I. R. IRAN

ISSN: 1728-144X
e-ISSN: 1735-9244

International Journal of Engineering

Journal Homepage: www.ije.ir

MERC

TRANSACTIONS B: APPLICATIONS

Volume 39, Number 05, May 2026

Materials and Energy Research Center

ISSN: 1728-144X, e-ISSN: 1735-9244

International Journal of Engineering, Volume 39, Number 05, May 2026

CONTENTS

Transactions B: Applications

S. M. Sahebdad, M. Eftekhari, M. Eftekhari	Physics-Informed Neural Networks Techniques for Analyzing Forced Vibrations of Simply Supported Beams Featuring Variable Cross-Sections	1047-1062
W. Younus Abeid Al-Zubaidi, A. A. Ranjbar	Experimental and Numerical Investigations for Improvement of Building Ventilation Performance by Using Solar Chimney Integrated with Different Phase Change Materials and Nanoparticles	1063-1076
G. V. Buslaev, A. V. Konoplyannikov	Mathematical Modeling of a Hydro-mechanical System with a Positive Displacement Motor and Hydraulic Thruster to Optimize the Drilling Process for Extended-reach Drilling Wells	1077-1087
M. Ramezani, H. Falaghi, A. Azimi	Distribution Substation Allocation Considering Load Uncertainty based on Information Gap Decision Theory	1088-1100
R. K. Khoman, H. M. Owaid	Sustainable High-performance Self-compacting Concrete with Multi Blended of Materials Powders: Fresh and Hardened Properties	1101-1114
V. V. Nikishin, A. V. Mysin, D. S. Kuznetsova	Investigation of Cement Compositions of Backfill Systems for Lining Wells with Inclined and Horizontal Sections	1115-1121
E. Fatapour, A. Afrous, B. Aminnejad, A. Saremi, A. Khosrojerdi	Mapping Flood Zone in Khorram Abad River by Integrating Hydrological and Hydraulic Models	1122-1138
M. A. Ali Bash, A. N. Jasim, A. M. Resin, S. A. Ajeel	Effect of Radio Frequency Sputtering and Laser Remelting on Ti6Al4V Behavior in Medical Applications	1139-1151
A. V. Mysin, M. V. Bykova, E. E. Soloveva	Evaluation of Efficiency of Associated Petroleum Gas Utilization under Conditions of Low Atmospheric Humidity	1152-1160
P. Sahebgouda, S. Maradithaya, A. Jadhav	Optimized Deep Learning Model for Pomegranate Disease Detection: A Convolutional Neural Network Long Short-Term Memory Approach	1161-1175
A. Sadeghi, T. B. Gorji, M. Gorji-bandpy	Comparative Performance Evaluation of Savonius Vertical Axis Wind Turbines: Two and Three-bladed Profiles	1176-1190
V. V. Maksarov, M. S. Sinyukov	Methods of Ensuring the Quality of Assembly of Non-removable Joints from Dissimilar Materials	1191-1199

ISSN: 1728-144X, e-ISSN: 1735-9244

International Journal of Engineering, Volume 39, Number 05, May 2026

S. T. Y. AL-Qaicy	Predicting Optimum Project by Using Structural Equation Modeling	1200-1210
S. Nazari, F. Sogandi	An Improved Control Chart Pattern Classification using a Transfer Learning-Based VGG-16 Network	1211-1225
I. N. Budilov, S. N. Tyulyandin	Ensuring Vibration Strength of Mainline Pumping Unit Systems Using Digital Prototypes	1226-1237
S. A. Olamaei, M. Baghaei Jezehei, A. Broumandnia	Achieving Enhanced Efficiency and Security in IoT: A Hybrid Encryption Approach with ECC and Blowfish	1238-1253
M. S. Shokoohfar, M. Khanmohammadi	Numerical Evaluation of Behavior of Tapered Piled Raft in Loose Sand	1254-1266
V. V. Maksarov, V. D. Nguyen	Machining of Austenite Steels Using the Method of Preliminary Local Cryogenic Treatment	1267-1274
Z. Mortezaie, Z. Amiri, M. Mortezaee	Adaptive Fractional-order Differentiation for Enhanced Image Contrast Utilizing Caputo Masks	1275-1292
R. Azis, M. Arsyad Thaha	Contribution of Perforated Pipe Roughness to Improve Sediment Transport in Hybrid Fluidization Method as an Alternative Dredging Technique	1293-1304

In the Name of God

INTERNATIONAL JOURNAL OF ENGINEERING
Transactions B: Applications

DIRECTOR-IN-CHARGE
H. Omidvar

EDITOR IN CHIEF
G. D. Najafpour

ASSOCIATE EDITOR
A. Haerian

EDITORIAL BOARD

S.B. Adeloju, Charles Sturt University, Wagga, Australia
K. Badie, Iran Telecomm. Research Center, Tehran, Iran
M. Balaban, Massachusetts Ins. of Technology (MIT), USA
M. Bodaghi, Nottingham Trent University, Nottingham, UK
E. Clausen, Univ. of Arkansas, North Carolina, USA
W.R. Daud, University Kebangsaan Malaysia, Selangor, Malaysia
M. Ehsan, Sharif University of Technology, Tehran, Iran
J. Faiz, Univ. of Tehran, Tehran, Iran
H. Farrahi, Sharif University of Technology, Tehran, Iran
K. Firoozbakhsh, Sharif Univ. of Technology, Tehran, Iran
A. Haerian, Sajad Univ., Mashhad, Iran
H. Hassanpour, Shahrood Univ. of Tech., Shahrood, Iran
W. Hogland, Linnaeus Univ, Kalmar Sweden
A.F. Ismail, Univ. Tech. Malaysia, Skudai, Malaysia
M. Jain, University of Nebraska Medical Center, Omaha, USA
M. Keyanpour rad, Materials and Energy Research Center, Karaj, Iran
A. Mahmoudi, Bu-Ali Sina University, Hamedan, Iran
O.P. Malik, University of Calgary, Alberta, Canada
G.D. Najafpour, Babol Noshirvani Univ. of Tech., Babol, Iran
F. Nateghi-A, Int. Ins. Earthquake Eng. Seis., Tehran, Iran
S. E. Oh, Kangwon National University, Korea
M. Osanloo, Amirkabir Univ. of Tech., Tehran, Iran
M. Pazouki, MERC, Karaj, Iran
J. Rashed-Mohassel, Univ. of Tehran, Tehran, Iran
S. K. Sadrnezhaad, Sharif Univ. of Tech, Tehran, Iran
R. Sahraeian, Shahed University, Tehran, Iran
A. Shokuhfar, K. N. Toosi Univ. of Tech., Tehran, Iran
R. Tavakkoli-Moghaddam, Univ. of Tehran, Tehran, Iran
T. Teng, Univ. Sains Malaysia, Gelugor, Malaysia
P. Tiong, Nanyang Technological University, Singapore
X. Wang, Deakin University, Geelong VIC 3217, Australia
H. Omidvar, Amirkabir Univ. of Tech., Tehran, Iran

EDITORIAL ADVISORY BOARD

S. T. Akhavan-Niaki, Sharif Univ. of Tech., Tehran, Iran
M. Amidpour, K. N. Toosi Univ of Tech., Tehran, Iran
M. Azadi, Semnan university, Semnan, Iran
M. Azadi, Semnan University, Semnan, Iran
F. Behnamfar, Isfahan University of Technology, Isfahan
R. Dutta, Sharda University, India
M. Eslami, Amirkabir Univ. of Technology, Tehran, Iran
H. Hamidi, K.N.Toosi Univ. of Technology, Tehran, Iran
S. Jafarmadar, Urmia Univ., Urmia, Iran
A. Kheyroddin, Semnan Univ., Semnan, Iran
N. Latifi, Mississippi State Univ., Mississippi State, USA
H. Oraee, Sharif Univ. of Tech., Tehran, Iran
S. M. Seyed-Hosseini, Iran Univ. of Sc. & Tech., Tehran, Iran
M. T. Shervani-Tabar, Tabriz Univ., Tabriz, Iran
E. Shirani, Isfahan Univ. of Tech., Isfahan, Iran
A. Siadat, Arts et Métiers, France
C. Triki, Hamad Bin Khalifa Univ., Doha, Qatar

TECHNICAL STAFF
M. Khavarpour; M. Mohammadi; V. H. Bazzaz, R. Esfandiar; T. Ebadi

DISCLAIMER

The publication of papers in International Journal of Engineering does not imply that the editorial board, reviewers or publisher accept, approve or endorse the data and conclusions of authors.

