.05	0.738	10572.61	124.326
		0.814		0.718	11168.82	127.784
		0.833		0.702	11699.32	130.783
		1.031		0.567	17909.32	161.812

	1.114	0.525	20916.85	174.872

3.3 Robustness to mistuning:

Table 3 and Table 4 show the reduction in peak inter-story drift and peak roof acceleration under mistuning of the damper stiffness k_d, which was the most critical parameter. The STMD outperformed the MTMD because the lower-floor dampers move little in the first mode; with equal total mass, the effective mass ratio is 0.05 for the MTMD against 0.078-0.089 for the STMD. Reduction of peak inter-story drift and reduction of peak roof acceleration plots are given in Figs. 3 and 4, respectively.

Table 3. Reduction percentage in peak inter-story drift

k_1/k	TMD	Perfectly tuned	Worst ±10%	Worst ±30%	Maximum loss
1.0	S-DH	26.1	25.1	22.2 (−30 %)	3.9
	S-JA	27.9	26.2	20.2 (+30 %)	7.7
	M-DH	19.0	17.4	12.5 (+30 %)	6.5
	M-JA	24.0	21.4	12.3 (+30 %)	11.7
0.7	S-DH	24.1	23.3	20.4 (−30 %)	3.7
	S-JA	25.0	24.3	20.8 (+30 %)	4.2
	M-DH	18.9	17.9	14.9 (−30 %)	4.0
	M-JA	23.2	20.8	13.1 (−30 %)	10.1
0.5	S-DH	19.4	19.0	16.9 (+30 %)	2.5
	S-JA	19.7	18.9	15.2 (−30 %)	4.5
	M-DH	14.1	13.6	11.4 (+30 %)	2.7
	M-JA	19.6	17.9	12.9 (−30 %)	6.7

Table 4. Reduction percentage in peak acceleration

k_1/k	TMD	Perfectly tuned	Worst ±10%	Worst ±30%	Maximum loss
1.0	S-DH	25.6	23.9	19.8 (−30 %)	5.8
	S-JA	27.6	25.2	19.4 (−30 %)	8.2
	M-DH	20.5	18.8	14.3 (−30 %)	6.2
	M-JA	22.3	20.1	14.5 (−30 %)	7.8
0.7	S-DH	23.1	21.6	18.4 (−30 %)	4.7
	S-JA	24.7	23.3	19.6 (−30 %)	5.1
	M-DH	19.2	17.4	13.3 (−30 %)	5.9
	M-JA	19.1	16.2	9.5 (−30 %)	9.6
0.5	S-DH	18.7	18.5	16.1 (−30 %)	2.6
	S-JA	18.1	17.2	13.6 (−30 %)	4.5
	M-DH	14.9	14.5	12.0 (−30 %)	2.9
	M-JA	17.1	15.4	10.7 (−30 %)	6.4

3.4 Damper stroke

For MTMD cases, the Jaya algorithm requires the largest damper stroke, 0.22 to 0.27 m, compared with 0.11 to 0.17 m for other designs, as a direct result of its 5% damping. The detailed damper stroke values are illustrated in Table 5.

Table 5. Mean peak damper stroke

k_1/k	STMD		MTMD	
	Den Hartog	Jaya	Den Hartog	Jaya
1.0	0.111	0.135	0.142	0.222
0.7	0.123	0.126	0.153	0.254
0.5	0.138	0.165	0.167	0.267

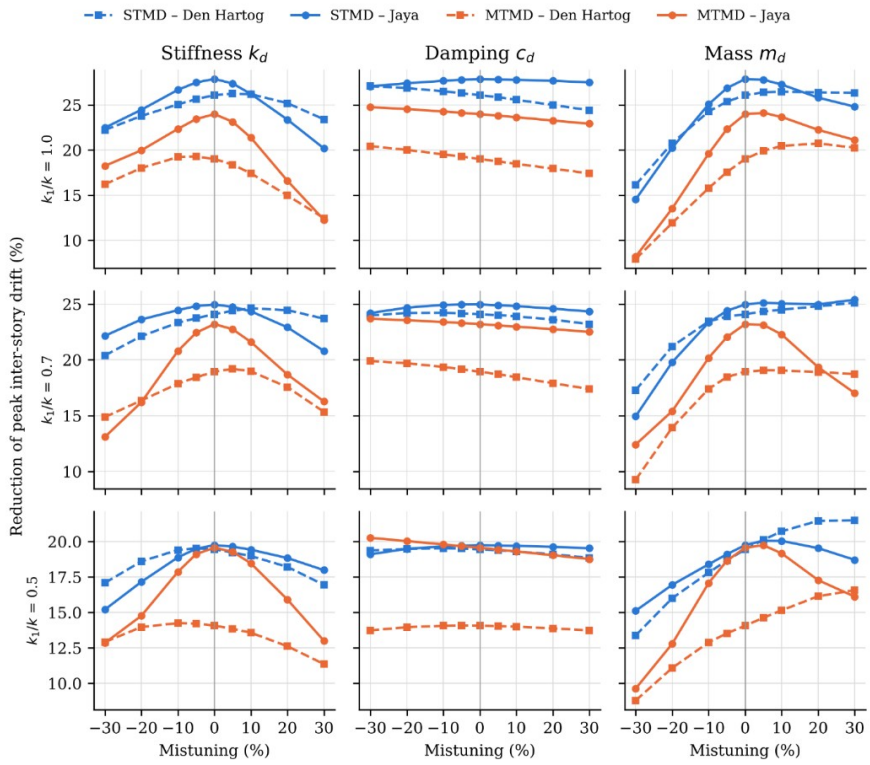


Fig. 3. Reduction of peak inter-story drift

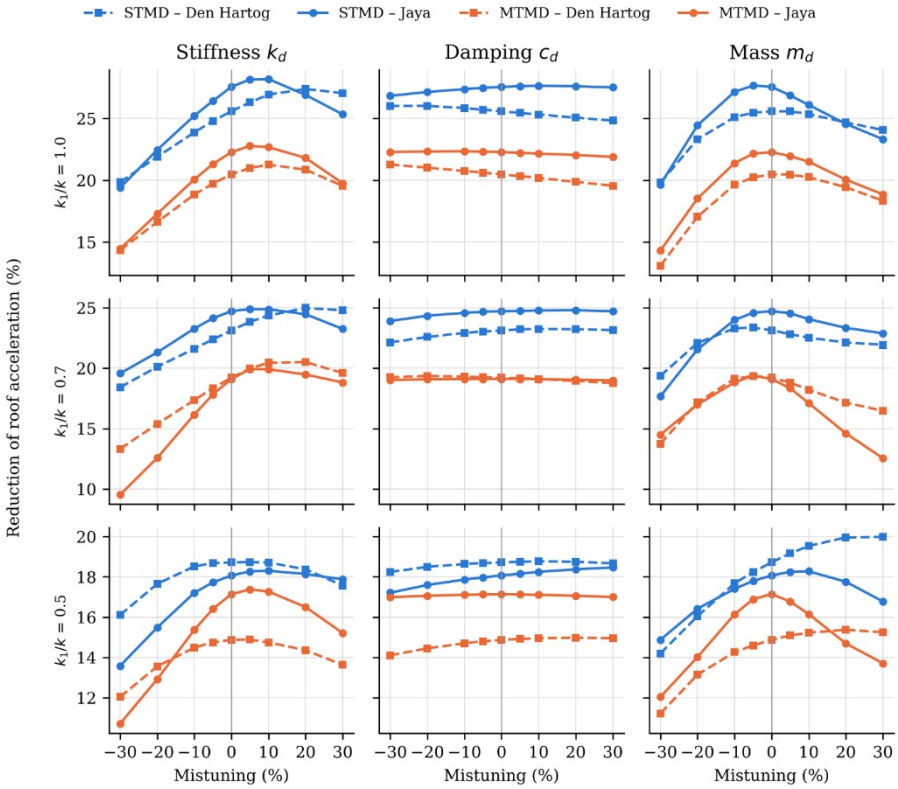


Fig. 4. Reduction of peak roof acceleration

4 Discussion:

The results show a trade-off: the design that performs the best when perfectly tuned is the one that degrades fastest when mistuned. MTMD case using the Jaya algorithm gains its advantages by spreading lightly damped dampers over a small frequency range. This range covers the first mode response well; however, it has sharp edges. A uniform stiffness error shifts the whole range away from ω_1, and with only 5% damping, each damper has little bandwidth to compensate. Den Hartog designs carry 13 to 18% damping, which flattens the response curve around the optimum and makes them tolerant of frequency error. The Jaya STMD with 10 to 13 % damping lies between the two in robustness.

Frequency errors, whether from stiffness or mass, dominate the loss of performance; damping errors of ±30% are almost irrelevant. Design tolerances and monitoring should therefore target the damper frequency.

With the presence of the soft story, the dampers become less effective in percentage terms. This is due to the concentration of deformations in that floor, whereas dampers target the building's overall first vibration mode. However, the absolute reduction in

drifts saved in these stories is still larger. Also, the soft story reduces the performance gap between different layouts.

In conclusion, while Den Hartog formulas remain a robust and near-optimal choice for a single roof damper, metaheuristic optimization for distributed dampers only succeeds when frequency errors are kept within ±10% and optimization objectives explicitly incorporate robustness and stroke limits to prevent fragile, lightly damped designs.

5 Conclusion:

This study compared single and multiple TMDs, designed by both the Den Hartog and Jaya algorithm, in a five-story shear building with a soft first story under 44 far-field records, and tested their robustness to ±5% to ±30% parameter errors. The main conclusions are:

- With equal total mass, a single roof TMD reduced the peak inter-story drift (19.4%-27.9%) more than five floor-mounted TMDs (14.1%-24%).
- Jaya algorithm improved on Den Hartog the most using MTMDs by spreading the damper frequencies and using low damping.
- The presence of the soft story reduces the best drift reduction from 27.9% to 19.7%.
- Using the Jaya algorithm with MTMD was the least robust design at ±20% - ±30% stiffness errors, and also required the largest strokes 0.22- 0.27 m).
- Using Den Hartog with STMD was the most robust design.
- None of the mistuned configurations amplifies the response of the uncontrolled building.

References

[1] *Minimum Design Loads and Associated Criteria for Buildings and Other Structures*. Reston, VA: American Society of Civil Engineers, 2021. doi: 10.1061/9780784415788.

[2] E. Çelebi and O. Kırtel, “Seismic performance of reinforced concrete framed buildings with ribbed slabs at the affected region by 2023 Kahramanmaraş earthquakes,” *Bulletin of Earthquake Engineering*, vol. 23, no. 9, pp. 3623–3646, Jul. 2025, doi: 10.1007/s10518-025-02191-0.

[3] D. J. P. Hartog, *Mechanical Vibrations*, 4th ed. McGraw-Hill Book Company, 1956.

[4] F. Sadek, B. Mohraz, A. W. Taylor, and R. M. Chung, “A method of estimating the parameters of tuned mass dampers for seismic applications,” *Earthq. Eng. Struct. Dyn.*, vol. 26, no. 6, pp. 617–635, 1997, doi: 10.1002/(SICI)1096-9845(199706)26:6<617::AID-EQE664>3.0.CO;2-Z.

[5] R. Rana and T. T. Soong, “Parametric study and simplified design of tuned mass dampers,” *Eng. Struct.*, vol. 20, no. 3, pp. 193–204, 1998.

[6] A. Kareem and S. Kline, “Performance of Multiple Mass Dampers under Random Loading,” *journal of structural engineering*, pp. 348–361, 1995.

[7] A. S. Joshi and R. S. Jangid, “OPTIMUM PARAMETERS OF MULTIPLE TUNED MASS DAMPERS FOR BASE-EXCITED DAMPED SYSTEMS,” *J. Sound Vib.*, vol. 202, no. 5, pp. 657–667, 1997.

[8] G. Bekdaş and S. M. Nigdeli, "Estimating optimum parameters of tuned mass dampers using harmony search," *Eng. Struct.*, vol. 33, no. 9, pp. 2716–2723, Sep. 2011, doi: 10.1016/j.engstruct.2011.05.024.

[9] G. Bekdaş, S. M. Nigdeli, and X. S. Yang, "A novel bat algorithm based optimum tuning of mass dampers for improving the seismic safety of structures," *Eng. Struct.*, vol. 159, pp. 89–98, Mar. 2018, doi: 10.1016/j.engstruct.2017.12.037.

[10] R. Venkata Rao, "Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems," *International Journal of Industrial Engineering Computations*, vol. 7, no. 1, pp. 19–34, Dec. 2016, doi: 10.5267/j.ijiec.2015.8.004.

[11] A. Mubuli, S. M. Nigdeli, and G. Bekdaş, "Optimum Design of Tuned Mass Dampers for Structures with Torsional Irregularity via Jaya Algorithm," *Journal of Vibration Engineering and Technologies*, vol. 12, no. 2, pp. 2279–2293, Feb. 2024, doi: 10.1007/s42417-023-00979-5.

[12] F. Pellizzari, G. C. Marano, A. Palmeri, R. Greco, and M. Domaneschi, "Robust optimization of MTMD systems for the control of vibrations," *Probabilistic Engineering Mechanics*, vol. 70, Oct. 2022, doi: 10.1016/j.probengmech.2022.103347.

[13] E. Matta, "Seismic effectiveness and robustness of tuned mass dampers versus nonlinear energy sinks in a lifecycle cost perspective," *Bulletin of Earthquake Engineering*, vol. 19, no. 1, pp. 513–551, Jan. 2021, doi: 10.1007/s10518-020-00973-2.

[14] FEMA. (2009). Quantification of Building Seismic Performance Factors (FEMA P-695). Prepared by the Applied Technology Council for the Federal Emergency Management Agency, Washington, D.C.

[15] D. J. P. Hartog, *Mechanical Vibrations*, Fourth edition. McGraw-Hill, 1956.

Physics-Informed Learning of Anisotropic Epidemic Dynamics from Chicago Bike-Share Mobility

Rasha Almshekhs[1], Suhaib Abdul Sattar Shukur[2]

[1] Department of Mechanical and Aerospace Engineering, University of Dayton, Dayton, OH 45469, USA
[2] State Company for Gas Filling and Services, Ministry of Oil, Republic of Iraq

Abstract

For analysis, the Chicago COVID-19 dataset and 3.05 million Divvy rides in 46 ZCTAs spanning 38 weeks were applied to estimate the learned anisotropic SIRS model, with twelve weeks set aside for validation purposes. In the process, the estimated model attained an RMSE of 145.63 cases per 100,000 population, performing worse than the isotropic SIRS model, persistence and transmission-only approaches. While there was significant transport inequality, neither Turing nor Hopf instability was observed during the 200 sensitivity analyses conducted.

Keywords: epidemic dynamics; anisotropic diffusion; bike-share mobility; graph Laplacian; physics-informed learning; SIRS; modal stability.

1 Introduction

Spatial disease spread is a consequence of the interplay of local epidemiological processes and human mobility. Whereas classical compartment models consider homogeneously mixed populations, reaction-diffusion models account for population spatial movement, such that states of the disease evolve spatially and temporally [1–8].

Diffusion may affect stability properties of diseases, leading to Turing patterns, Hopf bifurcations, and nonhomogeneous steady states [3–11]. Network dynamics are affected by graph topology, cross-diffusion, delay terms, and heterogeneity in causing instability in spatial modes. Additional works have shown that nonlinearities, stochastic perturbations, reinfection, and anisotropic diffusion can change the instability conditions [12–20]. Yet, the direction of transportation is typically prescribed mathematically and not derived from human movements.