International **J**ournal of **E**ngineering *Transactions A: Basics* (ISSN 1728-1431) (EISSN 1735-9244)
International **J**ournal of **E**ngineering *Transactions B: Applications* (ISSN 1728-144X) (EISSN 1735-9244)
International **J**ournal of **E**ngineering *Transactions C: Aspects* (ISSN 2423-7167) (EISSN 1735-9244)
Web Sites: www.ije.ir & www.ijeir.info E-mails: ije.editor8@gmail.com, Tel: (+9821) 88771578, Fax: (+9821) 88773352
Materials and Energy Research Center (MERC)

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1047-1062

International Journal of Engineering

Journal Homepage: www.ije.ir

Physics-Informed Neural Networks Techniques for Analyzing Forced Vibrations of Simply Supported Beams Featuring Variable Cross-Sections

S. M. Sahebdad[a], M. Eftekhari*[a], M. Eftekhari[b]

[a] *Department of Mechanical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran*
[b] *Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran*

PAPER INFO

Paper history:
Received 31 March 2025
Received in revised form 17 May 2025
Accepted 06 June 2025

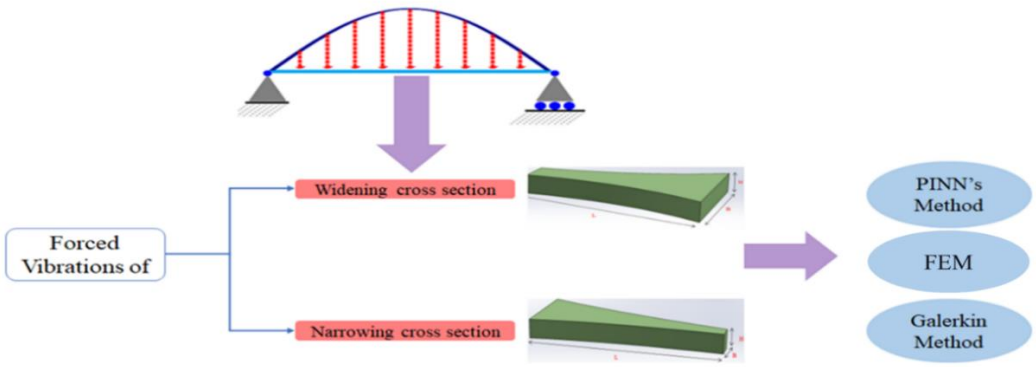
Keywords:
Free and Forced Vibration
Isotropic Beam
Variable Cross Section
Physics-informed Neural Networks
Galerkin Method

ABSTRACT

This research explores the forced vibrations of isotropic beams with variable cross-sections, modeled by the Euler-Bernoulli beam theory. Using Hamilton's principle, the governing partial differential equations are derived, and the complex vibrational behaviors were analyzed. By introducing physics-informed neural networks (PINNs) as an innovative, mesh-free solution technique, the study highlights their ability to provide rapid and precise results by integrating physical laws directly into the machine learning framework. Compared to traditional methods like finite element (FEM) or finite difference schemes, PINNs significantly streamline the computational process by eliminating the need for mesh generation, which simplifies implementation and reduces computational effort, validation with the 6th-order Galerkin method and FEM confirms the high accuracy and efficiency of the proposed approach for analyzing vibrations in beams with varying cross-sections. Overall, this work enhances the application of PINNs in vibration assessment and offers valuable insights for optimizing design and performance across diverse engineering domains, including structural and mechanical systems.

doi: *10.5829/ije.2026.39.05b.01*

Graphical Abstract



1. INTRODUCTION

Isotropic beams stand out for their uniform mechanical properties in all directions, which is why they are widely used in various engineering fields. This uniformity allows these beams to consistently respond to stresses and deformations, making them particularly vital in sectors such as aerospace (1, 2). As highlighted in a

*Corresponding author email: mo.eftekhari@uk.ac.ir (M. Eftekhari)

Please cite this article as: Sahebdad SM, Eftekhari M, Eftekhari M. Physics-Informed Neural Networks Techniques for Analyzing Forced Vibrations of Simply Supported Beams Featuring Variable Cross-Sections. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1047-62.

recent study by Li and Shen (3), they proposed an optimized seismic design method for civil structures combined with robotic construction system technology. Their study emphasized the importance of integrating kinematic analysis of installation robots with internet-based seismic monitoring data to identify hidden safety hazards and implement effective safety management. This approach addresses the challenges of manual installation in high-rise projects and leverages smart robotic equipment to improve construction accuracy, reduce errors, and enhance the seismic resistance of structures. The results include the determination of optimal safety weights and thresholds, demonstrating a significant improvement in the strength and safety of the structures.

The reliability of isotropic beams under various loading conditions underscores their superiority in the design and construction of structures requiring high performance and strength, as exemplified by a very practical and recent study by Makrup et al. (4) conducted a study comparing structural responses derived from probabilistic seismic hazard analysis (PSHA) with those obtained using design spectral ground motion approaches. Their research highlights the importance of integrating these methodologies to achieve a comprehensive understanding of seismic impacts on structures. Results indicate that the structural responses predicted by both methods are quite similar, supporting the reliability of using either approach in seismic design.

Their work emphasized the value of combining probabilistic and deterministic methods to enhance earthquake-resistant design and ensure structural safety in seismic-prone areas., which illustrates the importance of these beams in various applications.

Despite extensive research conducted on the analysis of transverse vibrations of isotropic beams, Canales et al. (5) Moreno-García et al. (6), Azzara, Ghugal et al. (7), Ghugal et al. (8)) there is an urgent need to investigate non-uniform beams that possess variable mass and stiffness distributions. These non-uniform beams can exhibit unique characteristics that significantly differ from their isotropic counterparts. The analysis of vibrations in beams with variable cross-sections not only addresses diverse engineering needs but also contributes to the optimization of design and performance.

This study emphasizes the examination of transverse vibrations in isotropic beams featuring non-uniform cross-sections in this regard, the Euler-Bernoulli beam theory acts as the foundational framework for analysis, simplifying the governing equations of beam motion under various loading conditions. To derive the relevant equations, Hamilton's principle is utilized, which allows for the extraction of a partial differential equations (PDEs) governing the motion of the beam. The application of Hamiltonian mechanics in structural analysis contributes to a more comprehensive understanding of energy distribution throughout the system, thereby simultaneously addressing both kinetic and potential energy. The resulting PDEs encompasses the complexities of the transverse vibration states of the beam while considering variations in the cross-sectional geometry.

In recent years, considerable progress has been achieved in the field of machine learning, particularly concerning neural networks. Physics-informed neural networks (PINNs) idea, initially introduced by Raissi et al. (9) are recognized as a major advancement in numerical solution methods that have created a link between neural networks and engineering physical models. These networks derive the necessary information through differential equations, boundary conditions, and initial conditions governing engineering systems.

The primary incentive for employing PINNs is their capacity to deliver significantly simpler solutions for a variety of problems, requiring very few training data points and operating at a much faster pace compared to traditional numerical approaches such as Finite Element Method (FEM) (10, 11), Finite Difference Method (FDM) (12, 13) Galerkin method discussed by Wang and Wu (14), and other computational techniques (15).

As a mesh-free approach, PINNs can solve different issues several times faster than these conventional methods. In this research, the outcomes derived from PINNs are verified and contrasted with the results obtained via Galerkin method.

Various issues have been studied using PINNs yielding very suitable responses, and different studies in the field of vibrations can be addressed.

Chen et al. (16) explored an enhanced Physics-Informed Neural Network (PINN) approach for structural vibration analysis in "long-duration" simulations. They presented a significant technical hurdle that remains unaddressed within the realm of PINNs. The performance of both the standard PINN (STD-PINN) and conventional time-marching PINN (CT-PINN) techniques in solving vibration equations, especially free-vibration equations, tends to decline over extended simulation periods. This method tackles structural vibration challenges across successive time intervals by implementing four crucial strategies: normalizing the spatiotemporal domain for each interval, employing a reactivating optimization algorithm, utilizing transfer learning, and adopting the sine activation function. The findings reveal that the Advanced time-marching PINN (AT-PINN) method provides accurate solutions while maintaining a reduced computational expense, even throughout prolonged simulations.

Bazmara et al. (17) framework for PINNs was created to explore the nonlinear bending of three-dimensional (3D) functionally graded (FG) beams. The researchers established a mathematical model employing Euler-

Bernoulli beam theory in conjunction with an inhomogeneous beam model, confirming their results against findings from the finite difference method. Additionally, the PINNs technique has demonstrated the ability to forecast the nonlinear bending of the system up to 37 times more swiftly than traditional numerical methods.

In another work, Bazmara et al. (18) PINNs framework has been developed to investigate the nonlinear buckling characteristics of a three-dimensional (3D) functionally graded (FG) porous slender beam supported by a Winkler-Pasternak foundation. This PINN framework utilizes data derived from the governing differential equations of the beam system, along with boundary condition information, to ascertain the critical nonlinear buckling load. The findings from the PINNs framework are corroborated against results obtained from the generalized differential quadrature method (GDQM), illustrating that the proposed PINNs accurately capture the nonlinear buckling behavior of 3D FG porous slender beams. Furthermore, the PINNs approach forecasts nonlinear buckling behavior up to 71 times faster than conventional numerical methods.

Models utilizing PDEs to represent Euler-Bernoulli and Timoshenko beams on a Winkler foundation were executed through a causal physics-informed neural network (PINN) combined with transfer learning, as demonstrated by Kapoor et al. (19). The main contribution of this paper lies in the creation of a causality-respecting PINN loss function specifically designed for structural engineering, combined with transfer learning to improve the adaptability of PINNs in simulating the dynamics of beams on elastic foundations. Numerical tests on Euler-Bernoulli beam demonstrate the efficacy of the proposed approach across various initial conditions, including cases with noise in the initial data. Additionally, the method's capabilities are illustrated for the Timoshenko beam over an expanded spatial and temporal range.

Moreover, Kapoor et al. (20) developed an innovative framework that utilizes PINNs to model complex structural systems consisting of single and double beams, applying Euler-Bernoulli and Timoshenko theories. In this framework, the double beams rest on a Winkler foundation. The results suggest that PINNs offer a promising solution for tackling challenges associated with engineering structures and machinery involving beam systems.