On the other hand, mechanistic epidemic models have become increasingly refined with real surveillance data being used for estimating transmission dynamics and intervention impact [21-22]. Mobility studies have also shown relationships between human mobility and COVID-19 transmission, as OD network models enhance the representation of interregional disease transmission [23]. However, mobility networks are often pre-defined, and prediction rather than analysis of the stability of transportation dynamics is a key focus.

The physics-informed ML approach combines mechanistic modeling and prediction with machine learning. Current ML approaches integrate epidemiological equations within neural network and graph models [22-24]; yet, most of them emphasize state estimation and prediction rather than spatial transport operator learning.

This paper fills this gap by proposing a physics-guided approach that learns anisotropic epidemic transport from public urban mobility data. The directional origin-destination flows are mapped to a temporal symmetric conductance graph, with its Laplacian being embedded into a locally-conservative SIRS reaction-diffusion system. The performance of the proposed method is compared to geographic diffusion, static and raw mobility models, persistence, and a graph neural network baseline on temporally withheld data. In the end, spectral analysis is conducted to check if the learned dynamics have the potential to be diffusive Turing or Hopf unstable.

2 Materials and Methods

2.1 Study design and public data

We utilized the epidemiological weeks running from Sunday to Saturday for the period April 5 to December 26, 2020 based on the case and testing data of Chicago (City of Chicago, n.d.a). There were multiple timeouts on attempts to get the ride-hailing OD aggregations from the city, and therefore, following our prespecified contingency plan, we used the public Divvy bike rides as a mobility proxy and not as a proxy for the overall amount of traveling (City of Chicago, n.d. b; Divvy Bikes, n.d.). Public datasets used in the analysis.

Public source	Resolution	Primary role
Chicago COVID historical, yhhz-zm2v	ZIP by week	reported cases and effective initial state
Divvy monthly trip history (fallback)	anonymized timed trip	directional origin-destination proxy

Public source	Resolution	Primary role
Census 2020 TIGER ZCTA520	2020 polygon	point assignment and contiguity
ACS 2020 5-year B01003	ZCTA population	case/mobility normalization

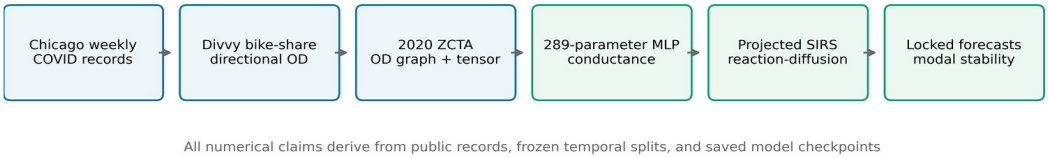


Figure 1. Public-data alignment learned conservative transport, one-week forecasting, and deterministic modal stability workflow.

2.2 Epidemiological preprocessing and cohort

From the total number of 1,748 ZIP-weeks identified for inclusion in the dataset, 8 (0.46%) cases with missing case numbers were eliminated from analysis due to outcome loss and performance measures. ZIP codes that were “Unknown” and 60666 did not have the corresponding ZCTA polygons in 2020. Among the 58 ZCTAs identified on the map, only 46 had the required number of 100 or more resolved bike trips.

Cohort measure	Verified value
Weeks; ZCTAs; ZIP-week records	38; 46; 1748
Suppressed/missing weekly cases	8 (0.46%)
Weekly case-rate range	0.0-1371.7 per 100,000
Spatially resolved Divvy trips	3,051,971 (98.78% of study-window trips)
Cohort trips; median per week	3,051,348; 83,636 (After the 46-ZCTA filter: 3,051,348)
Directed OD pairs; graph components	1673; 1
ACS population range	729-94,646
Train/validation/test weeks	26/4/8

2.3 Mobility preprocessing and spatial alignment

There were 3,114,796 rows across nine official Divvy data sources. Out of these, 3,089,684 started within the analysis weeks. The start and end coordinates were assigned based on point-in-polygon to 2020 ZCTAs in WGS84. Further calculations for geometry and centroid were performed using NAD83 Illinois East (EPSG:26971, meters). Only the trips where either one of the endpoints was not available or fell out of the mapped study area were removed, resulting in 3,051,971 (98.78%) trips within the 58-ZCTA crosswalk. OD counts for each week were saved for descriptive purposes

$$F_{ij}(t) = \#\{\text{trips } i \to j \text{ in week } t\},\ Q_{ij}(t) = F_{ij}(t) + F_{ji}(t) \tag{1}$$

$$q_{ij}(t) = \frac{Q_{ij}(t)}{\sqrt{N_i N_j}} \tag{2}$$

The ACS denominator N_i is fixed over the study and q_{ij} is a mobility intensity proxy; neither implies that one observed bike trip corresponds to one epidemiological contact.

2.4 Conductance graph and local anisotropy

The candidate edges were the union of the observed OD pairs during the training period and ZCTA adjacency, and no OD pair from the test period was included in the candidate edges. Edge attributes included log symmetric normalized OD pair strength, center distance, ranked trip and population volume of nodes, and axis coefficients cos(2θ_ij), sin(2θ_ij). The mean and standard deviation of continuous attributes were estimated from the training period alone. A symmetric 8-16-8-1 tanh MLP with 289 trainable parameters yielded positive edge conductance. A trip presence indicator will mobility edge. The small geographic backbone $\varepsilon_{\text{geo}} = 0.02$ applies before a per-week maximum-degree normalization; that normalization bounds integration stiffness while preserving symmetry and relative heterogeneity.

$$\widetilde{w}_{ij}(t) = \mathbf{1}_{\{Q_{ij}(t)>0\}} softplus\left(g_\phi(x_{ij}(t))\right) + \varepsilon_{geo} A_{ij}^{geo}, \qquad w_{ij}(t) = \frac{\widetilde{w}_{ij}(t)}{max\{1, \max_i \sum_j \widetilde{w}_{ij}(t)\}} \tag{3}$$

$$D_{ii}(t) = \sum_j w_{ij}(t), \qquad L(t) = D(t) - W(t), \quad W_{ij}(t) = w_{ij}(t) \tag{4}$$

We confirmed nonnegative finite weight values, symmetric Laplacians, row sums of zeros, and nonnegative Laplacian eigenvalues. The graph generated from this process is the discrete version of anisotropic diffusion, and not the approximation of a signed advection function. The second moment tensor is computed using the projection of the centroid unit vector.

$$A_i(t) = \frac{\sum_j w_{ij}(t) e_{ij} e_{ij}^{\top}}{\sum_j w_{ij}(t) + \varepsilon} \tag{5}$$

$$AI_i(t) = \frac{\lambda_{1,i}(t) - \lambda_{2,i}(t)}{\lambda_{1,i}(t) + \lambda_{2,i}(t) + \varepsilon}, \quad \lambda_{1,i} \geq \lambda_{2,i} \geq 0 \tag{6}$$

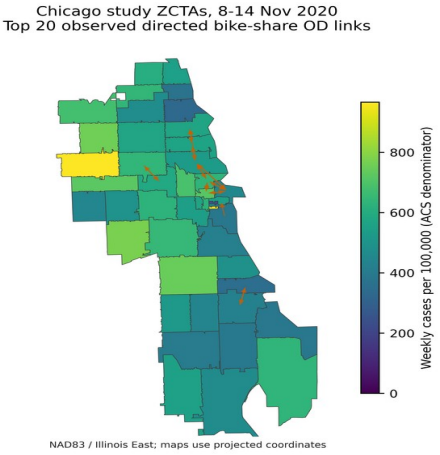


Figure 2. Study ZCTAs, observed cases, and the strongest directed bike-share OD links in a representative test week.

2.5 Projected epidemic reaction-diffusion dynamics

Compartment fractions $x_i = (s_i, i_i, r_i)$ are fit to the reported data but not understood as a precise biological quantity. The saturating incidence function employs a globally smooth six-term temporal basis (constant, trend, and two Fourier harmonics). An unconstrained temporal network is not permitted. Removal rate γ, effective return rate ω, saturation rate κ, and transport scaling coefficients are transformed to positive numbers. We assumed d_S=d_R=d_M but estimated d_I independently. Different diffusions for compartments would create a violation of the sum of fractions locally; an analytic projection onto tangent space ensures s_i+i_i+r_i=1 at all nodes while maintaining global mass balance.

$$h_i(t) = \frac{\beta(t) s_i(t) i_i(t)}{1 + \kappa i_i(t)}, \qquad \beta(t) = softplus(B(t)c) \tag{7}$$

$$\dot{x}_i = f(x_i, t) - \sum_j L_{ij}(t) P\, diag(d_M, d_I, d_M) x_j, \quad f(x_i, t) = (-h_i + \omega r_i,\ h_i - \gamma i_i,\ \gamma i_i - \omega r_i)^{\top}, \quad P = I_3 - \tfrac{1}{3}\mathbf{1}\mathbf{1}^{\top} \tag{8}$$

For each week-long rolling forecast, the previously reported values of i and r on a weekly basis and cumulatively initialize the respective values of effective i and r with regularization applied to the global infected fraction and explicit bounds for the admissible states. The susceptible population is then computed as 1-i-r. Next, the model integrates one week of time using four RK4 steps, while aggregating the incidence

$$\hat{y}_{i,t} = 10^5 \rho \int_t^{t+1} h_i(u)\, du, \qquad \rho = sigmoid(\rho_{raw}) \tag{9}$$

2.6 Training, comparators, and locked validation

The split was 26 weeks of training (5 April-3 October), 4 weeks of validation (4-31 October), and 8 weeks of locked final period (1 November-26 December). All the models predicted week t based on reported cases up to t-1 and mobility up to t-1; that is, a rolling one-week surveillance procedure, not a prediction of the future based on unknown future mobility. The masked robust loss function and the penalty on parameters/weights/states is defined in Equation 10. The Adam training procedure involved a maximum of 250 epochs and a patience of 30 epochs in the validation stage. The sweep of beta-roughness values (0.001, 0.01, 0.1) provided identical validation RMSE at seed 11; hence 0.01 was fixed prior to the test period.

$$\mathcal{L} = Huber\left(\frac{\hat{y} - y}{300}\right) + \lambda_\beta \overline{(\Delta^2 \beta)^2} + \lambda_w \parallel \phi \parallel_2^2 + \lambda_p R_{param} + \lambda_s R_{state} \tag{10}$$

Baselines considered were persistence, geographic diffusion, static and weekly mobility SIRS models, ablations, and a GCN-GRU with the same lagged inputs. Specifications for each model can be found in CSV and LaTeX in the reproducibility package.

2.7 Forecast metrics and uncertainty

The primary evaluation measures considered included the pooled RMSE and MAE of weekly cases per 100,000 in the observed, unsuppressed ZCTA-weeks. The secondary evaluation measures included the sMAPE, spatial correlation of weekly maps using Pearson/Spearman correlation, hotspot Jaccard/F1, peak timing error, and difference in absolute Moran's I on the contiguity network. The Pearson/Spearman was considered undefined for almost constant prediction maps (maximum standard deviation being one case per 100,000). A paired 1,000-resample moving two-week temporal block bootstrap resulted in 95% conditional confidence intervals for the eight-week holdout and model pair differences.

2.8 Stability, uncertainty and stochastic extension

Firstly, the deterministic SIRS model was fitted. A stochastic infection propagation model was subsequently fitted by adding a stochastic component to the infection transfer, using the Euler-Maruyama method with 20 substeps, 200 sample paths and noise scaling obtained from the training residual data only for the infected state, which was subtracted from the susceptible state as well to preserve the population balance locally.

$$di_i = b_i(x,t)\,dt + \sigma i_i\,dB_i(t),$$

$$ds_i = a_i(x,t)\,dt - \sigma i_i\,dB_i(t) \tag{11}$$

We used representative fitted training-period beta, solved for a feasible homogeneous equilibrium and computed the two variable reaction Jacobian, eliminating r=1-s-i to account for conservation. A comparison of the analytic Jacobian with centered finite differences was made. For primary transport-only stability comparison we eigen-decomposed the learned training-mean and uniform-geographic Laplacians and took the same reaction/diffusion parameters. One is not a posterior distribution, it is a local 200-sample multiplicative sensitivity ensemble (5% reaction and beta, 10% diffusion perturbations) which tests classification robustness. The 40x40 beta/diffusivity ratio continuation is called counterfactual. Stable homogeneous growth and positive real-dominant spatial growth is Turing-consistent, while positive complex-conjugate crossing with a nonzero imaginary part is Hopf-consistent.

$$J = \begin{pmatrix} -h_s - \omega & -h_i - \omega \\ h_s & h_i - \gamma \end{pmatrix}_{x^*}, \qquad h_s = \frac{\beta i^*}{1+\kappa i^*}, \quad h_i = \frac{\beta s^*}{(1+\kappa i^*)^2} \tag{12}$$

$$M_k = J - \mu_k D_{red}, \qquad D_{red} = \begin{pmatrix} d_M & -(d_I - d_M)/3 \\ 0 & (2d_I + d_M)/3 \end{pmatrix} \tag{13}$$

$$g(\mu_k) = \max_{\zeta \in eig(M_k)} Re(\zeta) \tag{14}$$

For the same reaction terms, the continuous reaction dispersion relation g(μ) is identical for both of the graphs, but the discrete spatial eigenvalue spectrum and eigenvectors, by which the relation is sampled, is modified by anisotropic mobility. We did not assume that instability was just because a case map was observed and it was shaped that way.

2.9 Reproducibility and public-data ethics

This study involved only publicly available data that did not contain identifiable information and did not involve linking any individuals or inferences. Missing values in epidemiological data were considered missing. Reproducibility was possible by using fixed temporal partitions, deterministic random number generators, proper preprocessing, and model evaluation. Therefore, individual-level ethical approvals were not necessary.

3 Results

3.1 Dataset characteristics

Frozen data comprised 46 ZCTAs within 38 weeks (1,748 ZIP-weeks) with 8 missing case counts and 0–1,371.7 cases per 100,000 persons for weekly rates. The OD data consisted of 3,051,348 trips between 1,673 directed pairs with 83,636 average weekly trips. The graphs created from the candidate list and geography were both connected. While most of the observations came before the rollout of vaccines, late December coincided with vaccine rollout, and bike share was limited to the center.

3.2 Learned mobility anisotropy

A mean anisotropy index of 0.407 (interquartile range 0.263-0.567; median 0.398) was obtained over ZCTA-weeks from the learned 2nd moment tensor. The 15 nodes were those with a time-averaged index greater than 0.5. The inferred conductance property, principal axes, often paralleled the north-south oriented lakefront bike network (Figure 3), which

is not directly measured or a true pathogen flow but a representation of the direction of conductance.

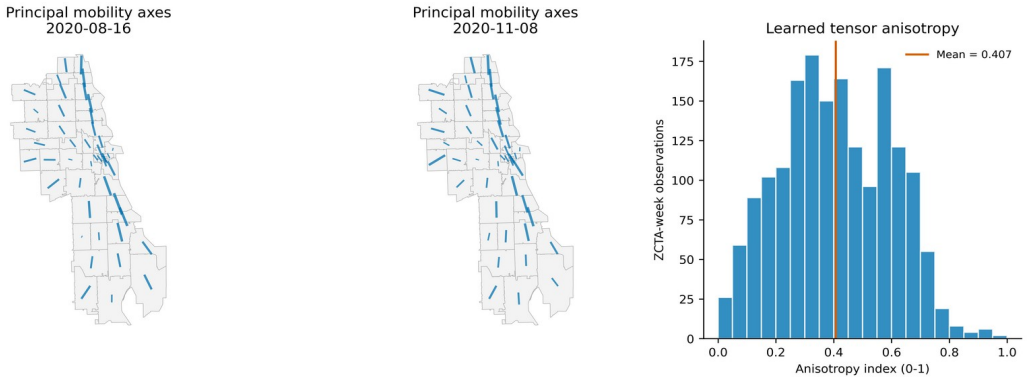


Figure 3. Inferred local conductance axes and the distribution of ZCTA-week anisotropy indices.

3.3 Locked-test forecasting performance

The learned model reached test RMSE 145.63 (95% two-week block interval 90.60-181.83) and MAE 109.51 per 100,000

Model	RMSE (95% block CI)	MAE	Spatial Pearson	Hotspot Jaccard
Persistence	135.1 (105.6-153.8)	107.4	0.734	0.589
Isotropic SIRS	145.9 (90.8-182.2)	109.8	0.719	0.589
Static mobility SIRS	146.0 (90.8-182.3)	109.9	0.720	0.589
Raw weekly SIRS	146.0 (90.8-182.2)	109.9	0.720	0.589
Learned mobility SIRS	145.6 (90.6-181.8)	109.5	0.720	0.589
Constant beta ablation	131.3 (83.2-164.4)	96.4	0.719	0.589
GCN-GRU	265.4 (162.1-349.8)	207.7	0.431	0.345

across 368 observed test records. The five seed RMSE standard deviation was 0.014. Isotropic SIRS reached RMSE 145.93; the paired learned-minus-isotropic difference was -0.303 (-0.436 to -0.184), or a 0.208% gain. The static and raw weekly mobility SIRS also did not differ much, with 145.97 and 145.95, respectively. Even with those minor differences depending on the condition, persistence achieved RMSE 135.15 while the constant-beta ablation reached 131.32, which is better on the held-out weeks, although the former was validated as the better candidate. GCN-GRU reached RMSE 265.44. The complete comparisons are presented in Table 3 and Figures 4-5.

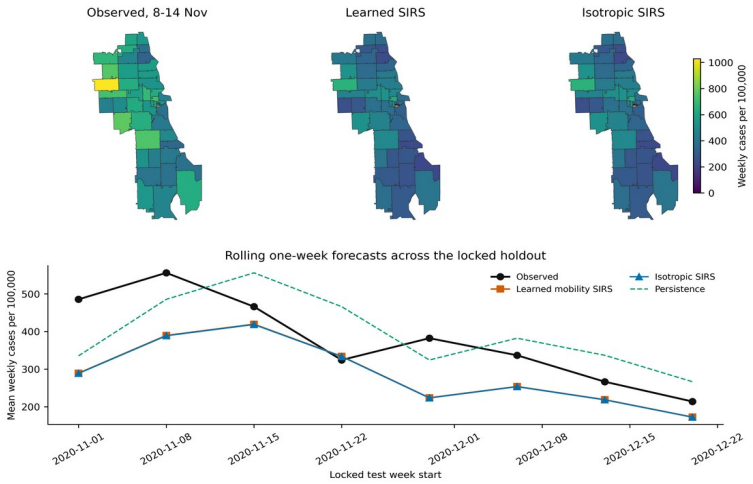


Figure 4. Representative observed and mechanistic prediction maps and mean weekly rates through the locked test.

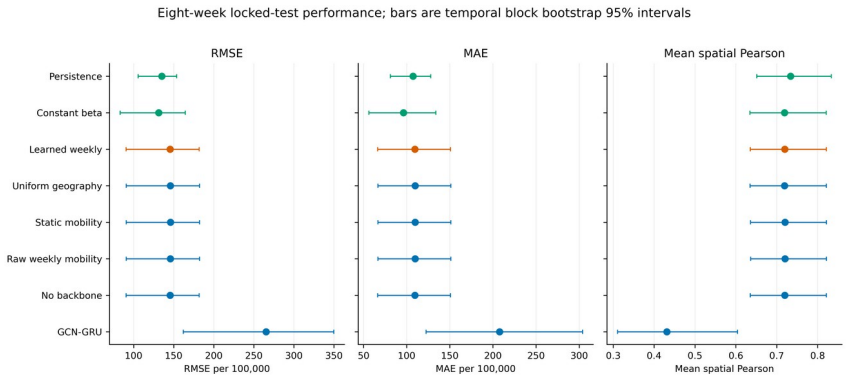


Figure 5. Locked-test predictive performance with conditional moving-block bootstrap intervals.

3.4 Ablations and spatial pattern fidelity

After removing the backbone, the value of RMSE does not change significantly, -0.0005 per 100,000, which is like the connected observed graph. The increase in RMSE from replacing learned conductance with static or raw weekly mobility was 0.345 and 0.324; the replacement of smooth beta with constant beta improved RMSE by a much larger amount of 14.308. The complete source table for ablations and Supplementary Figure S1 are provided with the project. Learned model and persistence weekly mean spatial Pearson correlation were 0.720 and 0.734, respectively and attained hotspot Jaccard 0.589. It had a mean absolute Moran's I discrepancy of 0.075, which was very close to the isotropic value of 0.075. The spatial rankings were similar between prediction and observation, and there was no visible improvement in the level of agreement between the two after including transport learning.

3.5 Deterministic stability and counterfactual continuation

The conservation-reduced feasible equilibrium turned out to be effective, with beta 0.752 week^(-1) at the representative, (s^*,i^*,r^*)=(0.629,0.045,0.326). The analytic Jacobian was within 3.26e-12 of the centered finite differences. The maximum real growth, both homogeneous and learned, were -0.05238 week^(-1) and -0.05490 at graph eigenvalue 0.05360, respectively. The uniform geography shift was -0.05294, which was followed by a learned-minus-uniform shift of -0.00196. Although the spatial mode index was the same for both graphs, 1, the spatial eigenvector was more localized under learned mobility (Figure 6). No Turing or Hopf consistency was observed, and every observed mode was stable. The 200 local perturbation samples were stable for both graphs. A 1,600-point counterfactual beta/diffusivity scan revealed no Turing/Hopf classifications; the nearby 1-point instability at β=γ occurred at a multiplier of 0.620 on representative β, but not at a diffusion-driven instability.

Conservation-reduced modal stability under identical fitted reaction parameters; growth rates are per week.

Transport graph	Homogeneous growth	Largest spatial growth	Dominant mode	Classification
Learned mobility	-0.05238	-0.05490	1	stable all modes
Uniform geography	-0.05238	-0.05294	1	stable all modes

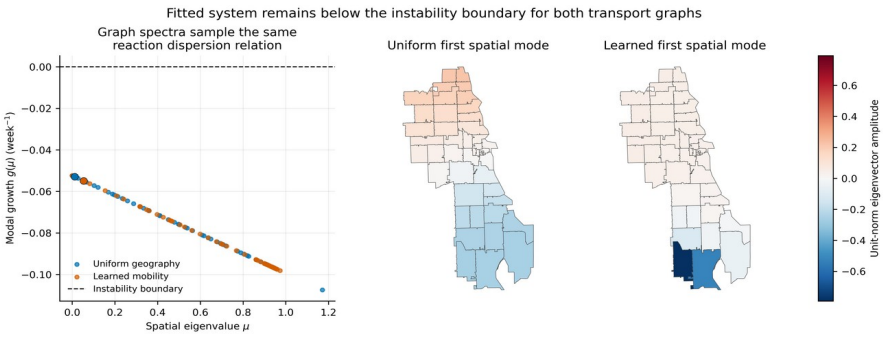


Figure 6. Two transport spectra sampling the same fitted dispersion relation, with least-damped mode maps.

3.6 Stochastic predictive uncertainty

Based on 200 stochastic runs, test RMSE was 147.83 per 100,000, with 75.8% interval coverage. While numerically stable, the absence of full coverage indicated that stochastic modeling could not bridge surveillance and model inconsistencies.

4 Discussion

The bike-share trip data gave us a graph that's uneven across space and sensitive to direction. But the extra mobility signal it captured, compared to just assuming uniform or static transport, wasn't much, around three-tenths of a case per 100,000 in RMSE. A paired bootstrap interval did exclude zero for that difference, sure. But that's the local result. It doesn't tell you the method will hold up in a different city, or even a different season in the same one.

Here's the part that undercuts the model: a much simpler constant-beta version did far better on the test block, even though it scored worse during validation. If edge conductance were really the thing driving accuracy, you wouldn't expect that. It looks more like shifting transmission rates and reporting quirks are doing most of the work not the graph structure. Both a plain persistence baseline and the constant-beta version beat the proposed model outright. And comparing favorably to the GCN-GRU doesn't mean much either, since that network never had enough training data to be a fair benchmark. Beating an undertrained model isn't proof that physics-informed learning is pulling its weight.

On the epidemiology side, anisotropic transport did shift the graph's spectrum and its slowest-decaying spatial mode but every growth margin stayed negative the whole time. Complex eigenvalues with negative real parts aren't a Hopf bifurcation, no matter how you frame it. A COVID map that looks visually patterned isn't a Turing pattern, and running a counterfactual scan tells you about the model's sensitivity to its own inputs, not about how the epidemic behaved. In the SIRS projection, the fitted diffusivity for infected versus non-infected compartments came out nearly the same. Which alone limits how much diffusion-driven instability is even mathematically possible. Some of the limitations of the study include: Divvy has limitations in terms of exclusion of most forms of transportation and inconsistent coverage, while ride-hailing and case data are affected by suppression, rounding off of timestamps and inconsistent reporting. The mismatched geography limits the comparison of mobility and cases to ecological relationship and not individual transmission. Network Symmetrizing eliminates any form of directionality while biases in reporting and vaccination roll out complicates SIRS estimation.

5 Conclusions

The anisotropic transport model was better than isotropic SIRS by just 0.208% in the testing RMSE, while being inferior to both persistence and constant-β models. No instability was observed for all spatial models – neither Turing nor Hopf. Bike-sharing mobility and ZCTA data thus reflect anisotropic transport structure, which has not yet been proven to enhance forecasting capabilities. The framework is presented as a diagnostic tool, not as a forecast method..

Data Availability

Public source records are available from the City of Chicago historical COVID and TNP data pages, Divvy's public monthly system trip archives, and the U.S. Census Bureau 2020 TIGER and ACS B01003 pages (City of Chicago n.d.a, n.d.b; Divvy Bikes n.d.; U.S. Census Bureau n.d.b, n.d.a).

Code Availability

The computational analyses in this study were performed using custom scripts developed by the authors. The code is not publicly available. The methodological details required to understand and reproduce the computational workflow are provided in the manuscript.

Ethics and Public Data Statement

This study uses public, anonymized aggregate or trip-history data without personal identity or linked infection/travel records. Suppressed surveillance and source-tract values were respected, and no individual-level transmission claim is made.

References

1. A. M. Turing, "The chemical basis of morphogenesis," *Philosophical Transactions of the Royal Society of London B*, vol. 237, no. 641, pp. 37–72, 1952, doi: 10.1098/rstb.1952.0012.
2. V. Davydovych, V. A. Dutka, and R. Cherniha, "Reaction–Diffusion Equations in Mathematical Models Arising in Epidemiology," *Symmetry*, vol. 15, no. 11, art. 2025, 2023, doi: 10.3390/sym15112025.
3. G.-Q. Sun, Z. Jin, Q.-X. Liu, and L. Li, "Pattern formation in a spatial S–I model with non-linear incidence rates," *Journal of Statistical Mechanics: Theory and Experiment*, art. P11011, 2007, doi: 10.1088/1742-5468/2007/11/P11011.
4. Y. Cai and W. Wang, "Spatiotemporal dynamics of a reaction–diffusion epidemic model with nonlinear incidence rate," *Journal of Statistical Mechanics: Theory and Experiment*, art. P02025, 2011, doi: 10.1088/1742-5468/2011/02/P02025.
5. W. Wang, Y. Lin, H. Wang, H. Liu, and Y. Tan, "Pattern selection in an epidemic model with self and cross diffusion," *Journal of Biological Systems*, vol. 19, no. 1, pp. 19–31, 2011, doi: 10.1142/S0218339011003555.