Söyleyici and Ünver (21) investigated an innovative framework based on PINN for solving the theory of free vibrations of Euler-Bernoulli beams and demonstrates its capabilities in identifying damping constants. The results obtained indicate that the use of multi-scale Fourier features and a neural tactile kernel can enhance prediction accuracy in beam vibration analysis, particularly at high frequencies.

Teloli et al. (22) developed an innovative method that utilizes PINNs to tackle inverse problems in structural analysis. This technique is particularly applied to the fourth-order PDEs derived from Euler-Bernoulli theory to evaluate beam deflection and determine critical structural parameters such as damping and elastic modulus. The approach incorporates PDEs into the loss function of the neural network during training, ensuring adherence to physics-based constraints. This method streamlines intricate structural assessments, even in cases where boundary conditions are ambiguous. The proposed technique successfully extracts structural properties using experimental data and confirms its accuracy against sophisticated methodologies.

Despite existing research, Physics-Informed Neural Networks (PINNs) have significant shortcomings. Therefore, this study investigates the forced vibration problem in Euler-Bernoulli beams with variable cross-sections using Physics-Informed Neural Networks. Furthermore, the obtained responses will be validated against results from the Galerkin method and Finite Element Analysis.

In this study, the deep learning library PyTorch version 2.5.0 was used to construct a Physics-Informed Neural Network (23-25). This library is comprehensive in the field and facilitates the development of Physics-Informed Neural Networks.

The main objective of this research is to develop a framework of Physics-Informed Neural Networks (PINNs) for investigating the vibrations of Euler-Bernoulli beams with variable cross-sections. This study notably differs from other research in the field of Physics-Informed Neural Networks dealing with beam vibrations, as it examines beams with variable cross-sections under forced vibration scenarios. Moreover, the responses obtained through the developed Physics-Informed Neural Networks are validated using the Galerkin method and Finite Element Analysis. This approach is very simple, as it does not require discretizing derivatives, approximating a set of basics functions to satisfy boundary conditions, or solving complex frequency equations. It also eliminates the need for refining the number of elements, the complexities of weak form meshing and time-space discretization, and solving systems of equations with various numerical methods; it only requires the differential equation and its boundary and initial conditions to solve the problem with the network. The clear distinction of this study lies in the construction and design of an efficient and fast network for solving forced vibration problems for beams with variable cross-sections. This study can provide a fast, efficient, and less cumbersome process for solving such problems.

Therefore, this work advances the field by combining forced vibration analysis, variable cross-sections, and PINNs in a way that is not fully replicated in the related

literature (15-23).

More details about this research are explained in the following sections: Section 2 is devoted to the mathematical formulation. The principle of Physics-Informed Neural Networks (PINNs) is introduced in section 3. Section 4 pertains to the Finite Element Method theory. Results and analyses are described in section 5. Finally, section 6 concludes the paper.

2. MATHEMATICAL MODELING

This study investigates a structural element consisting of an isotropic beam with a specified length L, oriented along the x-axis. The beam has a variable cross-section with a width B along the y-axis and a height H along the z-axis, as illustrated in Figure 1. Figure 1.c, a simply supported beam subjected to a prescribed load is depicted. additionally, Figures 1.a and 1.b respectively show an isotropic wide-section beam and an isotropic narrow-section beam.

According to the Euler-Bernoulli beam theory, the relationship between axial strain displacement can be expressed as follows:

$$\varepsilon_{xx} = -z\frac{\partial^2 w}{\partial x^2} + \frac{\partial u}{\partial x} + \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2 \quad (1)$$

In the context of an isotropic beam, where u and w represent the axial and transverse displacements, respectively, the normal stress can be derived using the principles of linear elastic constitutive law:

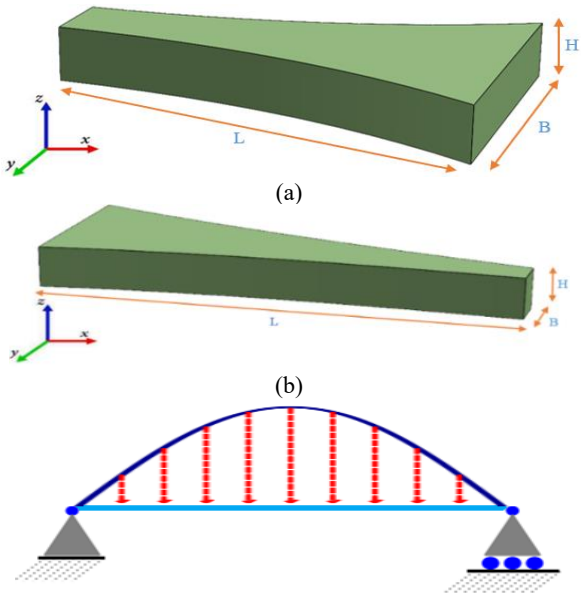


Figure 1. Schematic of a simply supported beam with variable cross sections a) widening b) narrowing

$$\sigma_{xx} = E\left(-z\frac{\partial^2 w}{\partial x^2} + \frac{\partial u}{\partial x} + \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2\right) \quad (2)$$

where E represents Young's modulus. Variation of potential energy is explained as:

$$\delta U = \int_V \sigma_{xx}\delta\varepsilon_{xx}\, dV = \int_0^L \int_A \sigma_{xx}\, \delta\left(-z\frac{\partial^2 w}{\partial x^2} + \frac{\partial u}{\partial x} + \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2\right) dA dx = \int_0^L N_{xx}\delta\left(\frac{\partial u}{\partial x} + \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2\right) - M_{xx}\delta\frac{\partial^2 w}{\partial x^2}\, dx \quad (3)$$

In which, A refers to the beam's cross-sectional area, while N_{xx} and M_{xx} denote the axial load and moment of bending per unit length, respectively:

$$N_{xx} = \int_A \sigma_{xx}\, dA \quad (4)$$

$$M_{xx} = \int_A \sigma_{xx}\, z\, dA \quad (5)$$

The variational representation of the external force (δw_{nc}) is expressed as follows:

$$\delta w_{nc} = \int_0^L P sin\left(\frac{\pi x}{L}\right)\cos(\Omega t)\, \delta w dx \quad (6)$$

In Equation 6, P, L, and Ω represent the load, the length of the beam, and the frequency, respectively. variation of kinetic energy (δT):

$$\delta T = \int_0^L \frac{1}{2}\rho A(x)\left(\left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial w}{\partial t}\right)^2\right) = \int_0^L \rho A(x)\frac{\partial u}{\partial t}\delta\frac{\partial u}{\partial t}dx + \int_0^L \rho A(x)\frac{\partial w}{\partial t}\delta\frac{\partial w}{\partial t}dx \quad (7)$$

Hamilton's principle was employed to derive the governing equations for beams with non-uniform cross-sections., which are expressed as follows:

$$\int_{t_1}^{t_2}(\delta T - \delta U + \delta W_{nc})\, dt = 0 \quad (8)$$

By substituting Equations 3, 6, and 7 into Equation 8 and employing integration by parts with respect to the variables, the following results are obtained, leading to two equations for lateral and axial displacements:

$$\delta u: \frac{\partial N_{xx}}{\partial x} - \rho A(x)\frac{\partial^2 u}{\partial t^2} = 0 \quad (9)$$

$$\delta w: N_{xx}\frac{\partial^2 w}{\partial x^2} - \frac{\partial N_{xx}}{\partial x}\frac{\partial w}{\partial x} + \frac{\partial^2 M_{xx}}{\partial x^2} + \rho A(x)\frac{\partial^2 w}{\partial t^2} + P sin\left(\frac{\pi x}{L}\right)cos(\Omega t) = 0 \quad (10)$$

From Equation 9, conclude $\frac{\partial N_{xx}}{\partial x} = \rho A(x)\frac{\partial^2 u}{\partial t^2}$. By assuming the axial vibration is much smaller than the lateral vibration, the time-dependent longitudinal component of Equation 9 can be disregarded:

$$\frac{\partial N_{xx}}{\partial x} = 0, \quad (11)$$

The displacement $u(x,t)$ is determined as follows:

$$\frac{\partial}{\partial x}\left[\int_A \sigma_{xx}\ dA\right]=0\rightarrow\frac{\partial}{\partial x}\left[\int_A\left(E\ \left(-z\frac{\partial^2 w}{\partial x^2}+\ \frac{\partial u}{\partial x}+\ \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2\right)\right)dA\right]=0\rightarrow\int_A z\ dA=0,\frac{\partial}{\partial x}\left[EA\left(\frac{\partial u}{\partial x}+\ \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2\right)\right]=0,\rightarrow,$$

$$\frac{\partial}{\partial x}\left[EA\left(\frac{\partial u}{\partial x}+\frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2\right)\right]=0,\rightarrow,EA\left(\frac{\partial u}{\partial x}+\ \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2\right)=c_1$$

$$EA\left(u+\int_0^x\frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2dx\right)=c_1x+c_2,\quad(12)$$

$$u(0,t)=0\rightarrow c_2=0\ ,u(L,t)=0,\rightarrow$$

$$c_1=\frac{EA}{L}\left(\int_0^L\frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2dx\right),\rightarrow u(x,t)=\frac{c_1}{EA}x-\int_0^x\frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2dx,\rightarrow N_{xx}=\frac{EA}{L}\left(\int_0^L\frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2dx\right),$$

$$u(x,t)=\frac{c_1}{EA}x-\int_0^x\frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2dx$$

$$N_{xx}=\frac{EA}{L}\left(\int_0^L\frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2dx\right)$$