6. W. Wang, Y. Cai, M. Wu, K. Wang, and Z. Li, “Complex dynamics of a reaction–diffusion epidemic model,” *Nonlinear Analysis: Real World Applications*, vol. 13, no. 5, pp. 2240–2258, 2012, doi: 10.1016/j.nonrwa.2012.01.018.
7. M. Duan, L. Chang, and Z. Jin, “Turing patterns of an SI epidemic model with cross-diffusion on complex networks,” *Physica A: Statistical Mechanics and its Applications*, vol. 533, art. 122023, 2019, doi: 10.1016/j.physa.2019.122023.
8. Q. Zheng, J. Shen, V. Pandey, L. Guan, and Y. Guo, “Turing instability in a network-organized epidemic model with delay,” *Chaos, Solitons & Fractals*, vol. 168, art. 113205, 2023, doi: 10.1016/j.chaos.2023.113205.
9. J. Guo, X. Li, R. He, X. Luo, Z.-G. Guo, and G.-Q. Sun, “Pattern dynamics of networked epidemic model with higher-order infections,” *Chaos*, 2024, doi: 10.1063/5.0224187.
10. A. K. Singh, S. Ramakrishnan, and M. Kumar, “Instabilities and self-organization in spatiotemporal epidemic dynamics driven by nonlinearity and noise,” *Physical Biology*, vol. 21, no. 4, art. 046001, 2024, doi: 10.1088/1478-3975/ad5d6a.
11. A. K. Singh, N. Boltz, M. Kumar, and S. Ramakrishnan, “Turing-Type Instabilities and Pattern Formation Induced by Saturation Effects and Randomness in Nonlinear, Diffusive Epidemic Spread,” in *2024 American Control Conference (ACC)*, pp. 4711–4716, 2024, doi: 10.23919/ACC60939.2024.10644593.
12. A. K. Singh, C. Buschmeyer, S. Ramakrishnan, and M. Kumar, “Instabilities and Pattern Formation in Epidemic Spread Induced by Nonlinear Saturation Effects and Ornstein–Uhlenbeck Noise,” *ASME Letters in Dynamic Systems and Control*, vol. 5, no. 1, art. 011002, 2025, doi: 10.1115/1.4066628.
13. R. Almshekhs, A. K. Singh, and S. Ramakrishnan, “Anisotropic Diffusion and the Emergence of Instabilities and Self-Organized Pattern Formation in Spatiotemporal Epidemic Dynamics,” *ASME Letters in Dynamic Systems and Control*, vol. 6, no. 2, art. 021010, 2026, doi: 10.1115/1.4070496.
14. A. K. Singh, R. Almshekhs, M. Kumar, and S. Ramakrishnan, “The Emergence of Turing Instability and Pattern Formation in a Nonlinear Stochastic Spatiotemporal Epidemic Model with Reinfections,” *Acta Biotheoretica*, vol. 74, art. 10, 2026, doi: 10.1007/s10441-026-09518-7.
15. D. M. Busiello, G. Planchon, M. Asllani, T. Carletti, and D. Fanelli, “Pattern formation for reactive species undergoing anisotropic diffusion,” *European Physical Journal B*, 2015, doi: 10.1140/epjb/e2015-60269-0.
16. Y. Önder and R. A. Van Gorder, “Modification of Turing patterns through the use of time-varying anisotropic diffusion,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2023, doi: 10.1098/rspa.2023.0487.
17. C. Stoicescu, “On a nonlocal and nonlinear second-order anisotropic reaction-diffusion system with in-homogeneous Cauchy–Neumann boundary conditions: Applications on epidemic infection spread,” *Discrete and Continuous Dynamical Systems – S*, 2023, doi: 10.3934/dcdss.2022172.
18. F. Majid, A. M. Deshpande, S. Ramakrishnan, S. Ehrlich, and M. Kumar, “Analysis of epidemic spread dynamics using a PDE model and COVID-19 data from Hamilton County OH USA,” *IFAC-PapersOnLine*, vol. 54, no. 20, pp. 322–327, 2021, doi: 10.1016/j.ifacol.2021.11.194.
19. F. Majid, M. Gray, A. M. Deshpande, S. Ramakrishnan, M. Kumar, and S. Ehrlich, “Non-Pharmaceutical Interventions as Controls to mitigate the spread of epidemics: An analysis using a spatiotemporal PDE model and COVID-19 data,” *ISA Transactions*, vol. 124, pp. 215–224, 2022, doi: 10.1016/j.isatra.2021.02.038.
20. D. A. David, C. Liu, L. Street, S. Ehrlich, M. Kumar, and S. Ramakrishnan, “A novel approach to forecasting reproduction numbers of spatiotemporal stochastic epidemic spread using a PDE-based model and real-time infection data,” *Scientific Reports*, vol. 15, art. 9760, 2025, doi: 10.1038/s41598-025-91811-5.
21. M. Zhang, S. Wang, T. Hu, X. Fu, X. Wang, Y. Hu, B. Halloran, Z. Li, Y. Cui, H. Liu, Z. Liu, and S. Bao, “Human mobility and COVID-19 transmission: a systematic review and future directions,” *Annals of GIS*, vol. 28, no. 4, pp. 501–514, 2022, doi: 10.1080/19475683.2022.2041725.
22. Anupriya, P. Bansal, and D. J. Graham, “Modelling the propagation of infectious disease via transportation networks,” *Scientific Reports*, vol. 12, art. 20572, 2022, doi: 10.1038/s41598-022-24866-3.
23. A. Rodríguez, J. Cui, N. Ramakrishnan, B. Adhikari, and B. A. Prakash, “EINNs: Epidemiologically-Informed Neural Networks,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 12, pp. 14453–14460, 2023, doi: 10.1609/aaai.v37i12.26690.
24. S. Han, L. Stelz, H. Stoecker, L. Wang, and K. Zhou, “Approaching epidemiological dynamics of COVID-19 with physics-informed neural networks,” *Journal of the Franklin Institute*, vol. 361, no. 6, art. 106671, 2024, doi: 10.1016/j.jfranklin.2024.106671.

Different Generations of Metaheuristic Algorithms in Structural Optimization: Tubular Column and I-Beam Applications

Yaren Aydın[1], Gebrail Bekdaş[1] and Sinan Melih Nigdeli[1]

[1]Department of Civil Engineering, Istanbul University-Cerrahpaşa,
34320 Avcılar/Istanbul/Turkey
yaren.aydin@iuc.edu.tr, bekdas@iuc.edu.tr, melihnig@iuc.edu.tr

Abstract. In this study, the problems of minimizing the cost of a tubular column and minimizing the vertical displacement of an I-section beam are addressed using various meta-heuristic optimization algorithms. Within this scope, solutions obtained by applying the Bat Algorithm (BA), Archimedes Optimization Algorithm (AOA), and Competition-Amensalism Optimization (CAO) were compared. During the optimization process, design variables related to column and beam cross-sections were optimized subject to constraints determined by cross-sectional area and stress conditions. The performance of the algorithms was evaluated based on minimum, mean, and standard deviation values obtained from 20 independent runs, as well as the number of iterations required to reach the optimum solution. The results revealed that all three algorithms produced feasible solutions but differed in terms of solution stability and convergence behavior. Notably, while CAO achieved the same minimum objective function value as BA, it stood out due to its lower standard deviation and fewer average iterations. The findings indicate that CAO can be an effective and robust alternative for the engineering optimization problem under consideration.

Keywords: Metaheuristic methods, Benchmarks, Algorithms, Optimization, Column, Beam.

1 Introduction

Optimization is the process of obtaining the best solution from among the feasible solutions available under specific constraints. In optimization problems, the variables that influence performance are termed "decision variables," while the structure that mathematically expresses the impact of these variables on the objective is known as the "objective function". Metaheuristic methods are problem-independent techniques that aim to generate approximate solutions to complex problems within an acceptable timeframe by effectively exploring the search space [1]. However, these algorithms do not guarantee finding the global optimum.

There are numerous studies in the literature in which Metaheuristic optimization algorithms are applied in civil engineering. Bekdaş and Nigdeli (2014) [2] used the Harmony Search (HS) algorithm for the optimum design of slender reinforced concrete columns. They concluded that the optimum total material cost increases non-linearly as column length increases and that slenderness has a significant impact on column cost. Bekdaş and Nigdeli (2016) [3] investigated the cost-optimal design of reinforced concrete columns using the Teaching-Learning-Based Optimization (TLBO) algorithm. Consequently, the proposed method offered an effective approach for the optimal design of reinforced concrete columns. Bekdaş et al. (2022) [4] used the Harmony Search algorithm for the optimum sizing of reinforced concrete circular columns, considering the total cost of steel and concrete as the objective function. Sojobi and Liew (2023) [5] optimized high performance carbon fibre reinforced polymer (CFRP) composites for strengthening recycled concrete columns using the Taguchi-Response surface methodology. In the optimized columns strengthened with only two layers of CFRP wrapping, increases of 36.3% and 36.19% were achieved in fracture toughness and fracture energy, respectively, along with improvements of 1.4 times better in the displacement ductility index and 2.78 times in the strain ductility index. Aydın et al. (2024) [6] carried out the optimal design of rectangular reinforced concrete columns subjected to uniaxial bending in terms of cost and CO_2 emissions. Tunca and Çarbaş (2024) [7] used the Coati Optimization Algorithm, Fox Optimizer, and Pelican Optimization Algorithm (POA) to minimize the design cost of reinforced concrete columns. The results presented a comparative analysis of the three algorithms' performances and demonstrated that the POA yielded the most successful solution. Luo et al. (2025) investigated the carbon emissions and costs of prefabricated concrete-filled steel tubular columns (PCSPC) during the construction phase using the LCA method. The results indicated that 98.845% of the total carbon emissions originated from the production phase, with steel plates having the greatest impact. As a result of optimization, carbon emissions were reduced by 10.32% and costs by 21.55%, while maintaining load-bearing capacity. Tunca and Çarbaş (2026) [8] carried out the optimization of the design cost for reinforced concrete columns using the Gaining-sharing Knowledge-based Algorithm (GSKA), Human Conception Optimizer (HCO), and War Strategy Optimization (WSO) algorithms. The results demonstrated that GSKA, in particular, is an effective optimization approach for reinforced concrete column design in terms of cost and design efficiency. Xiong et al. (2026) proposed a smart design method that integrates topology optimization into column layout design for reinforced concrete frames. Validated through numerical and practical examples, the method reduced material usage by 16-18% compared to conventional designs while maintaining structural rigidity.

An examination of the cited studies reveals that these methods yield successful results in the optimal design of columns. Given the diversity of problem types and solution spaces, the absence of a single Metaheuristic algorithm that outperforms others across all problems ensures that research in this field remains relevant. New Metaheuristic algorithms are being developed, and efforts to enhance the performance of existing algorithms continue, all with the aim of achieving more effective and reliable solutions.

In this study, the Bat Algorithm (developed by Xin-She Yang [9] in 2010), the Archimedes Optimization Algorithm (AOA) (developed by Hashim et al. [10] in 2021), and the Competition-Amensalism Optimization (CAO) algorithm -a very new algorithm (developed by Yilma and Woldegiorgis [11] in 2026) were utilized and compared.

The distinct parameters inherent to each algorithm can directly influence their local and global search capabilities. In this study, three different Metaheuristic algorithms were applied to achieve the optimal solution, and the performance of the algorithms was compared based on the obtained results. Two different structural optimization problems were addressed for all three algorithms, and MATLAB [12] was used during the optimization process. The first problem involves the cost optimization of a tubular column subjected to a compressive load, while the second concerns the displacement optimization of an I-section steel beam.

2 An Overview of BA, AOA and CAO Algorithms

2.1 Bat Algorithm (BA)

Bat Algorithm (BA) is a population-based Metaheuristic algorithm developed in 2010 based on the way bats use an acoustic radar system known as echolocation to navigate in the dark and detect their prey [9] in 2010. In the BA, the position of each bat represents a solution to an optimization problem. In the BA, each individual represents a bat, and the bats' position, velocity, frequency, loudness, and pulse emission rates are taken into account.

2.2 Archimedes Optimization Algorithm (AOA)

The Archimedes Optimization Algorithm (AOA) is a population-based optimization method developed in 2021, inspired by Archimedes' principle. In the algorithm, objects are moved within the solution space using information regarding volume, density, acceleration, and position. In each iteration, volume and density values are updated, acceleration is calculated based on the collision state, and acceleration is normalized. The best fitness value is determined by updating the positions of the objects according to their moving or stationary states [10].

2.3 Competition-Amensalism Optimization (CAO)

Competition-Amensalism Optimization (CAO) is a Metaheuristic optimization algorithm developed in 2026, inspired by competitive interactions in ecological systems. The algorithm employs a competition phase for global exploration and an amensalism phase for local exploitation. This two-phase structure aims to effectively scan the search space and achieve a balance between exploration and exploitation [11].

The code for the algorithms employed in this study is based on the algorithmic steps and implementation code provided in the original publications where these algorithms were first proposed [[9],[10],[11]]. Code and implementation specifications for the

algorithms used can be found in the relevant original sources. Table 1 gives the parameters of Metaheuristic algorithms used in structural optimization problems.

Table 1. Parameters of Metaheuristic algorithms used in structural optimization problems.

Algorithm	Parameter	Value
All	Population size	20
	Maximum iterations	10000
	Runs	20
BA	A_0, A_{min}, α, f_{min}, f_{max}	1.0, 0.0, 0.9, 0, 100
AOA	Density (den), Volume (vol), Acceleration (acc)	rand [0,1]
CAO	Interaction effect	[1,2]

3 Structural Optimization Problems

To evaluate the performance of the Metaheuristic algorithms, two problems within the scope of structural engineering were addressed. The comparison of the algorithms was based on identical design variables, objective functions, and design constraints. Thus, the success of each algorithm in reaching the optimum solution for the structural design problem and the quality of the solutions obtained were comparatively assessed.

The first problem is to minimize the cost of a tubular column subjected to a compressive load (Fig. 1). The design variables are the column's cross-sectional diameter (d) and thickness (t). P represents the compressive load and l the column length, while σy, E, and ρ denote the material's yield stress, modulus of elasticity, and density, respectively.

For this problem, the objective function representing total cost minimization is given in Equation 1, and the variation ranges for design parameters are provided in Equations 2 and 3; the problem constraints relate to axial load capacity (g_1) (Equation 4) and the column buckling limit (g_2) (Equation 5).

$$f(d,t) = 9.8 \times d \times t + 2 \times d \quad (1)$$

$$2 \le d \le 14 \quad (2)$$

$$0.2 \le t \le 0.9 \quad (3)$$

$$g_1 = \frac{P}{\pi \times d \times t \times \sigma_y} - 1 \le 0 \quad (4)$$

$$g_2 = \frac{8 \times P \times l^2}{\pi^3 \times E \times d \times t \times (d^2 + t^2)} - 1 \le 0 \quad (5)$$

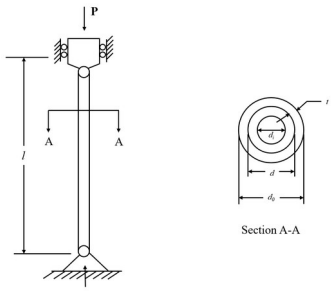


Fig. 1. Tubular column under compression and section detail [13].

The other optimization problem is to minimize the vertical displacement of an I-section steel beam under applied loads (Fig. 2). The design variables are beam height h, flange width b, flange thickness t_f and web thickness t_w, all measured in centimeters. A vertical load P is applied at the beam's midpoint, and a horizontal load Q is applied at the center of the web. Vertical displacement is calculated using Equation 6:

$$f(x) = \frac{P \times L^3}{48 \times E \times I} \tag{6}$$

where L represents the beam span, E the modulus of elasticity, and I the moment of inertia. The objective function has been formulated using this relationship and the moment of inertia of the I-section. The objective function determined for the purpose of minimizing vertical displacement was obtained by incorporating the constants P, L, and E and the moment of inertia relationship into the displacement equation (Equation 7).

$$\min f\left(h,b,t_w,t_f\right) = \frac{5000}{\left(\frac{\left(t_w \times (h-2\times t_f)\right)^3}{12} + \frac{b \times t_f^3}{6} + 2 \times b \times t_f \times \left(\frac{h-t_f}{2}\right)^2\right)} \tag{7}$$

In the optimization problem, the constraint g1 stipulates that the beam cross-sectional area must not be less than 300 cm², while the constraint g_2 specifies that the bending stress must not exceed the permissible value of 6 kN/cm².

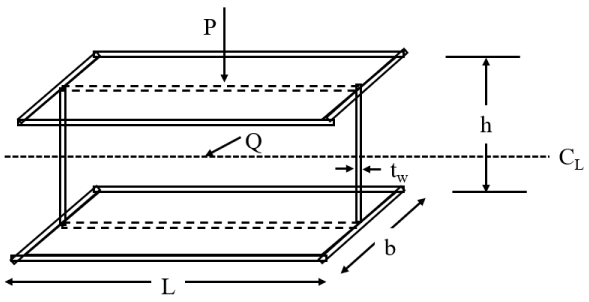


Fig. 2. The I beam problem [14] .

4 Results

4.1 Comparison Results for Tubular Column Cost Minimization Problem

During the optimization process, each algorithm was tested through 20 independent runs, and the obtained results were evaluated. Table 2 presents the minimum, mean, and standard deviation values for the algorithms, as well as the average number of iterations required to reach the optimal solution. Additionally, the convergence behaviors of the algorithms were compared using convergence curves (Fig. 3).

Across 20 independent runs, CAO achieved the lowest average and minimum fitness values (26.4995) and yielded the most stable results, with a standard deviation of 0.00003. BA ranked second with an average fitness value of 26.5207, while AOA produced the highest average value at 26.6277. AOA was the algorithm with the shortest average execution time, at 1.022 seconds (Table 2).

Table 2. Performance comparison of the algorithms for tubular column cost minimization problem.