By substituting Equation 11 in Equation 10 and neglecting the nonlinear cubic terms:

$$\frac{\partial^2 M_{xx}}{\partial x^2}+\rho A(x)\frac{\partial^2 w}{\partial t^2}-Psin\left(\frac{\pi x}{L}\right)cos(\Omega t)=0\quad(13)$$

$\frac{\partial^2 M_{xx}}{\partial x^2}$ is presented for a variable cross section as follows:

$$M_{xx}=\ EI(x)\frac{\partial^2 w}{\partial x^2},\rightarrow\frac{\partial^2 M_{xx}}{\partial x^2}=EI''(x)\frac{\partial^2 w}{\partial x^2}+2EI'(x)\frac{\partial^3 w}{\partial x^3}+\ EI(x)\frac{\partial^4 w}{\partial x^4}\quad(14)$$

By substituting Equation 14 into Equation 13 one obtains:

$$EI(x)\frac{\partial^4 w}{\partial x^4}+\ 2EI'(x)\frac{\partial^3 w}{\partial x^3}+\ EI''(x)\frac{\partial^2 w}{\partial x^2}+\rho A(x)\frac{\partial^2 w}{\partial t^2}=\ Psin\left(\frac{\pi x}{L}\right)cos(\Omega t)\quad(15)$$

Equation 16 is utilized to define dimensionless variables, allowing the governing equation to be represented in a dimensionless form as follows:

$$\hat{t}=\frac{1}{L^2}\sqrt{\frac{EI_0}{\rho A_0}},\hat{x}=\frac{x}{L},\widehat{w}=\frac{w}{L},\hat{I}=\frac{I}{I_0},\hat{P}=\frac{PL^3}{\hat{A}EI_0},\quad\widehat{\Omega}=\Omega L^2\sqrt{\frac{\rho A_0}{EI_0}},\hat{A}=\frac{A}{A_0}\quad(16)$$

y substituting dimensionless variables in Equation 16 into Equation 15 one obtains:

$$\frac{\hat{I}(\hat{x})}{\hat{A}(\hat{x})}\frac{\partial^4\widehat{w}}{\partial\hat{x}^4}+\ 2L\frac{\hat{I}'(\hat{x})}{\hat{A}(\hat{x})}\frac{\partial^3\widehat{w}}{\partial\hat{x}^3}+\ L^2\frac{\hat{I}''(\hat{x})}{\hat{A}(\hat{x})}\frac{\partial^2\widehat{w}}{\partial\hat{x}^2}+\frac{\partial^2\widehat{w}}{\partial\hat{t}^2}=\hat{P}\sin(\pi\hat{x})\,cos\left(\widehat{\Omega}\hat{t}\right)\quad(17)$$

In this study, the boundary conditions are defined as simply supported on both ends, as indicated in Equation 18.

$$\widehat{w}\ (0,\hat{t})=0\qquad\widehat{w}\ (1,\hat{t})=0$$

$$\frac{\partial^2\widehat{w}(0,\hat{t})}{\partial\hat{x}^2}=0\qquad\frac{\partial^2\widehat{w}(1,\hat{t})}{\partial\hat{x}^2}=0\quad(18)$$

3. PRINCIPLE OF PHYSICS-INFORMED NEURAL NETWORKS (PINNs)

Physics-informed neural networks (PINNs) signify a notable progression in the convergence of machine learning and scientific computing. In this essay, essential concepts that form the basis of PINNs will be outlined, focusing on three primary aspects: 3.1 Fundamentals of PINNs, 3.2 Automatic Differentiation in PINNs, and 3.3 Adaptive Activation functions**.** By integrating incorporate fundamental physical principles directly into the training process, PINNs enhance model accuracy and generalization. This approach allows for more reliable predictions in complex systems where traditional machine learning methods may struggle. Figure 2 shows the Framework of a PINN approach.

3. 1. Fundamental of PINNs Examine a typical PDE system parameterized by using Λ, the solution can be found $j(x,t)$ defined across a domain $\Gamma\subset\mathbb{R}$ with $x=(x_1,\cdots,x_e)$ and t representing time (9, 26). This system can be expressed in partial differential equations that characterize the behavior of the solution $j(x,t)$ across the spatial domain and over time, allowing for various boundary and initial conditions to be applied.

$$F(x,t;\frac{\partial j}{\partial x_1},\dots\frac{\partial j}{\partial x_e},\frac{\partial j}{\partial x_1};\frac{\partial j}{\partial t},\frac{\partial^2 j}{\partial t^2};\frac{\partial j}{\partial x_1\partial x_1},\dots,\frac{\partial j}{\partial x_1\partial x_e},\dots\frac{\partial j}{\partial x_1\partial x_e};\frac{\partial j}{\partial t\,\partial x_1},\dots,\frac{\partial j}{\partial t\,\partial x_e};\dots,\Lambda)=0$$

$$j(x,t)=\alpha_0(x,t)\quad(19)$$

$$j\ (x,t)=\ \alpha_\Gamma(0,t)=0$$

$$x\in\ \Gamma,x\in\ \partial\Gamma$$

Let Γ represent the spatiotemporal domain, while $\partial\Gamma$ indicates the boundary. The spatial coordinate is denoted by $x\ \in\Gamma$, and F signifies the residual of the partial differential equation (PDE), which includes the differential operators. $[\frac{\partial j}{\partial x_1},\dots,\frac{\partial j}{\partial x_e},\dots,\frac{\partial j}{\partial t},\frac{\partial^2 j}{\partial t^2}]$; $\Lambda=[\Lambda_1,\Lambda_2,\cdots]$. The parameters of the partial differential equation (PDE) include $j(x,t)$, which denotes the solution of the PDE given certain boundary conditions $\alpha_0(x,t)$. These boundary conditions can be of different types, including Robin, Neumann, and Dirichlet, in addition to the initial conditions $\alpha_\Gamma(0,t)$.

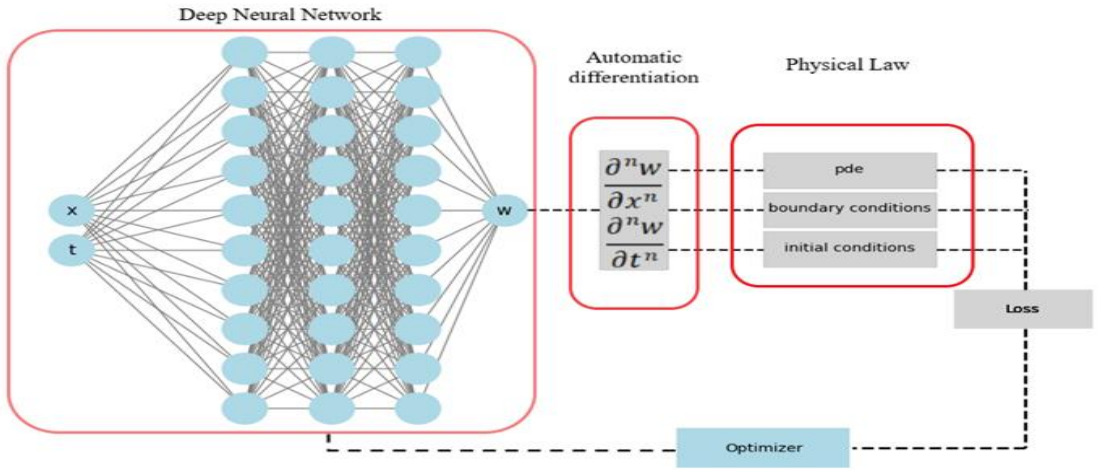


Figure 2. Framework of a PINN approach

In the realm of PINNs, a feed-forward neural network with multiple hidden layers is designed to estimate the solution of a partial differential equation (PDE) over time. It does this by using spatial coordinates as inputs and producing an output vector that matches the dimensionality of the problem, $j(x,t)$, as Shown in Figure 2. In Equation 20, the term h^0 within the equation incorporates the variables x and t, reflecting the input variable from the initial layer. This input is then transmitted to the subsequent layer as Z^1, t in the neural network, represented by Equations 20 and 21 below:

$$Z^1 = w^1 h^0 + b^1,\ O^1 = \psi(Z^1) \quad \text{For } k=1$$
$$Z^k = w^k O^{k-1} + b^k \quad \text{For } 2 \le k \le L\text{-}1 \tag{20}$$

$$O^k = \psi(Z^k) \quad \text{For } 2 \le k \le L\text{-}1$$
$$O^L = w^L O^{L-1} + b^L \quad \text{For } k = L \tag{21}$$

Where L represents a neural network with L-layer, and L refers to the total count of hidden layers; w^k and α^k signify the weights and bias terms for the k^{th} layer; ψ denotes a nonlinear activation function. Z^k and O^k depict the input and output of the activation function ψ for the k^{th} layer respectively. From the sequence of Relations 20 and 21, it can be observed that O^k is used as the input for the k+1 layer.