Algorithm	Minimum	Mean	Std. deviation	Average time (s)	Iteration no.
BA	26.5066	26.5207	0.0103	1.123	8929
AOA	26.5282	26.6277	0.1073	1.022	37
CAO	26.4995	26.4995	0.00003	1.480	329

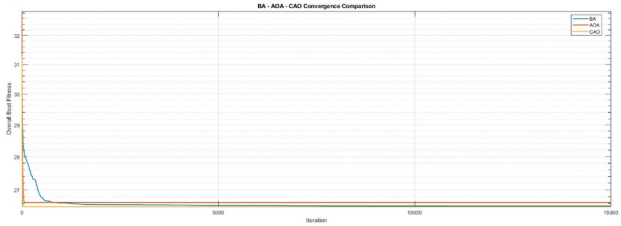


Fig. 3. Convergence curves of the algorithms for tubular column cost minimization problem.

4.2 Comparison Results for of an I-section Steel Beam Displacement Minimization Problem

To evaluate the performance of optimization algorithms for the—the I-section steel beam displacement minimization problem, the BAT, AOA, and CAO algorithms were tested across 20 independent runs. The same problem definition and stopping criteria were used in each run. The obtained results were compared based on minimum, mean, and standard deviation values (Table 3). Convergence behaviors of the algorithms were compared using convergence curves (Fig. 4).

An examination of the results in Table 3 reveals that the CAO algorithm outperforms the other algorithms in terms of the average objective function value and standard deviation. The average value of 0.013433 obtained for CAO indicates that the algorithm

exhibits stable performance across 20 independent runs. The standard deviation values further demonstrate that CAO has the lowest variability, at 0.000649.

In terms of the minimum objective function value, BA and CAO are quite close; the minimum value obtained was 0.0130741251 for BA and 0.0130741189 for CAO. The fact that CAO achieves lower average and standard deviation values indicates superior performance not only regarding the best solution found but also in terms of repeatability and stability. When iteration numbers are evaluated, the fact that CAO reaches its best solution in an average of 2460 iterations demonstrates faster convergence compared to BA and AOA. These values were found to be 12,729 and 5,151 iterations for BA and AOA, respectively.

Table 3. Performance comparison of the algorithms for I-section steel beam displacement minimization problem.

Algorithm	Minimum	Mean	Std. deviation	Iteration no.
BA	0.013074	0.018923	0.008636	12729
AOA	0.013928	0.020029	0.004373	5151
CAO	0.013074	0.013433	0.000649	2460

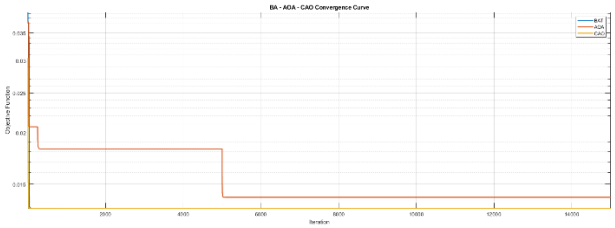


Fig. 4. Convergence curves of the algorithms for I-section steel beam displacement minimization problem.

5 Conclusions

In this study, the problems of cost minimization for tubular columns and vertical displacement minimization for I-section beams were solved using the Bat Algorithm (BA), Archimedes Optimization Algorithm (AOA), and Competition-Amensalism Optimization (CAO), and the algorithms were compared based on 20 independent runs.. The results revealed that while all three algorithms could generate feasible solutions for these structural problems, they differed in terms of solution quality and convergence behavior. Notably, CAO stood out regarding solution stability and convergence speed; it achieved the same minimum objective function value as BA but did so with a lower standard deviation and fewer average iterations. These results demonstrate that CAO - the most recently developed algorithm among those compared- can serve as an effective and stable metaheuristic optimization method for structural problems.

In future works, given that minor differences between algorithms were observed even in benchmark problems involving a small number of variables, plans are in place

to adapt the algorithms to more complex problems, subject them to comprehensive testing, and evaluate their performance. Such detailed testing is considered important for determining the algorithm's effectiveness across different problem structures.

References

[1] E. G. Talbi, *Metaheuristics: From Design to Implementation*. John Wiley & Sons, 2009.

[2] G. Bekdas and S. M. Nigdeli, "The Optimization of Slender Reinforced Concrete Columns," *PAMM*, vol. 14, no. 1, pp. 183–184, Dec. 2014, doi: 10.1002/pamm.201410079.

[3] "Optimum design of reinforced concrete columns employing teaching-learning based optimization," *Challenge Journal of Structural Mechanics*, 2016, doi: 10.20528/cjsmec.2016.11.030.

[4] G. Bekdaş, C. Cakiroglu, S. Kim, and Z. W. Geem, "Optimization and Predictive Modeling of Reinforced Concrete Circular Columns," *Materials*, vol. 15, no. 19, p. 6624, Sep. 2022, doi: 10.3390/ma15196624.

[5] A. O. Sojobi and K. M. Liew, "Multi-objective optimization of high performance concrete columns under compressive loading with potential applications for sustainable earthquake-resilient structures and infrastructures," *Compos. Struct.*, vol. 315, p. 117007, Jul. 2023, doi: 10.1016/j.compstruct.2023.117007.

[6] Y. Aydın, S. M. Nigdeli, G. Bekdaş, S. Kim, and Z. W. Geem, "A Review of Metaheuristic-Based Optimum Design of Reinforced Concrete Structures," 2024, pp. 116–151. doi: 10.4018/979-8-3693-2161-4.ch006.

[7] O. Tunca and S. Carbas, "Design cost minimization of a reinforced concrete column section using overnew swarm-based optimization algorithms," *Neural Comput. Appl.*, vol. 36, no. 27, pp. 16941–16958, Sep. 2024, doi: 10.1007/s00521-024-09998-z.

[8] O. Tunca and S. Carbas, "Human-Inspired Optimization Algorithms for Cost Minimization in Reinforced Concrete Column Design," 2026, pp. 63–92. doi: 10.1007/978-3-032-07738-7_4.

[9] X.-S. Yang, "A New Metaheuristic Bat-Inspired Algorithm," 2010, pp. 65–74. doi: 10.1007/978-3-642-12538-6_6.

[10] F. A. Hashim, K. Hussain, E. H. Houssein, M. S. Mabrouk, and W. Al-Atabany, "Archimedes optimization algorithm: a new metaheuristic algorithm for solving optimization problems," *Applied Intelligence*, vol. 51, no. 3, pp. 1531–1551, Mar. 2021, doi: 10.1007/s10489-020-01893-z.

[11] A. A. Yilma and B. H. Woldegiorgis, "Competition-amensalism optimization: A novel metaheuristic optimization algorithm," *Results in Engineering*, vol. 30, p. 111048, Jun. 2026, doi: 10.1016/j.rineng.2026.111048.

[12] M. Natick, "MATLAB," 2018.

[13] "Introduction to Optimization," in *Engineering Optimization*, Wiley, 2009, pp. 1–62. doi: 10.1002/9780470549124.ch1.

[14] X.-S. Yang, G. Bekdaş, and S. M. Nigdeli, Eds., *Metaheuristics and Optimization in Civil Engineering*, vol. 7. Cham: Springer International Publishing, 2016. doi: 10.1007/978-3-319-26245-1.

Machine Learning-Based Prediction of Concrete Compressive Strength Using a Multi-source Modified Concrete Dataset

Yaren Aydın[1], Gebrail Bekdaş[1], Sinan Melih Nigdeli[1] and Ümit Işıkdağ[2]

[1]Department of Civil Engineering, Istanbul University-Cerrahpaşa,
34320 Avcılar/Istanbul/Turkey
yaren.aydin@iuc.edu.tr, bekdas@iuc.edu.tr, melihnig@iuc.edu.tr
[2] Department of Architecture, Mimar Sinan Fine Arts University, 34427 Istanbul, Turkey
umit.isikdag@msgsu.edu.tr

Abstract. This study aims to predict concrete compressive strength using machine learning methods applied to an open-source dataset compiled from various sources. The dataset encompasses diverse variables, such as concrete constituent properties and curing age, and underwent necessary data preprocessing steps prior to analysis. Multiple machine learning algorithms were employed, and their performances were compared. Model success was evaluated using cross-validation and an independent test dataset, based on RMSE, MAE, and R^2 performance metrics. The results revealed significant differences in predictive performance among the models. The comparisons identified the XGBoost model as the most successful, demonstrating low error values and high R^2 scores on both the cross-validation and test datasets. Additionally, SHAP analysis was conducted to determine which variables most significantly influenced the predictions of the best-performing model. The SHAP results indicated that the water-to-binder (W/B) ratio and curing age were the variables contributing most substantially to the prediction of concrete compressive strength.

Keywords: Machine learning, XGBoost, Compressive strength, SHAP, Prediction.

1 Introduction

Artificial intelligence (AI) is a research field that aims to enable machines to perform cognitive functions -such as learning and decision-making- characteristic of human intelligence. Today, AI is widely used in various technological applications and offers significant capabilities for analyzing multidimensional and complex data, particularly through machine learning (ML) [1]. Machine learning encompasses algorithms capable of making predictions or classifications by learning from data, representing an approach that allows computers to learn without being explicitly programmed [2].

To date, various studies have been conducted on the use of ML models in civil engineering problems. Examples of these include compressive strength prediction of blast furnace slag and fly ash concrete [3], fatigue prediction of composite materials for wind turbine blades [4], structural optimization of lattice shells [5], carbon emission prediction of Alkali-Activated Concrete [6], prediction of the load-carrying capacity of engineered cementitious composites (ECC)-strengthened RC beams, design and behavior of steel fiber-reinforced concrete flat slabs [7], prediction of splitting tensile strength of basalt fiber-reinforced concrete [8], predicting the bond strength between steel and CFRP plate [9], soil classification [10], real-time structural integrity assessment of bridges [11], prediction of the ultimate flexural capacity of BFRP-reinforced concrete beams [12], cooling load prediction of a double-story terrace house [13]. These studies from the literature demonstrate the potential of ML algorithms for application in civil engineering.

2 Research Methodology

The primary research question of the study is to observe the performance of various machine learning regression algorithms in predicting compressive strength using concrete mixture components and curing age. In the dataset used in the study, a total of 10 input variables were employed to predict concrete compressive strength (CS): cement, fly ash (FA), ground granulated blast-furnace slag (GGBS), silica fume (SF), desulfurized gypsum (DG), water/binder ratio (W/B), sand, coarse aggregate (CA), superplasticizer (SP), and curing age. The target output variable is concrete compressive strength (CS).

During the data preprocessing stage, completely empty rows were first removed from the dataset. Records containing missing data were excluded without applying any imputation methods, thereby preventing any potential impact of missing values on model results. Following the cleaning process, the dataset was randomly split into two groups: an 80% training set and a 20% independent test set. While the training set was used for model development and selection, the test set was kept entirely separate from the development process and used solely for the final performance evaluation.

To compare prediction performance, five different regression algorithms were evaluated: Linear Regression, Random Forest, XGBoost, LightGBM, and CatBoost. Model comparison was conducted using 5-fold cross-validation on the training dataset. Root Mean Square Error (RMSE) served as the primary metric for evaluating model performance; for each model, the mean RMSE and standard deviation across the five folds were calculated. The model with the lowest mean cross-validation RMSE was identified as the most successful model for final evaluation. The best model identified was retrained using the entire training dataset and evaluated on an independent test dataset that had not been used during the model development process. Model performance was assessed using R^2, RMSE, and MAE metrics. Visualizations were generated to examine the prediction behavior of the selected best model in detail.

2.1 Dataset

The dataset used in the study was obtained from an open-source dataset [14]. This dataset, compiled from various sources by Tan (2026) [14], originally contains information on a total of 2,578 mix proportions and curing durations. However, after removing incomplete data, the dataset used for the analyses in this study consists of 1,720 rows. The dependent variable of the model is concrete compressive strength (CS). The independent variables are cement, fly ash (FA), ground granulated blast-furnace slag (GGBS), silica fume (SF), desulfurized gypsum (DG), water-to-binder ratio (W/B), sand, coarse aggregate (CA), superplasticizer (SP), and curing age.

2.2 Data Preprocessing

As part of the data cleaning process prior to analysis, the dataset was first examined for missing values. Records containing missing values were identified and removed from the dataset, and only complete observations were used in the analyses.

2.3 Computing Environment and Libraries Used

In this study, data analysis, data preprocessing, modeling, and visualization tasks were carried out in the Google Colab environment. The Pandas and NumPy libraries were utilized for data processing and analysis; Scikit-Learn [15], XGBoost, LightGBM, and CatBoost were used for constructing and training machine learning models; and Matplotlib [16] and Seaborn [17] were employed for visualization. All data processing, model training, cross-validation, and performance evaluation procedures were performed within the same Google Colab environment, ensuring consistency across the computational environment and the methods employed.

2.4 ML Models

In this study, various machine learning and ensemble learning models were employed to predict compressive strength. The models were selected with the aim of achieving the best results by comparing the performance of different prediction methods. Each model utilizes distinct mathematical and statistical techniques to capture the complex relationships and patterns within the dataset. The primary models used in the study are described below.

Linear Regression is used to model the linear relationship between a dependent variable and one or more independent variables. Random Forest is an ensemble learning method composed of a combination of multiple decision trees. Each decision tree is trained on a random subset of the dataset, and the final prediction is made by averaging the predictions of all the trees [18]. Extreme Gradient Boosting (XGBoost) is an ensemble learning method based on the gradient boosting algorithm; it improves prediction accuracy by sequentially combining weak learners to minimize the error function [19]. Categorical Boosting (CatBoost) is a model developed for gradient boosting with decision trees that can utilize categorical data directly without the need for preprocessing [20]. Light Gradient Boosting Machine (LightGBM) was developed

with the aim of enhancing the speed and predictive efficiency of the GBM algorithm. By adopting a leaf-wise strategy, the method minimizes the error rate [21].

2.5 Model Performance Evaluation

The dataset was split into 80% for training and 20% for testing to train and evaluate the models. 5-fold cross-validation was employed to assess the model's generalization capability. In this method, the dataset was divided into five equal parts; each part was used as the test set in turn, while the remaining four parts served as the training set. Coefficient of Determination (R^2), Root Mean Squard Error (RMSE), and Mean Absolute Error (MAE) metrics were used to evaluate the training, testing, and cross-validation performance of the models.

MAE calculates the average of the errors. RMSE, being the square root of MSE, assigns greater weight to large errors and measures the standard deviation of the errors. R^2 is the coefficient of determination that indicates the extent to which independent variables explain the variance in the dependent variable [22].

3 Results

Table 1 compares the performance of the models based on both cross-validation and test set results. The XGBoost model outperformed the others, achieving an average value of 4.5551. XGBoost also demonstrated the best performance on the test set, yielding the lowest error rates (RMSE=4.3316, MAE=3.0189) and an R^2 of 0.9489. The LightGBM model also exhibited high predictive performance on the test set, with RMSE and MAE values of 4.5514 and 3.2078, respectively, and explained approximately 94% of the variance in the target variable with an R^2 value of 0.9435. Conversely, the Linear Regression model produced higher error values compared to the other models -with an RMSE of 10.2793 and an MAE of 7.8681- and achieved an R^2 value of 0.7122. This result indicates that, under the examined dataset and experimental conditions, the linear model performed less effectively than tree-based ensemble methods in capturing non-linear relationships.