PINNs address the specified system of PDE system outlined in Equation 21 by refining the problem through the iterative modification of the trainable model parameters, which encompass w^k and α^k, with the goal of minimizing the loss function $J(\theta)$ as specified in Equation 22.

$$J(\theta) = \lambda_{pde} L_{pde}(\theta) + \lambda_{ic} L_{ic}(\theta) + \lambda_{bc} L_{bc}(\theta) \tag{22}$$

where

$$L_{pde} = \frac{1}{N_{pde}} \sum_{i=1}^{N_{pde}} \left(\mathcal{F}\left(\hat{x}^i{}_{pde}, \hat{t}^i{}_{pde} \right) \right)^2 \tag{23}$$

$$L_{ic} = \frac{1}{N_{ic}} \sum_{i=1}^{N_{ic}} \left(\hat{w}\left(\hat{x}^i{}_{ic}, 0 \right) - K^i{}_0\left(\hat{x}^i{}_{ic} \right) \right)^2 + \left(\frac{\partial \hat{w}(\hat{x}^i{}_{ic}, 0)}{\partial t} - K^i{}_1\left(\hat{x}^i{}_{ic} \right) \right)^2 \tag{24}$$

$$L_{bc} = \frac{1}{N_{bc}} \sum_{i=1}^{N_{bc}} \left(\hat{w}\left(\hat{x}^i{}_{bc}, \hat{t}^i{}_{bc} \right) - \hat{w}^i \right)^2 + + \left(\frac{\partial^n \hat{w}(\hat{x}^i{}_{bc}, \hat{t}^i{}_{bc})}{\partial x^n} - \frac{\partial^n \hat{w}^i}{\partial x^n} \right)^2 \tag{25}$$

In this context, λ_{pde}, λ_{ic} and λ_{bc} represent the weights. L_{pde} evaluates the residuals of the governing equations. $\Theta = \{w^k, b^k\}_{k=1}^{L}$ stands for trainable parameters. Moreover, to comply with L_{ic} and L_{bc} are the initial conditions for the predictions of PINNs align with the data and boundary conditions, initial conditions are incorporated. Automatic differentiation (AD) is employed to compute the derivatives of the neural network and its components through the application of the chain rule Equation 18. Specifically, the loss function in Equation 22 is characterized as sampling the Mean Square Error (MSE) in Equation 23, where the term $F(x,t)$ pertains to the PDE equation, and N_{pde} signifies the number of training points in the PDE domain. In Equation 24, $k_0(x)$ indicates the actual initial value, while $k_1(x)$ represents the actual initial velocity, with N_{ic} denoting the number of training points for the initial conditions. Additionally, Equation 25 generally illustrates various types of boundary conditions, which may exhibit different states based on the problem's requirements, with N_{bc} indicating the number of training points for the boundary conditions. Typically, in the MSE

approach, the predicted value of the grid is subtracted from the actual value and raised to the power of two.

Two different optimizers were used independently to fine-tune the weights and biases of each layer, resulting in varying outcomes, with one demonstrating superior performance. The first optimizer, which is more commonly used, and the second optimizer, which we used to solve the problem, are compared in performance in Figures 4 and 5. Figure.3 related to the optimizer ADAM Mehmood et al. (27) and Figure 4 to the optimizer L-BFGS Tong et al. (28) which has shown significantly better performance.

Establishing an appropriate loss function is paramount in the development of neural network algorithms, as it encapsulates the fundamental physical laws governing the problem at hand. This study formulates the loss function by integrating PDEs, initial conditions, and boundary conditions. Furthermore, it explores the interplay between these elements and their impact on the model's performance. Ultimately, this framework aspires to enhance the precision and dependability of the solutions generated by the neural network.

3. 2. Automatic Differentiation (AD) in PINNs

Automatic Differentiation (AD) is pivotal for computing partial derivatives, a fundamental step in formulating the loss function outlined in Equation 21. By leveraging the chain rule integrated into machine learning frameworks like PyTorch, AD facilitates the direct computation of derivatives of outputs concerning the inputs to the network within the computational graph. In contrast to traditional numerical methods, AD eliminates the necessity of calculating discretization and truncation errors by deriving derivatives through explicit expressions. It is important to recognize that neural networks are utilized to parameterize the partial differential equations (PDEs) and their derivatives, leading to what are termed generalization and optimization errors, influenced by the training dataset and the optimizer, respectively. The schematic framework of PINNs and their critical components employed in this study are depicted. A physical problem is examined, as illustrated in various sections of Figure 5.

3. 3. Adaptive Activation Function (AAF)

As detailed in sections (3.1) and (3.2), the derivative of the loss function depends on the optimization parameters, highlighting the importance of the activation function. Different activation functions, such as sine and hyperbolic tangent (tanh), and others, are utilized within the PINNs framework, tailored to address specific challenges. Although the training initially employed a standard tanh function, it encountered convergence issues with the loss function. To address this challenge, this study introduces an adaptive activation function, incorporating a hyperparameter β into the activation mechanism, as delineated in Equation 26. This approach effectively mitigates the convergence difficulties experienced by Jagtap et al. (29, 30).

This approach enables greater flexibility and adaptability during training. By incorporating the hyperparameter, more stable convergence behavior across various problem types is anticipated. The findings indicate a significant enhancement in the efficiency and effectiveness of the model, ultimately leading to better performance in solving the target equations.

$$O^k = \psi(\beta Z^k) = \frac{e^{\beta Z^k} - e^{-\beta Z^k}}{e^{\beta Z^k} + e^{-\beta Z^k}} \tag{26}$$

In this approach, the parameter β is determined through the simultaneous reduction of the loss function along with the weights and biases. This methodology substantially improves the accuracy of the model. and therefore, it leads to the optimization of $\theta = \{w^k, b^k, \beta\}_{k=1}^{L}$ as:

$$\theta^* = agrmin_{\beta \in \Lambda \geq 0}(J(\theta)) \tag{27}$$

Additionally, the hyper-parameter β is adjusted along with weights and biases as follows:

$$\theta^{i+1} = \theta^i - \alpha \nabla_\theta J^i(\theta) \tag{28}$$

where $0 < \alpha \leq 1$ is the learning rate to regulate the impact of the learning factor on the activation Function 25, which is responsible for to achieve convergence to minima, one should scale the hyper-parameter β through a multiplier. Additionally, this adjustment can enhance the model's performance during training.$\eta > 1$, which produces a new activation function.

$$O^k = \psi(\eta \beta Z^k) \tag{29}$$

Comprehensive and detailed information regarding the impact of this parameter η can be found in the literature [29,30].

4. FINITE ELEMENT METHOD

The Finite Element Method (FEM) is a powerful numerical technique for solving partial differential equations (PDEs) governing dynamic systems, such as forced vibrations of beams in structural engineering. This section presents a comprehensive mathematical framework for analyzing forced vibrations of an Euler-Bernoulli beam with variable cross-sectional area $A(x)$, variable moment of inertia $I(x)$, and constant damping, using FEM. The formulation includes the derivation of the weak form, spatial discretization, assembly of system matrices, and numerical solution techniques. Detailed explanations are provided for element discretization, degrees of freedom (DOFs), and shape functions, ensuring clarity for applying FEM to a general forced

vibration problem. the sections include 4.1 Governing Equation, 4.2 Weak Form Derivation, 4.3 Element Discretization and Degrees of Freedom, 4.4 Shape Functions, 4.5 Element Matrices and Force Vector, 4.6 Global Assembly.

4. 1. Governing Equation The governing PDE for an Euler-Bernoulli beam undergoing forced vibrations in a one-dimensional domain $x \in [0, L]$ over time $t \in [0, T]$ is:

$$\frac{\partial^2}{\partial x^2}\left(EI(x)\frac{\partial^2 w(x,t)}{\partial x^2}\right) + \rho A(x)\frac{\partial^2 w(x,t)}{\partial t^2} + c\frac{\partial w(x,t)}{\partial t} = f(x,t) \quad (30)$$

where: $w(x,t)$ transverse displacement of the beam, E Young's modulus, I(x)moment of inertia, varying along the beam length, A(x) cross-sectional area, varying along the beam length, ρ mass density (constant), c constant damping coefficient per unit length, $f(x,t)$ external forcing function, $x \in [0, L]$ spatial coordinate, $t \in [0,T]$ time. The first term represents bending stiffness, the second term is the inertial force, and the third term accounts for viscous damping (constant). The forcing function f (x,t) drives the forced vibration response. Boundary Conditions: For generality, consider the supported boundary conditions that are in Equation 18, although the framework can be adapted for other conditions, Initial Conditions

$$w(x,0) = w_0(x)\ , \frac{\partial w(x,0)}{\partial t} = v_0(x) \quad (31)$$

where $w_0(x)$ and $v_0(x)$ are the initial displacement and velocity distributions, respectively.

4. 2. Weak Form Derivation To apply FEM, the strong form of the PDE is converted into a weak form by multiplying by a test function $v(x) \in H^2(x)$, satisfying the essential boundary conditions ($v(0) = v(L) = 0$), and integrating over the domain. The weak form is:

$$\int_0^L \left[\frac{\partial^2}{\partial x^2}\left(EI(x)\frac{\partial^2 w(x,t)}{\partial x^2}\right) + \rho A(x)\frac{\partial^2 w(x,t)}{\partial t^2} + c\frac{\partial w(x,t)}{\partial t} - f(x,t)\right] v(x)dx = 0 \quad (32)$$

this must hold for all test functions $v(x) \in H^2(x)$ satisfying the boundary conditions.

4. 3. Element Discretization and Degrees of Freedom The domain $[0,L]$ is discretized into n finite elements of equal length $h = \frac{L}{n}$, forming a mesh with nodes at $x_i = ih$ (for i=0,1,…,n). Each element spans $[x_e, x_e+h]$, where $x_e = eh$ for element e = 0,1,…,n−1.

The number of elements n is chosen to balance accuracy and computational efficiency; a larger n (smaller h) captures variations in $A(x)$ and $I(x)$ more accurately but increases computational cost. Each element has two nodes (one at each end), and each node has two degrees of freedom, displacement (w_i): the transverse displacement at node i, rotation ($\theta_i = \frac{\partial w_i}{\partial x}$): the slope of the displacement at node i.

Thus, each element has a total of four DOFs, represented by the element DOF vector:

$$q^e = [w_1, \theta_1, w_2, \theta_2]^T \quad (33)$$

where w_1, θ_1 are the displacement and rotation at the left node ($x = x_e$), and w_2, θ_2 are at the right node ($x = x_e + h$). For a mesh with n elements, there are n+1 nodes, resulting in 2(n+1) DOFs before applying boundary conditions.

4. 4. Shape Functions The displacement within an element is approximated using Hermitian shape functions, which ensure C^1continuity (continuity of displacement and its first derivative) across element boundaries, necessary for the fourth-order PDE. The displacement is expressed as follows:

$$w(x,t) \approx w_h(x,t) = N(x)\, q^e(t) \quad (34)$$

where $N(x) = [N_1(x), N_2(x), N_3(x), N_4(x)]$is the vector of shape functions. In local coordinates $\xi = \frac{x - x_e}{h}$ ($0 \le \xi \le 1$):

$$N_1(\xi) = 1 - 3\xi^2 + 2\xi^3 \qquad N_2(\xi) = h(\xi - 2\xi^2 + \xi^3)$$
$$N_3(\xi) = 3\xi^2 - 2\xi^3 \qquad N_4(\xi) = h(-\xi^2 + \xi^3) \quad (35)$$

These functions satisfy:

$$N_1(0) = 1\ , N_1(1) = 0, \frac{dN_1(0)}{d\xi} = 0, \frac{dN_1(1)}{d\xi} = 0$$

displacement at ξ= 0

$$N_2(0) = 0, N_2(1) = 0, \frac{dN_2(0)}{d\xi} = h, \frac{dN_2(1)}{d\xi} = 0$$

rotation at ξ= 0

$$N_3(0) = 0\ , N_3(1) = 1, \frac{dN_3(0)}{d\xi} = 0, \frac{dN_3(1)}{d\xi} = 0$$

displacement at ξ=1

$$N_4(0) = 0\ , N_4(0) = 0, \frac{dN_4(0)}{d\xi} = 0, \frac{dN_4(1)}{d\xi} = h$$

rotation at ξ= 1

The test function is similarly approximated:

$$v(x) = N(x)v^e \quad (36)$$

The Hermitian shape functions are cubic polynomials, ensuring that both w_h and $\frac{\partial w_h}{\partial x}$are continuous across elements, which is critical for modeling the beam's bending behavior.

4. 5. Element Matrices and Force Vector Substituting $w_h = N\, q^e$ and $v_h = N\, v^e$ into the weak form for an element spanning $[x_e, x_e+h]$, the element-level equations are derived. The cross-sectional area

$A(x)$ and moment of inertia $I(x)$ re-assumed to vary linearly within each element:

$$A(x) = A_1 + (A_2 - A_1)\xi$$
$$I(x) = I_1 + (I_2 - I_1)\xi \quad (37)$$

where $A_1 = A(x_e)$,$A_2 = A(x_e h)$,$I_1\ I(x_e)$ and $I_2 = I(x_e + h)$.Integrals are expressed in local coordinates ξ with $dx = hd\xi$.

The Stiffness Matrix:

$$\int_{x_e}^{x_e+h} EI(x)\frac{\partial^2 w_h}{\partial x^2}\frac{\partial^2 v_h}{\partial x^2}dx = v^{eT}\left(\int_{x_e}^{x_e+h} EI(x)\frac{d^2N^T}{dx^2}\frac{d^2N}{dx^2}dx\right)q^e \quad (38)$$

$$K^e = \int_{x_e}^{x_e+h} EI(x)\frac{d^2N^T}{dx^2}\frac{d^2N}{dx^2}dx$$

Since $\frac{d\xi}{dx} = \frac{1}{h}$, $\frac{d^2N}{dx^2} = \frac{1}{h^2}\frac{d^2N}{d\xi^2}$

$$\frac{d^2N_1}{d\xi^2} = -6 + 12\xi \qquad \frac{d^2N_2}{d\xi^2} = h(-4 + 6\xi)$$
$$\frac{d^2N_3}{d\xi^2} = 6 - 12\xi \qquad \frac{d^2N_4}{d\xi^2} = h(-4 + 6\xi) \quad (39)$$

$$K^e = \int_0^L E\,[I_1 + (I_2 - I_1)\xi]\frac{1}{h^3}$$
$$\begin{bmatrix}(-6+12\xi)^2 & h(-6+12\xi)(-4+6\xi) & \cdots \\ \vdots & h^2(-4+6\xi)^2 & \cdots \\ \vdots & \vdots & \ddots\end{bmatrix}d\xi \quad (40)$$

Due to the linear variation of $I(x)$, the integral is typically evaluated numerically using quadrature methods Gauss-Legendre with 3 points to compute the matrix entries accurately.

The damping Matrix (C^e) is obtained as: (c is constant)

$$\int_{x_e}^{x_e+h} c\frac{\partial w_h}{\partial t}v_h dx = v^{eT}\left(\int_{x_e}^{x_e+h} cN^T N dx\right)\dot{q}^e$$
$$C^e = c\int_{x_e}^{x_e+h} N^T N\,dx = ch\int_0^L N^T N\,d\xi \quad (41)$$

$$C^e = \frac{ch}{420}\begin{bmatrix}156 & 22h & 54 & -13h \\ 22h & 4h^2 & 13h & -3h^2 \\ 54 & 13h & 156 & -22h \\ -13h & -3h^2 & -22h & 4h^2\end{bmatrix} \quad (42)$$

The mass Matrix (M^e) is obtained as:

$$\int_{x_e}^{x_e+h}\rho A(x)\frac{\partial^2 w_h}{\partial t^2}v_h dx = v^{eT}\left(\int_{x_e}^{x_e+h}\rho A(x)N^T N dx\right)\ddot{q}^e$$
$$M^e = \int_{x_e}^{x_e+h}\rho A(x)N^T N dx = \int_0^L \rho[A_1 + (A_2 - A_1)\xi]N^T N\,hd\xi \quad (43)$$

$$M^e = \rho h\int_0^L [A_1 + (A_2 - A_1)\xi] \quad (44)$$

$$\begin{bmatrix}N_1N_1 & N_1N_2 & N_1N_3 & N_1N_4 \\ N_2N_1 & N_2N_2 & N_2N_3 & N_2N_4 \\ N_3N_1 & N_3N_2 & N_3N_3 & N_3N_4 \\ N_4N_1 & N_4N_2 & N_4N_3 & N_4N_4\end{bmatrix}d\xi$$

With linear variation in $A(x)$, this integral is also evaluated numerically using quadrature methods to account for the ξ-dependent term.

The Force Vector ($F^e(t)$) is obtained as:

$$F^e(t) = \int_{x_e}^{x_e+h} f(x,t)N^T dx = \int_{-1}^{1} f(x_e + h\gamma, t)N^T h d\gamma \quad (45)$$

Using three-point Gauss-Legendre quadrature:

$$\gamma_i = \left[-\sqrt{\frac{3}{5}}, 0, \sqrt{\frac{3}{5}}\right] \qquad \omega_i = \left[\frac{5}{9}, \frac{8}{9}, \frac{5}{9}\right] \quad (46)$$

$$F^e(t) = \sum_{i=1}^{3}\omega_i\ f(x_e + h\gamma_i, t)N^T(\gamma_i)h \quad (47)$$

4. 6. Global Assembly The global stiffness (K), damping (C), mass (M), and force ($F(t)$) matrices are assembled by summing contributions from all elements. For n elements, there are $n + 1$ nodes and $2(n + 1)$ DOFs before boundary conditions. The DOFs for element eee are mapped as:

$$\text{Dofs} = [2e\ , 2e + 1\ , 2(e + 1), 2(e + 1) + 1] \quad (48)$$

Simply supported boundary conditions $w(0, t) = w(L, t) = 0$ are enforced by eliminating displacement DOFs at $x = 0$ $(index\ 0)$ and $x = L$ $(index\ 2n + 1)$, reducing the system to$K_{red}, C_{red}, M_{red}$ and F_{red} .

The Global finite element equation" The global finite element equation for the reduced system is:

$$M_{red}\ \ddot{q}_{red}(t) + C_{red}\dot{q}_{red}(t) + K_{red}q_{red}(t) = F_{red}(t) \quad (49)$$

In this study, Equation 52 given next section is solved using fourth-order Runge-Kutta method.