Overall, the XGBoost and LightGBM models demonstrated the best performance, characterized by low error values and high R^2 results. While the CatBoost and Random Forest models also yielded comparatively successful results, the performance of the Linear Regression model was relatively lower. These findings indicate that, within the scope of the dataset and experimental framework used, tree-based ensemble models provided superior results compared to the linear regression model.

Table 1. Cross-validation (CV) and test metrics of the models used.

Model	CV RMSE (Mean)	CV RMSE (Std)	R^2 (Test)	RMSE (Test)	MAE (Test)
XGBoost	4.5551	0.2865	0.9489	4.3316	3.0189
LightGBM	4.6659	0.3860	0.9435	4.5514	3.2078
CatBoost	5.2846	0.2196	0.9264	5.1969	3.8509

Random Forest	5.8169	0.3039	0.9205	5.4029	3.7551
Linear Regression	10.5603	0.7190	0.7122	10.2793	7.8681

The most successful model, based on cross-validation results, was re-evaluated using an independent test dataset. The agreement between the actual and predicted concrete compressive strength values is shown in Fig. 1. Additionally, a residual plot was generated to assess the distribution of prediction errors (Fig. 2). SHAP analysis was conducted to determine the extent to which the model's predictions were influenced by specific input variables. The SHAP summary plot presented in Fig. 3 illustrates the relative effects of the variables on the model output.

An examination of the graphs in Figs. 2 and 3 -showing the predicted vs actual values and the residual plot for the XGBoost model- reveals that the model's predictions track the actual values well and exhibit low error rates. The model accurately captures trends in the dataset and successfully minimizes significant deviations.

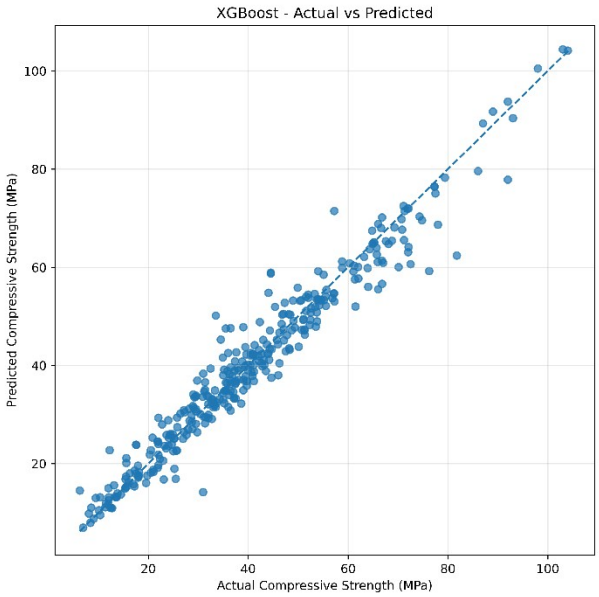


Fig. 1. Actual vs Predicted plot for XGBoost.

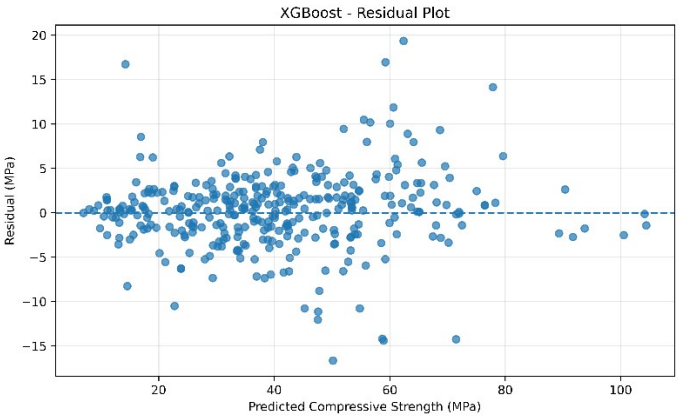


Fig. 2. Residual plot for XGBoost.

The SHAP importance ranking shown in Fig. 3 shows that the mixture's W/B ratio and curing age are the most dominant factors influencing the XGBoost model's prediction performance compared to other variables. The Mean Absolute SHAP value of 8.20 for the W/B ratio indicates that this variable makes the highest average absolute contribution to the model outputs. Curing age ranks second with a value of 7.39, while the impact of cement content is determined to be 3.47. Notably, the lower SHAP values associated with the DG, GGBS, and FA variables indicate that their contributions to prediction variability are relatively limited within the scope of the current data and model.

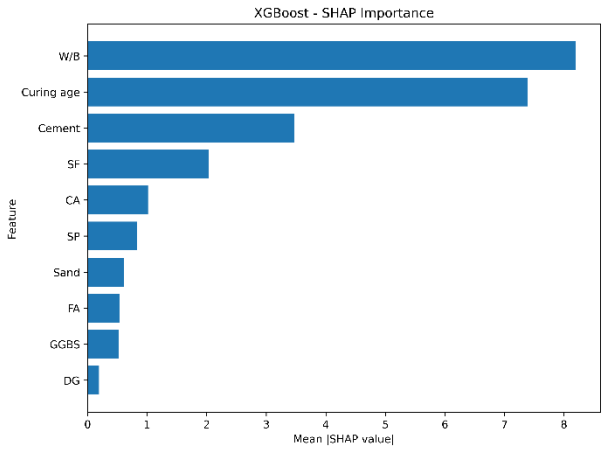


Fig. 3. SHAP importance for XGBoost.

4 Conclusions

This study aims to predict concrete compressive strength using machine learning methods, utilizing the open-source 'Multi-source Modified Concrete Mechanical and Durability Dataset [14]'. The predictive performances of various machine learning models were compared using cross-validation and an independent test dataset, and model success was evaluated based on RMSE, MAE, and R^2 metrics.

The results revealed that the machine learning models employed exhibited varying levels of performance in predicting concrete compressive strength. Among the models, XGBoost was identified as the most successful approach when considering results from both cross-validation and the independent test dataset. The low error values and high R^2 value achieved by the XGBoost model (RMSE=4.3316, MAE=3.0189, R^2=0.9489) demonstrate its ability to successfully capture the non-linear relationships between the dataset variables and compressive strength. It was observed that evaluating model performance based on success in the independent test dataset -rather than relying solely on training or cross-validation results- yields more reliable outcomes.

Additionally, SHAP analysis was conducted to assess the interpretability of the predictions made by the XGBoost model, which demonstrated the best performance. The results indicated that the water-to-binder ratio (W/B) and curing age were the variables with the greatest impact on the model's predictions.

In conclusion, the performance achieved by the XGBoost model, in particular, suggests that machine learning can serve as a data-driven tool during the preliminary assessment phase of experimental studies. The fact that the dataset used in this study aggregates data from numerous distinct sources has enabled the joint evaluation of information derived from diverse experimental conditions and concrete mixtures. In this respect, the study demonstrates the applicability of machine learning methods to data obtained from various sources, rather than relying on modeling based on a single experimental dataset. However, it would be more appropriate to view machine learning models as tools that support the planning of experimental programs and the preliminary assessment of suitable mixture alternatives.

References

[1] L. Oneto *et al.*, "Informed machine learning for complex data," *Neurocomputing*, vol. 669, p. 132505, Mar. 2026, doi: 10.1016/j.neucom.2025.132505.

[2] B. Ginsburg, "Application case study—machine learning," in *Programming Massively Parallel Processors*, Elsevier, 2017, pp. 345–367. doi: 10.1016/B978-0-12-811986-0.00016-9.

[3] M. I. Khan and Y. M. Abbas, "Robust extreme gradient boosting regression model for compressive strength prediction of blast furnace slag and fly ash concrete," *Mater. Today Commun.*, vol. 35, p. 105793, Jun. 2023, doi: 10.1016/j.mtcomm.2023.105793.

[4] Y. Aydın, C. Cakiroglu, G. Bekdaş, and Z. W. Geem, "Fatigue Predictive Modeling of Composite Materials for Wind Turbine Blades Using Explainable Gradient Boosting Models," *Coatings*, vol. 15, no. 3, p. 325, Mar. 2025, doi: 10.3390/coatings15030325.

[5] C. Kupwiwat and L. Prasittisopin, “Multi-task machine learning for structural optimizations of lattice shells,” *Eng. Struct.*, vol. 358, p. 122654, Jul. 2026, doi: 10.1016/j.engstruct.2026.122654.

[6] Y. Aydın *et al.*, “Neural Network Predictive Models for Alkali-Activated Concrete Carbon Emission Using Metaheuristic Optimization Algorithms,” *Sustainability*, vol. 16, no. 1, p. 142, Dec. 2023, doi: 10.3390/su16010142.

[7] A. Babiker, Y. M. Abbas, M. I. Khan, and R. Masmoudi, “Enhancing design and behavioral understanding of steel fiber-reinforced concrete flat slabs through a robust machine learning framework,” *Journal of Building Engineering*, vol. 111, p. 113133, Oct. 2025, doi: 10.1016/j.jobe.2025.113133.

[8] C. Cakiroglu, Y. Aydın, G. Bekdaş, and Z. W. Geem, “Interpretable Predictive Modelling of Basalt Fiber Reinforced Concrete Splitting Tensile Strength Using Ensemble Machine Learning Methods and SHAP Approach,” *Materials*, vol. 16, no. 13, p. 4578, Jun. 2023, doi: 10.3390/ma16134578.

[9] I. Afriadi, C. Thongchom, D. R. Kumar, S. Keawsawasvong, and W. Wipulanusat, “Ensemble boosting-based soft-computing models for predicting the bond strength between steel and CFRP plate,” *Sci. Rep.*, vol. 15, no. 1, p. 20459, Jul. 2025, doi: 10.1038/s41598-025-04866-9.

[10] Y. Aydin, G. Bekdas, U. Isikdag, S. M. Nigdeli, and Z. W. Geem, “Optimizing artificial neural network architectures for enhanced soil type classification,” *Geomechanics and Engineering*, vol. 37, no. 3, pp. 263–277, 2024.

[11] S. Jayasinghe *et al.*, “Application of Machine Learning for Real-Time Structural Integrity Assessment of Bridges,” *CivilEng*, vol. 6, no. 1, p. 2, Jan. 2025, doi: 10.3390/civileng6010002.

[12] S. Jueyendah and E. Ağcakoca, “An Interpretable Ensemble Machine Learning Framework for Predicting the Ultimate Flexural Capacity of BFRP-Reinforced Concrete Beams,” *Polymers (Basel).*, vol. 18, no. 5, p. 601, Feb. 2026, doi: 10.3390/polym18050601.

[13] C. Cakiroglu, Y. Aydın, G. Bekdaş, U. Isikdag, A. N. Sadeghifam, and L. Abualigah, “Cooling load prediction of a double-story terrace house using ensemble learning techniques and genetic programming with SHAP approach,” *Energy Build.*, vol. 313, p. 114254, Jun. 2024, doi: 10.1016/j.enbuild.2024.114254.

[14] T. Tan, “Multi-source Modified Concrete Mechanical and Durability Dataset,” Mendeley Data, V1. Accessed: Sep. 15, 2026. [Online]. Available: https://data.mendeley.com/datasets/3ph7nk4jzz/1

[15] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[16] J. D. Hunter, “Matplotlib: A 2D Graphics Environment,” *Comput. Sci. Eng.*, vol. 9, no. 3, pp. 90–95, 2007.

[17] Seaborn, “seaborn.pairplot — seaborn 0.13.2 documentation.” Accessed: Jul. 24, 2025. [Online]. Available: https://seaborn.pydata.org/generated/seaborn.pairplot.html

[18] L. Breiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[19] T. Chen and C. Guestrin, “XGBoost,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA: ACM, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[20] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: unbiased boosting with categorical features,” 2017.

[21] M. Hajihosseinlou, A. Maghsoudi, and R. Ghezelbash, “A Novel Scheme for Mapping of MVT-Type Pb–Zn Prospectivity: LightGBM, a Highly Efficient Gradient Boosting Decision Tree Machine Learning Algorithm,” *Natural Resources Research*, vol. 32, no. 6, pp. 2417–2438, Dec. 2023, doi: 10.1007/s11053-023-10249-6.

[22] K. Li, “Informativeness of Performance Measures: Coefficients or R-Squareds?,” *Journal of Risk and Financial Management*, vol. 17, no. 11, p. 481, Oct. 2024, doi: 10.3390/jrfm17110481.

Particle-Tuned Mass Dampers for Structural Vibration Control: A Review of Modeling, Design and Applications

Aşan Şenlik, Yaren Aydın, Sinan Melih Nigdeli and Gebrail Bekdaş

Department of Civil Engineering, Istanbul University-Cerrahpaşa,
Istanbul, 34320, Türkiye
asan.senlik@ogr.iuc.edu.tr, yaren.aydin@iuc.edu.tr, melihnig@iuc.edu.tr, bekdas@iuc.edu.tr

Abstract. Every structure exhibits a certain degree of vibration under dynamic loading. In some structures, these vibrations may reach significant levels. Tall and slender buildings and structures susceptible to torsional responses are examples of structures that may require vibration control measures. Different approaches have been developed to limit the vibrational response of structures. This review examines the modeling, design, and applications of particle-tuned mass dampers (PTMDs) for structural vibration control. A PTMD is a passive vibration control system consisting of a rigid container connected to the primary structure through flexible elements, with freely moving granular particles placed inside the container at a specified filling ratio. When the structure vibrates under dynamic excitation, the relative motion of the particles within the container generates collisions and frictional interactions. These interactions dissipate vibrational energy and contribute to the reduction of structural vibrations. The performance of PTMD systems depends on several parameters, including particle properties, filling ratio, container geometry, connection stiffness, and the dynamic characteristics of the primary structure. The vibration reduction capability of a PTMD system is strongly influenced by how these parameters are configured. Compared with conventional tuned mass dampers (TMDs), PTMDs provide an alternative approach to passive structural vibration control by incorporating additional energy dissipation mechanisms through particle interactions. However, the effectiveness of these systems depends on the characteristics of the dynamic loading and the system configuration. This review discusses the working principles, design characteristics, and structural applications of PTMD systems.

Keywords: The Particle-Tuned Mass Damper (PTMD), Structural Vibration Control, Granular Particles.

1 Introductıon

Structures are subjected to various types of loads. The most common of these are dynamic loads such as earthquakes, wind, traffic, and moving vehicle loads. Under the influence of these loads, structures vibrate. The need to limit these vibrations under dynamic conditions has led to the development of structural control systems. Structural control systems are classified as passive, active, semi-active, and hybrid

systems according to their operating principles [1]. Passive control systems operate by converting vibration energy without the need for an external power source [1]. These systems collect energy through masses and springs attached to the structure and use it to dampen the structural response [1]. Their lack of need for an external power source makes passive systems more economical than other alternatives [1]. Tuned mass dampers (TMDs) are among the most extensively researched and frequently implemented traditional passive control systems in structural engineering [2]. Nevertheless, existing TMD systems suffer from chronic disadvantages, including a narrow operational frequency bandwidth, limited structural lifespan, and a decline in damping capacity over time [2]. To eliminate these constraints, researchers have shifted their focus toward developing alternative passive control mechanisms capable of mitigating the shortcomings of conventional TMD [2]. As a result of these investigations, a particle-tuned mass damper system has been developed. By combining the mechanisms of traditional tuned mass dampers (TMDs) and particle dampers (PDs), the concept of a particle tuned mass damper (PTMD) has been established [3]. According to [4], particle dampers (PDs) are inexpensive to build, geometrically versatile, and highly resistant to environmental factors like to address these limitations, the researchers introduced the concept of a particle tuned mass damper (PTMD). By combining the frequency-insensitive robustness of PDs with the high energy dissipation capacity of TMDs, this hybrid device achieves enhanced structural performance and built-in damping adjustability.