5. RESULTS AND DISCUSSIONS

In this section, the forced vibrations of a beam with variable cross-sections are investigated using Physics-Informed Neural Networks (PINNs), and the results are validated using the Galerkin and finite element methods. This study extracts results related to the forced vibrations of a beam with variable cross-section from subsection 5.1 and employs the PINNs method. Furthermore, the findings are compared with classical methods to ensure their accuracy and reliability. To optimize the accuracy and convergence of the numerical solution, hyperparameters were selected through trial and error, and the network was trained with different architectures to select the most optimal network for this problem. Regarding the selection of the network architecture,

various architectures were used, but a satisfactory response was not obtained. Therefore, to increase the convergence speed, an adaptive activation function was used to increase the convergence speed by creating adaptive changes in the slope of the activation function. However, to understand precisely the hyperparameter choices offered, the following problem is selected, where $\hat{C}$ is a dimensionless damping coefficient.

$$\frac{\partial^4 \hat{w}}{\partial x^4} + \frac{\partial^2 \hat{w}}{\partial t^2} + \hat{c}\frac{\partial \hat{w}}{\partial t} = \hat{P}\sin(\pi \hat{x})\, cos(\hat{\Omega}\hat{t}) \tag{50}$$

$$\hat{c} = \frac{c\, l}{EI_0\sqrt{\frac{EI_0}{\rho A_0}}} \tag{51}$$

This problem is subsequently investigated using Physics-Informed Neural Networks under specified boundary conditions Equation 18 and the initial conditions laid out in Equation 52.

$$\hat{w}(\hat{x}, 0) = 0$$
$$\frac{\partial \hat{w}(\hat{x}, 0)}{\partial t} = 0 \tag{52}$$

Initially, the problem was examined using a conventional activation function in conjunction with two optimization algorithms, ADAM and LBFGS. It is worth noting that a random dataset was used in this study to solve all problems. This dataset consists of 1600 points., 1600 points have been included, consisting of N_{pde}= 1000 number of training points in the PDE domain, N_{ic} = 200 number of training points initial conditions and N_{bc}= 400 number of training points boundary conditions, as shown in Figure 3.

However, the results were unsatisfactory. Subsequently, an approach utilizing an adaptive activation function was adopted, resulting in a significant improvement in outcomes. The network was then trained with various architectures, as shown in Table 1. Figure 4 illustrates the impact of the adaptive activation function on the network's performance, which is revealing that LBFGS significantly outperforms ADAM, regardless of whether adaptive or non-adaptive activation functions are employed. The introduction of the adaptive activation function notably enhanced the effectiveness of the LBFGS optimizer, leading to its continued use in this study. The learning rate for the ADAM optimizer is set at to 1e-4, while for the LBFGS optimizer, it is set at to 0.1. The training parameters include a maximum of 50,000 evaluations and a maximum of 1 iteration. It is important to note that the final loss values reported in Table 1 are influenced by the use of the adaptive activation function. Figures 5 and 6 display the loss evolution during training with the adaptive activation function for both optimizers: ADAM exhibits significant fluctuations, whereas LBFGS demonstrates minimal fluctuations, converging more rapidly and smoothly.

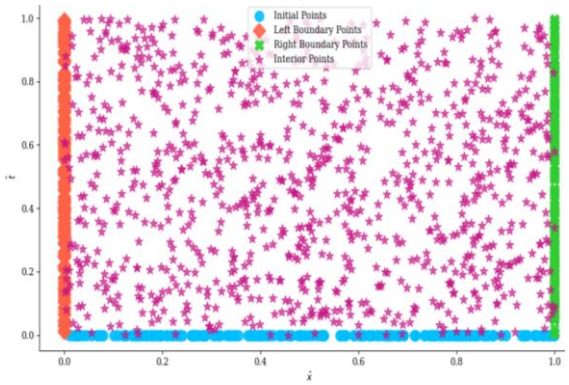


Figure 3. Data for learning beam parameters

TABLE 1. PINN Hyperparameters Across Different Cases

Architecture	Hidden layers N_h	Neuron layers N_n	Scale factor η	Final Loss Value
Model A	4	60	10	4.00e-6
Model B	8	80	10	3.65e-6
Model C	4	100	10	3.25e-6
Model D	3	100	10	3.18e-6

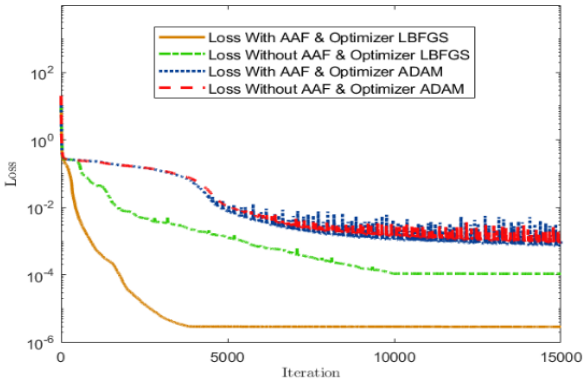


Figure 4. Impact of Activation Function Adaptivity on Optimization Efficiency with Adam and LBFGS

Based on all the results obtained throughout the training process of this study, the LBFGS optimizer and model architecture D will be utilized to analyze vibrations.

Now that the hyperparameters and optimizer of the designed network architecture have been set, we will examine the response obtained from solving problem in Equation 50. This solution has been validated with another issue resolved using Physics-Informed Neural Networks (PINNs) Chen et al. (16). As shown in Figure 7, the obtained solutions are in complete agreement with each other, and there are no significant errors. Additionally, Figure 7 includes higher-dimensional parameters.

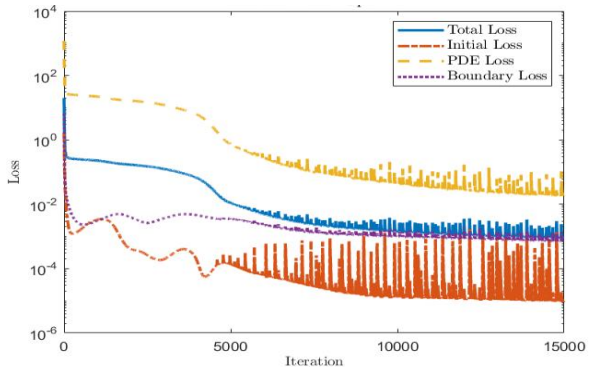


Figure 5. Training Loss Components: PDE, Initial Condition, and Boundary Condition with Adam Optimizer

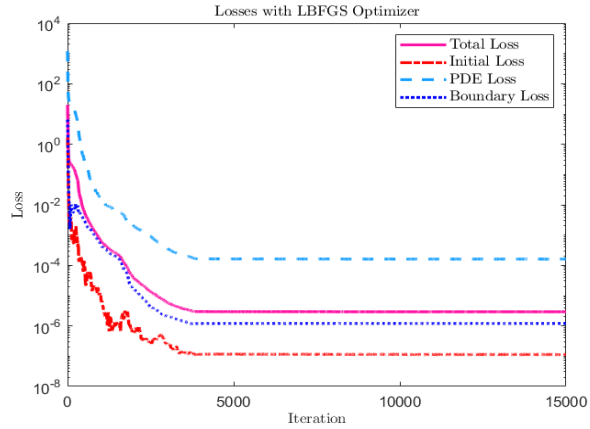


Figure 6. Training Loss Components Using LBFGS Optimizer: PDE, Initial Condition, and Boundary Condition

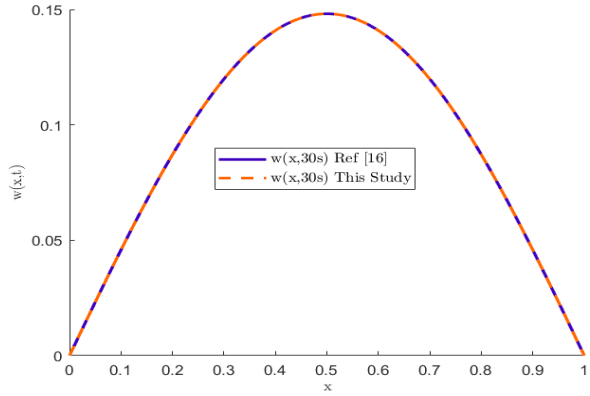


Figure 7. Comparison of responses from Chen et al. [16] and the PINNs method. Complete agreement indicates accuracy

In this study, the deep learning library PyTorch version 2.5.0 was used to construct a Physics-Informed Neural Network. This library is comprehensive in the field and facilitates the development of Physics-Informed Neural Networks. Also, a free Google Colab account with a T4 graphics processor, 16 GB of graphics memory, and 12 GB of RAM was utilized for conducting experiments.

After adjusting the hyperparameters according to the physical parameter issue, assuming that the height remains constant while the width changes $\hat{I}(\hat{x})= \hat{A}(\hat{x})$ in Table 2, and the curvature will be examined. specifically in section 5.1 concerning forced vibration analysis.

5. 1. Forced Vibration Analysis In this section, the analysis of forced vibrations is examined. is conducted based on the dimensionless parameters defined in Equation 14. The values are L = 1 m, P = 1.68 $\frac{N}{m}$, $\Omega = 6.0721\pi\ Hz$, $\rho A_0 = 0.1\frac{kg}{m}$ and EI_0= 0.16 $\frac{N}{m^2}$. By substituting these values into the parameters defined in Equation 16, the forced vibrations are examined according to Table 2.