1.1 Methodology

In this study, more than ten scientific publications were reviewed to investigate the application of particle-tuned mass dampers (PTMDs) in structural vibration control. The literature search was conducted using platforms such as ScienceDirect and Wiley Online Library. Keywords such as “particle tuned mass damper,” “particle damper,” “structural vibration control,” “granular damping,” and “tuned mass damper” were used to identify relevant publications. The reviewed studies focused on the mathematical modeling, design parameters, vibration mitigation performance, and applications of PTMD systems in different structural systems. Furthermore, the differences between PTMD systems and conventional tuned mass dampers (TMDs) were evaluated.

2 Modeling, Design, Applications of Particle Tuned Mass Dampers

Below, the modeling, design, and applications of particle tuned mass dampers, used as an alternative to other structural control systems, are described.

2.1 Modeling of Particle Tuned Mass Dampers

As a passive vibration reduction method, the particle-tuned mass damper (PTMD) basically consists of a spring, a damper, a container and particles inside it [5] ,[6]. Fig. 1. illustrates a two-dimensional (2D) cross-sectional model of a PTMD system installed in a building. Fig. 2. shows a schematic representation of an impact damper and a particle damper mounted on the end of a cantilever beam under dynamic load to better understand the PTMD system.

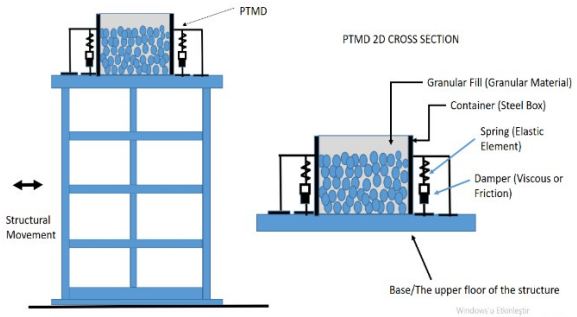


Fig.1. The PTMD system is located on the building.

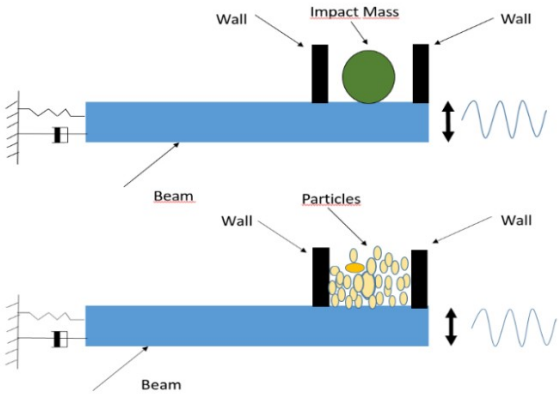


Fig.2. Schematic representation of an impact damper (upper) and a particle damper (lower) mounted on the tip of a cantilever beam under dynamic load [7].

2.2. Design Parameters of Particle Tuned Mass Dampers

In this section, in accordance with established literature, the "Equivalent Single Particle Damper" approach is investigated to understand the vibration reduction performance of the particle tuned mass damper (PTMD) in the primary structure under seismic excitation. The basic principle of this method is to simplify the highly

nonlinear, chaotic, and random collision dynamics between a large number of particles within the damper vessel. Under this approach, the entire group of particles is reduced to a single equivalent spherical mass, strictly adhering to the principles of conservation of mass, material density, and void volume [8], [9].

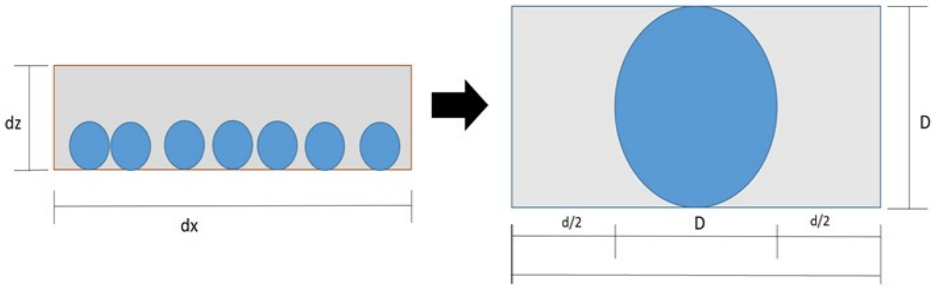


Fig.3.Schematic diagrams illustrating the particle tuned mass damper and the equivalent single-particle damper [8].

In the PTMD system, based on the assumptions presented here, a single particle replaces a group of particles. Regarding container geometry, the PTMD features a cuboid whose height is restricted to the particle diameter (Dp), whereas the single-particle damper utilizes a cylinder with a bottom diameter equivalent to that of a single particle (D). The void volume of the single-particle damper ($Veid$) is equivalent to that of the PTMD ($Vepd$). The combined mass of the particles with-in the PTMD (m) is identical to the mass of the single particle (m_p). Both damping systems utilize spherical particles with a uniform density designated as ρ. Fig. 3 shows the corresponding simplification diagram. The simplified method is as follows: By applying Eq. (1), the volume of particles for the PTMD $Vspd$ is obtained [8] :

$$Vspd = \frac{1}{6} N \pi D p^3 \tag{1}$$

The total particle count within the PTMD is denoted by N, which is determined by dividing the total mass (m) by the mass of a single particle ($m1p$). The void volume of the PTMD $Vepd$ is

$$Vepd = Vpd - Vspd = \left(\frac{1}{\rho p} - 1\right) Vspd \tag{2}$$

In this equation, the PTMD volume is given by (Vpd) and the ratio and $\rho_p = Vspd/Vpd$ defines the packing density of the PTMD. where $Veid$ is the void volume of the single-particle damper $Veid$ is calculated using the following formula:

$$Veid = \frac{\pi}{12} \times D^3 + \frac{\pi}{4} \times D^2 d \tag{3}$$

The value of d the clearance of the simplified single-particle damper can be obtained from Eq. (4) based on the equivalence principle (1), [8] :

$$\left(\frac{1}{\rho p}-1\right)\frac{m}{\rho}=\frac{mp}{2\rho}+\frac{\pi}{4}\left(6\frac{mp}{\pi\rho}\right)^{2/3}d \quad (4)$$

Where the material density is denoted by ρ, and mp defines the equivalent single particle mass $m_p = m$. By applying the equivalent single particle transformation, the matrix form governing equations of motion capturing the dynamic response of the combined primary structure PTMD system under ground motion $\ddot{x}_g$ are derived as [8]:

$$M\ddot{X}+C\dot{X}+KX=E\ddot{x}g+\phi(c_c\,\dot{y}_1+k_c y_1) \quad (5)$$

$$m_c\,\ddot{x}_c+c_c\dot{y}_1+k_c y_1-c_p H(y_2,\dot{y}_2)-k_p G(y_2)=0 \quad (6)$$

$$m_p\ddot{x}_p+c_p H(y_2,\dot{y}_2)+k_p G(y_2)=0 \quad (7)$$

In the equations given above, M represents the mass matrix, C represents the damping matrix, and K represents the stiffness matrix. The nonlinear boundary and impact functions, expressed as $G(y_2)$ and $H(y_2, \dot{y}_2)$ in Equations (6) and (7), come into effect during impacts between the equivalent particle and the container walls. The operational limits of these functions are established by the prior calculation of the free gap clearance (d).

The proposed single-particle approach facilitates highly efficient numerical solutions of the governing equations of motion within the MATLAB environment. By decoupling critical design parameters from intricate inter-particle friction, this framework enables a direct evaluation of their structural impacts. Ultimately, this research validates the capability of the modeling strategy to predict the structure's nonlinear response trends and peak behaviors [8].

As a lightweight assembly positioned at the upper part of the primary structure, the Particle Tuned Mass Damper (PTMD) remarkably suppresses both the peak and r.m.s. responses under seismic actions and free vibration. This dampening process is driven by the frictional interactions and collisions among the particles themselves, as well as against the container boundaries, alongside the frequency tuning provided by the added masses, which highlights its strong potential for structural engineering practices.

Parametric evaluations reveal that the particle filling ratio serves as the most decisive factor governing the operational behavior of the PTMD. The maximum levels of mitigation are predicted to occur when this ratio is maintained within the boundaries of 20–30%, 50–60%, or 80–90%. When an ideal filling ratio is established, the auxiliary mass ratio does not fundamentally dictate the overall dampening performance; however, escalating the particle density tends to degrade the efficiency of vibration suppression under an appropriate filling ratio.

Consequently, the particle filling ratio must be prioritized during the optimization process of a PTMD layout. The systematic design framework initiates with selecting

the particle density, wherein a lower density is clearly preferred. Subsequently, the allowable range for the auxiliary mass ratio is defined based on practical engineering constraints. Next, a suitable filling ratio is designated from the 20–30%, 50–60%, or 80–90% brackets. Following this, the introduced analytical simulation routine is employed to compute the optimal mass ratio that matches the selected filling ratio, and finally, the material composition along with the geometric features of the container are determined. Additionally, the developed simplified numerical strategy is intrinsically linked to the filling ratio, thereby offering a dependable tool for early-stage design configurations [10].

2.3 Applications of Particle Tuned Mass Dampers

PTMD has various application areas; depending on the fragility of the structures of these systems and the type of dynamic loads they are subjected to, they undertake critical protection operations in different engineering sectors. It is used in areas such as multi-story buildings, wind turbines, and energy distribution lines [11], [12],[13].

In multi-story buildings, it is used to control wind and earthquakes caused by environmental conditions. The dynamic energy from the outside is transferred to the damper, dissipated by hitting the cold container walls, and the acceleration and displacement values of the resistance under seismic shocks are reduced [11].

This technology is used in wind turbine towers to reduce the dangerous dynamic reactions created by combined wind and earthquake loads in the tower body and upper regions. To maximize the damping process, the system is placed at the top of the tower structure, and large-sized particles are preferred in the chamber [12].

It is used as a powerful protection tool to minimize seismic effects in energy distribution lines and lattice towers. This technology, which can reduce the response of transmission towers during an earthquake, ensures the rapid dissipation of seismic energy by transferring the energy to high frequencies thanks to the contamination of the particles inside [13].

3 Literature Review on Particle-Mounted Mass Dampers (PTMDs)

This section presents some previous studies on PTMD systems.

A study on the optimization of a PTMD system investigated the effects of some fundamental parameters on vibration control. The research showed that the amount of damping is largely dependent on the particle occupancy rate. Based on free vibration and shaking table tests conducted on a single-story steel frame, guidelines focusing on the filling ratio were proposed for optimal PTMD design. Furthermore, an approximate analytical method based on an equivalent single-particle damper concept was developed, showing good agreement with experimental results. Finally, this simplified approach was applied in a case study to optimize a PTMD system for a five-story steel structure [10].

In another study, the effectiveness of a PTMD system was investigated using a small, single-story steel structure. Two scenarios were examined in this experiment. In one scenario, the damper was suspended from the structure by a fixed weight in a container, while in the other, the fixed weight was replaced by particles of equal mass filling two compartments of different sizes. In both cases, the damper frequency was close to the natural frequency of the structure. Under random excitation, it was determined that the root mean square (RMS) of the response acceleration decreased in both systems, but the particle system provided more effective damping. The decrease in the RMS of the relativistic displacement was observed after the first few seconds of excitation, and it was found that the effect of the excitation level on the effectiveness of the PTMD was limited [5].

Another study investigates the energy dissipation performance of a particle-tuned mass damper (PTMD) under wind load in a 76-storey building using a simplified model and validates it with a wind tunnel experiment. Examining all energy components in the system and the energy transfer between the structure and the PTMD, the study found that the incoming wind energy is rapidly transferred to the damper in the first few seconds, then dissipated as a result of collisions between the particles and the container wall. Finally, the interaction of parameters such as volume occupancy, mass ratio, damping coefficient, and frequency ratio was investigated using full factorial
analysis, confirming the significant impact of these parameter combinations on damping performance and proposing a novel design procedure [11].

Another study investigates the vibration reduction performance of a Particle Tuned Mass Damper (PTMD) on a 1/20 scale wind turbine tower model using experimental and numerical methods. Shaking table tests and finite element simulations show that the PTMD significantly reduces the dynamic responses at the tower body and top under wind and seismic loads. Parametric analyses reveal that the system has good robustness and a wide frequency band, while simulations confirm that aerodynamic damping also limits vibration to a certain extent. The consistent experimental and numerical results demonstrate that the PTMD significantly mitigates vibration in wind turbine towers under combined loads [12].

Another article experimentally and analytically investigates the damping performance of a PTMD system under wind load, based on a 1:200 scale reference model. The study explores the effects of parameters such as pendulum length, wind speed, mass ratios, and particle material on vibration reduction and provides accurate predictions consistent with experimental results by developing an analytical solution based on the concept of an equivalent single-unit impact damper. Experimental and simulation results demonstrate the robustness of the novel damper while also showing that damping performance can be improved by controlling particle density, increasing the number of particles, and intensifying the effect of the particles [3].

A study investigated the parameters and mounting locations of particle-tuned mass dampers (PTMDs) for controlling structural vibrations in wind turbine tower structures. The findings revealed that mounting the PTMD system at the top of the tower structure provided a significant optimization in vibration damping, and that the

large-sized particles used within the damper maximized the vibration reduction performance [14].

A study investigates the seismic mitigation effectiveness of particle-tuned mass dampers (PTMDs) in transmission tower-to-line systems (TTLS) using a finite-discrete element matching simulation model. The model, validated by shake table tests, demonstrates that PTMDs reduce the seismic response of towers by a maximum of 45.6%, and that particle collisions within them accelerate energy dissipation by transferring vibrational energy to higher frequencies. Furthermore, the study highlights the limitations of simplification methods used in previous studies, showing that while lower and upper numerical limits for displacement can be found in single-particle PTMDs, these limits are less pronounced in multi-particle systems [13].

In a study, the vibration control effects of a particle-tuned mass damper (PTMD), proposed by integrating a conventional tuned mass damper and a particle damper, were investigated using experimental, analytical, and computational approaches. Shaking table tests were conducted on a five-story steel frame to investigate the effects of parameters such as mass ratio, gap width, frequency, and input amplitude level on damping. The results showed that the PTMD provides high damping and increases the frequency bandwidth. Furthermore, an approximate analytical solution based on the concept of an equivalent single-particle damper was presented, and the accuracy of this method was proven with experimental and numerical data [8].

A study presents a novel Abaqus and Python-based combined analysis method (ESM-FEM) to investigate the vibration control performance of particle-tuned mass dampers (PTMDs) used in multi-degree-of-freedom (MDOF) structures such as power transmission towers. The accuracy of this developed method is demonstrated by comparing it with previous academic studies, experimental data, and MATLAB simulations under free vibration and seismic loads. The proposed model greatly simplifies the generation of the damping force, leading to a better understanding of the system's operating mechanism, while the results show that the PTMD system has high robustness and exhibits reliable control performance even against unpredictable seismic shocks [15].

4 Conclusions

This review highlights that the Particle-Tuned Mass Damper (PTMD) represents a highly efficient and resilient evolution in the field of passive structural vibration control. While conventional Tuned Mass Dampers (TMDs) have historically served as the benchmark for mitigating dynamic responses in tall and slender structures, their practical utility remains constrained by narrow operational bandwidths and long-term mechanical degradation. By synergizing the frequency-tuned mass displacement of a traditional TMD with the multi-particle impact mechanics of a Particle Damper (PD), the PTMD effectively circumvents these classic tuning vulnerabilities.

The core strength of the PTMD system lies in its multi-layered energy dissipation framework. As dynamic forces like wind or seismic activity induce structural drift, the autonomous, chaotic movement of the internal granular particles triggers

thousands of microscopic collisions and frictional interfaces. This process successfully converts kinetic energy into thermal dissipation without requiring secondary external power, offering a highly sustainable, robust, and cost-effective alternative for real-world infrastructure protection. Designing an optimized PTMD demands precise calibration of interconnected variables, spanning container layout, connection stiffness, particle material density, and filling ratios. To bypass the extreme computational challenges of modeling these stochastic particle interactions, the Equivalent Single Particle Damper method serves as an invaluable analytical bridge. Reducing the distributed particle mass into a singular, geometrically equivalent sphere while maintaining strict mass and void volume boundaries allows structural engineers to accurately predict damping performance under severe transient loads. Ultimately, PTMD technology provides an adaptable, low-maintenance pathway toward next-generation structural resilience. To further advance its deployment, subsequent research should prioritize investigating irregular or irregular-shaped container configurations under multi-axis dynamic loads, alongside testing the long-term structural wear profiles of diverse granular alloys during continuous cyclic loading.

References

[1] A. Ocak, S. M. Niğdeli, and G. Bekdas, “Pasif Kontrol Sistemleri,” in *Ayarlı Kütle Sönümleyicilerinin Optimum Ayarlanması*, İstanbul Üniversitesi–Cerrahpaşa Yayınevi, 2024. doi: 10.5152/3001.

[2] Z. Lu, K. Li, and Y. Zhou, “Comparative Studies on Structures with a Tuned Mass Damper and a Particle Damper,” *J. Aerosp. Eng.*, vol. 31, no. 6, Nov. 2018, doi: 10.1061/(ASCE)AS.1943-5525.0000878.

[3] Z. Lu, D. Wang, S. F. Masri, and X. Lu, “An experimental study of vibration control of wind-excited high-rise buildings using particle tuned mass dampers,” *Smart Struct. Syst.*, vol. 18, no. 1, pp. 93–115, Jul. 2016, doi: 10.12989/sss.2016.18.1.093.

[4] S. Simonian and S. Brennan, “Particle Tuned Mass Dampers: Design, Test, and Modeling,” in *46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Reston, Virigina: American Institute of Aeronautics and Astronautics, Apr. 2005. doi: 10.2514/6.2005-2325.

[5] A. Papalou, “An Experimental Study on the Effectiveness of Particle Tuned Mass Dampers,” *KSCE Journal of Civil Engineering*, vol. 27, no. 10, pp. 4437–4443, Oct. 2023, doi: 10.1007/s12205-023-0261-3.

[6] B. Yao, Q. Chen, H. Y. Xiang, and X. Gao, “Experimental and theoretical investigation on dynamic properties of tuned particle damper,” *Int. J. Mech. Sci.*, vol. 80, pp. 122–130, Mar. 2014, doi: 10.1016/j.ijmecsci.2014.01.009.

[7] B. B. Prasad, F. Duvigneau, T. Reinboth, D. Juhre, and E. Woschke, “Design Strategies of Particle Dampers for Large-Scale Applications,” *Journal of Vibration Engineering & Technologies*, vol. 12, no. 4, pp. 5715–5746, Apr. 2024, doi: 10.1007/s42417-023-01214-x.

[8] Z. Lu, X. Chen, D. Zhang, and K. Dai, “Experimental and analytical study on the performance of particle tuned mass dampers under seismic excitation,” *Earthq. Eng. Struct. Dyn.*, vol. 46, no. 5, pp. 697–714, Apr. 2017, doi: 10.1002/eqe.2826.

[9] A. Papalou and S. F. Masri, “Performance of Particle Dampers Under Random Excitation,” *J. Vib. Acoust.*, vol. 118, no. 4, pp. 614–621, Oct. 1996, doi: 10.1115/1.2888343.

[10] Z. Lu, X. Chen, and Y. Zhou, “An equivalent method for optimization of particle tuned mass damper based on experimental parametric study,” *J. Sound Vib.*, vol. 419, pp. 571–584, Apr. 2018, doi: 10.1016/j.jsv.2017.05.048.

[11] Z. Lu, J. Zhang, and D. Wang, “Energy analysis of particle tuned mass damper systems with applications to MDOF structures under wind-induced excitation,” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 218, p. 104766, Nov. 2021, doi: 10.1016/j.jweia.2021.104766.

[12] Z. Lu, S. Zhao, C. Ma, and K. Dai, “Experimental and analytical study on the performance of wind turbine tower attached with particle tuned mass damper,” *Eng. Struct.*, vol. 294, p. 116784, Nov. 2023, doi: 10.1016/j.engstruct.2023.116784.

[13] K. Rong, J. Gong, Z. Lu, L. Tian, Y. Shen, and M. Zhou, “Seismic mitigation and configuration-based performance study on particle tuned mass damper applied to transmission tower-line system,” *Eng. Struct.*, vol. 345, p. 121490, Dec. 2025, doi: 10.1016/j.engstruct.2025.121490.

[14] W. Xiao, S. Wang, Z. Cai, and Z. Xie, “Experimental investigation of particle-tuned mass dampers on vibration damping characteristics of wind turbine tower structures,” *J. Phys. Conf. Ser.*, vol. 3068, no. 1, p. 012162, Aug. 2025, doi: 10.1088/1742-6596/3068/1/012162.

[15] K. Rong and Z. Lu, “An improved ESM-FEM method for seismic control of particle tuned mass damper in MDOF system,” *Applied Acoustics*, vol. 172, p. 107663, Jan. 2021, doi: 10.1016/j.apacoust.2020.107663.

ABSTRACT General Session

Challenges and Implications of Project Management Practices Implementation for Engineering Project Performance in the Liberian Construction Industry

Martins Randoulf[1], Samuel Adekunle[2]

[1,2] *Civil Engineering Department, College of Engineering and Technology, William V.S. Tubman University, Harper City, Liberia*

Effective project management practices (PMPs) are widely recognised as critical determinants of construction project success; however, their systematic and effective application remains inconsistent, particularly in developing-country contexts. In Liberia, substantial public and donor-funded investments in infrastructure development have yet to translate consistently into satisfactory project outcomes, with engineering projects continuing to experience delays, cost overruns, quality deficiencies, and inefficient resource utilisation. The study also assessed the challenges constraining the effective implementation of PMPs and identified strategies for enhancing project performance through improved project management practices. A quantitative approach employing a descriptive research design was adopted. Data were collected from 50 construction professionals through a structured questionnaire using a five-point Likert scale. Descriptive statistics, including frequencies, percentages, mean scores, and standard deviations, were used to analyse the data. The findings indicate that inadequate funding, corruption, lack of transparency, and political interference constitute the most significant barriers to the effective implementation of PMPs. The findings underscore the need to strengthen project financing mechanisms, institutional accountability, transparency, and governance capacity to improve the effectiveness of project management practices and, consequently, engineering project performance in Liberia.

Keywords: *project management practices; implementation challenges; developing country, PM; Liberia*

ABSTRACT General Session

The Use of Plantain Leaf Ash as a Natural Pozzolan in Sandcrete Blocks for Sustainable Buildings

Ifeyinwa Ijeoma Obianyo[1*], Chinonso Obika[1], Uwemedimo Nyong Wilson[2], Assia Aboubakar Mahamat[3], Azikiwe Peter Onwualu[4]

[1] *Department of Civil Engineering, Nile University of Nigeria, Abuja, Nigeria*
[2] *Department of Civil Engineering, Nigerian Defence Academy, Kaduna, Nigeria*
[3] *Department of Civil Engineering, African University of Science and Technology, Abuja, Nigeria*
[4] *Department of Materials Science and Engineering, African University of Science and Technology, Abuja, Nigeria*
ifeyinwa.obianyo@nileuniversity.edu.ng

Cement is the most important material in construction and is produced in large quantities. Due to the mass production of cement globally, there is a significant release of greenhouse gases such as CO2, which impacts the environment negatively. The study aims to address the environmental and economic issues associated with cement production by utilizing agricultural waste, such as plantain leaves, as a sustainable alternative. By using these plantain leaves, the study also hopes to cut down the cost of cement by introducing a cheap substitute that can be used in combination with cement. The plantain leaf ash (PLA) underwent an XRF test to determine its pozzolanic properties before the process of block production. The process involves making sandcrete blocks with PLA replacing cement at different proportions (2, 4, 6, 8 and 10%) and making a control sample made with 100% cement for comparison. The mechanical properties were derived using the compressive strength test while the durability properties were derived using the water absorption test. It was observed that there was a decrease in the compressive strength of the Plantain leaf ash sample with the initial addition of PLA but showed a steady increase with increase in the % with the 10% replacement being the highest. This study contributes to reducing the environmental footprint of cement production while offering a compressive strength of the control sample was 3.2274MPa with the 10% replacement sample falling just above it at 3.238MPa. These findings highlight the potential of PLA as a sustainable and cost-effective material for partial cement replacement in sandcrete block production. By utilizing an innovative agricultural waste pathway for waste management, an eco-friendly block can be produced using PLA.

Keywords: *Plantain leaf ash; sandcrete blocks; partial cement replacement; compressive strength; water absorption; sustainable construction.*

ABSTRACT General Session

Evaluating Stormwater Drainage Systems in Urban Roads: A Case Study of Abuja Ring Road 2

Ifeyinwa Ijeoma Obianyo[a*], Muhammad Muazu Ibrahim[a], Abubakar Dayyabu[a], Assia Aboubakar Mahamat[b], Martin Agbili[c,d]

[a] *Department of Civil Engineering, Nile University of Nigeria, Abuja, Nigeria*
[b] *Department of Civil Engineering, African University of Science and Technology, Abuja, Nigeria*
[c] *Department of Environmental Management, Nnamdi Azikiwe University, Anambra State, Nigeria*
[d] *Ministry of Works & Infrastructure, Anambra State, Nigeria*
ifeyinwa.obianyo@nileuniversity.edu.ng

Rapid urbanization has significantly increased stormwater runoff in many developing cities, resulting in frequent flooding and accelerated deterioration of road infrastructure. Abuja, Nigeria's Federal Capital Territory, has experienced increasing drainage-related challenges due to expanding impervious surfaces, inadequate drainage capacity, poor maintenance practices, and changing rainfall characteristics. This study evaluated the performance of stormwater drainage systems along the 4.4 km Ring Road 2 corridor using field observations, hydraulic analysis, and historical rainfall records obtained from the Nigerian Meteorological Agency (NiMet). Peak runoff was estimated using the Rational Method, while the hydraulic capacity of existing drainage channels was assessed using Manning's Equation. Results showed that drainage performance varied considerably across the study corridor. Airport Junction recorded the best hydraulic performance, with an estimated peak discharge of 1.4 m^3/s and adequate channel capacity despite localized siltation, vegetation growth, and refuse blockage. Nile Junction operated close to its design capacity, indicating a high risk of overflow during more intense rainfall events. In contrast, CITEC Estate, Major Oil Filling Station, and Idu Junction lacked functional drainage infrastructure, resulting in persistent surface ponding, localized flooding, and pavement deterioration. The average estimated peak discharge for the corridor was 1.37 m^3/s, highlighting the increasing hydraulic demand caused by rapid urbanization. The study concludes that the existing drainage system is hydraulically inadequate and recommends the construction of additional drainage channels, upgrading of undersized structures, routine desilting, and implementation of sustainable maintenance strategies to improve stormwater management and enhance the long-term performance of urban road infrastructure.

Keywords: *Stormwater drainage, Rational Method, Manning's Equation.*

ABSTRACT General Session

Performance Evaluation of Interlocking Paving Blocks Produced with Ground Low-Density Polyethylene (LDPE) as a Partial Fine Aggregate Replacement

Ifeyinwa Ijeoma Obianyo, Muhammad Abu-Imran Ozeto, Abubakar Dayyabu, Akeem Amuda and Abdulhameed Danjuma Mambo

Department of Civil Engineering, Nile University of Nigeria, Abuja, Nigeria
ifeyinwa.obianyo@nileuniversity.edu.ng

Sand mining carries a high environmental cost and the price of fine aggregate keeps rising, while non-biodegradable low-density polyethylene (LDPE) waste continues to accumulate. This study evaluated interlocking paving blocks produced with ground LDPE as a partial replacement for fine aggregate, to establish how the replacement level affects mechanical and microstructural performance and to identify the optimum level. Six mixes were produced at LDPE replacement levels of 0, 5, 10, 15, 20 and 25% by weight of fine aggregate, using a nominal 1:2:4 mix and a constant water/cement ratio of 0.55. Materials were characterised by sieve analysis, specific gravity and X-ray fluorescence (XRF). Cubes of 70 × 70 × 70 mm were cured in water for 7 and 28 days and tested for bulk density, water absorption and compressive strength, and examined by scanning electron microscopy (SEM). XRF showed the ground LDPE to be about 82% organic polymer with only trace heavy metals. The 28-day compressive strength fell from 17.84 N/mm^2 for the control to 6.07 N/mm^2 at 25% replacement, a loss of about 66%, while the 5 and 10% mixes retained roughly 83 and 74% of the control strength. Water absorption rose from 2.02 to 4.78% and bulk density fell from 2592 to 2156 kg/m^3. SEM linked these changes to poor bonding between the hydrophobic LDPE and the cement paste. Replacement of 5 to 10% gives the best balance for structural paving, up to 15% suits light pedestrian use, and 20 to 25% suits only non-structural applications.

Keywords: *Low-density polyethylene, Interlocking paving blocks, Fine aggregate replacement, Compressive strength, Recycled plastic, Sustainable construction.*

ICASE 2026

Conference Topics

- Emerging Trends in Artificial Intelligence and Machine Learning
- Advancements in Renewable Energy Technologies
- Nanoengineering and Nanotechnology Applications
- Robotics and Automation in Modern Industry
- Smart Cities and Internet of Things (IoT) Solutions
- Advanced Materials for Future Engineering Applications
- Sustainable Engineering and Green Technologies
- Cybersecurity and Data Protection in Engineering
- Quantum Computing and Its Engineering Applications
- Bioengineering and Biomedical Innovations
- Advanced Computational Techniques in Engineering
- Engineering Education and Innovation
- Additive Manufacturing and 3D Printing
- Data Science and Big Data Analytics in Engineering
- Climate Change and Environmental Engineering
- Space Engineering and Technologies
- Structural Engineering Innovations
- Automotive Engineering and Electric Vehicles
- Telecommunication and Networking Engineering
- Artificial Intelligence for Sustainable Development Goals

İSTANBUL UNIVERSITY
- CERRAHPAŞA

ACT Research Group
Advanced Computational Techniques
act.iuc.edu.tr

SCAN TO VISIT

International Conference on Applied Sciences and Engineering · ICASE 2026
9–10 November 2026 · Online · https://icase.iuc.edu.tr/en/

ISBN: 978-625-00-5623-3