It is assumed that the height remains constant, while the width changes exponentially ($e^{\delta_i x}$). Here, δ_i is a dimensionless scaling coefficient specified in Table 2. The initial conditions are outlined in Equation 52.

It is noteworthy that the validation conducted in this study was performed using two numerical methods: the sixth-order Galerkin method and finite element analysis with 22 elements. Ultimately, the system of equations was solved using the fourth-order Runge-Kutta method.

The absolute error was also calculated by comparing the results obtained from the finite element method with those obtained from the Physics-Informed Neural Networks (PINNs), which will be explained in detail below.

It is noteworthy that all graphs in this section represent dimensionless parameters.

$$\frac{\hat{I}(\hat{x})}{\hat{A}(\hat{x})}\frac{\partial^4\hat{w}}{\partial\hat{x}^4} + 2\frac{\hat{I}'(\hat{x})}{\hat{A}(\hat{x})}\frac{\partial^3\hat{w}}{\partial\hat{x}^3} + \frac{\hat{I}''(\hat{x})}{\hat{A}(\hat{x})}\frac{\partial^2\hat{w}}{\partial\hat{x}^2} + \frac{\partial^2\hat{w}}{\partial\hat{t}^2} = \hat{P}sin(\pi\hat{x})\,cos(\hat{\Omega}\hat{t}) \quad (53)$$

The effect of beam curvature, represented as $\delta_1 = 0$, illustrates the fixed bending patterns at various stages under applied forces, particularly in the context of forced vibrations. As shown in Figure 8, the maximum deflection consistently occurs at the center (1/2) of the beam, where both ends exhibit the same stiffness due to external loading. These observations highlight the intrinsic symmetry in the bending behavior of beams with uniform cross-sections under external forces. Following these findings, a comparison of beam behavior under different levels of applied force will be conducted. Figure 9 also depicts the behavior of this beam and its vibration

TABLE 2. Mechanical parameters for bending analysis

Mechanical parameters	$\hat{I}(\hat{x})$	$\hat{A}(\hat{x})$	δ_1	δ_2	δ_3	$\hat{P}$	$\hat{\Omega}$
Values	$e^{\delta_i\hat{x}}$	$e^{\delta_i\hat{x}}$	0	0.5	-0.5	$\frac{10.5}{e^{\delta_i\hat{x}}}$	4.8π

amplitude, which can be further analyzed using the finite element method as illustrated in Figure 10. The absolute error depicted in Figure 11 between these two methods, which is less than 0.07%, indicates that the PINNs method is reliable.

The impact of beam expansion with a profile of $\delta_2 = 0.5$ is described. We evaluate the deflection of the beam at various stages of bending, as shown in Figure 12. Observations indicate that the deflection shifts from the center towards the left, revealing a difference in stiffness between the two sides of the beam. Consequently, the maximum deflection is observed on the stiffer side, leading to a reduction in the vibration amplitude compared to beams with $\delta_1 = 0$. This reduction can be seen in Figure 13, along with the harmonic shape, and a comparison with the finite element method in Figure 14 shows that the results of both are identical. In Figure 15. The absolute error between the finite element results and PINNs is approximately 0.34%.

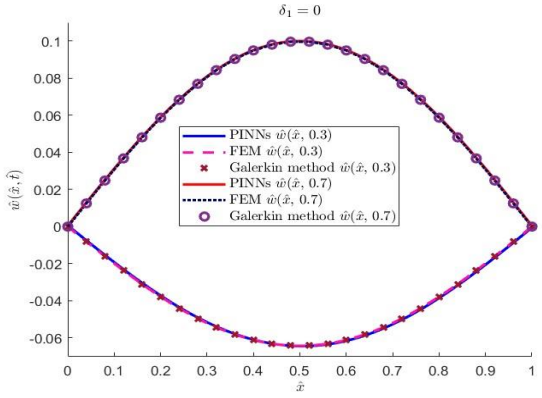


Figure 8. Forced vibration of beam with constant cross section area ($\delta_1 = 0$). comparison of the Galerkin, FEM and predicted PINNs

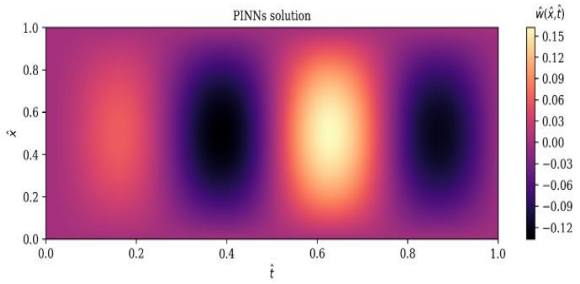


Figure 9. Transverse displacement of beam with constant cross section area ($\delta_1 = 0$). color bar represents PINNs solution

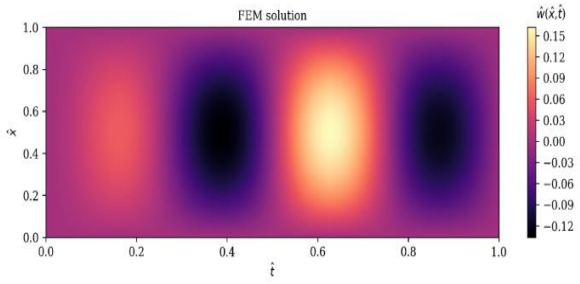


Figure 10. Transverse displacement of beam with constant cross section area ($\delta_1 = 0$). color bar represents FEM solution

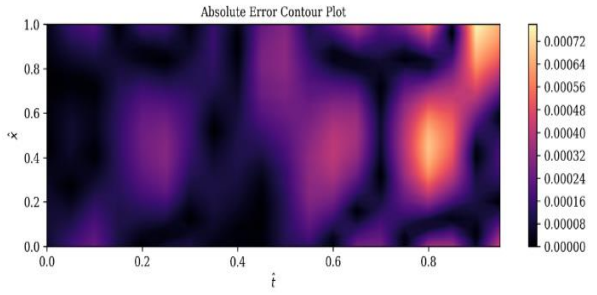


Figure 11. Absolute error in prediction PINNs & FEM $|w_{Pred} - w_{FEM}|$ for beam with constant cross section area ($\delta_1 = 0$)

between the two sides of the beam. Consequently, the maximum deflection is observed on the stiffer side, leading to a reduction in the vibration amplitude compared to beams with $\delta_1 = 0$. This reduction can be seen in Figure 13, along with the harmonic shape, and a comparison with the finite element method in Figure 14 shows that the results of both are identical. In Figure 15. The absolute error between the finite element results and PINNs is approximately 0.34%.

The impact of beam expansion with a profile of $\delta_2 = 0.5$ is described. We evaluate the deflection of the beam at various stages of bending, as shown in Figure 12. Observations indicate that the deflection shifts from the center towards the left, revealing a difference in stiffness between the two sides of the beam. Consequently, the maximum deflection is observed on the stiffer side, leading to a reduction in the vibration amplitude compared to beams with $\delta_1 = 0$. This reduction can be seen in Figure 13, along with the harmonic shape, and a comparison with the finite element method in Figure 14 shows that the results of both are identical. In Figure 15, the absolute error between the finite element results and PINNs is approximately 0.34%.

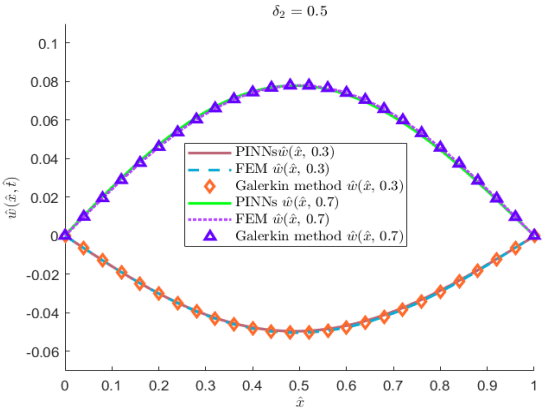


Figure 12. Forced vibration of beam with variable cross section area ($\delta_2 = 0.5$). comparison of the Galerkin, FEM and PINNs

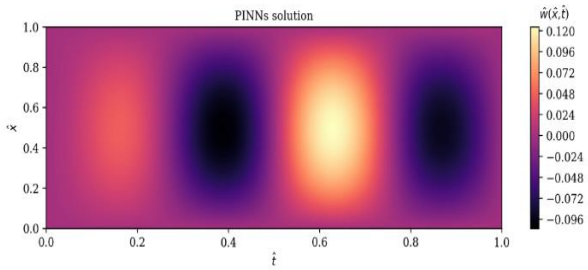


Figure 13. Transverse displacement of beam with constant cross section area ($\delta_2 = 0.5$). color bar represents PINNs solution

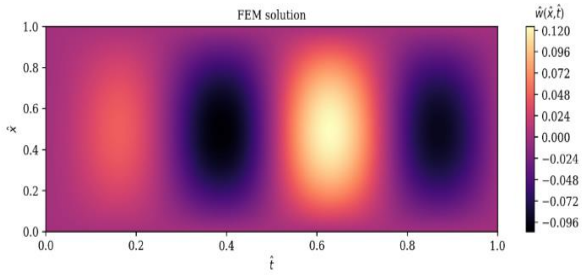


Figure 14. Transverse displacement of beam with constant cross section area ($\delta_2 = 0.5$). color bar represents FEM solution

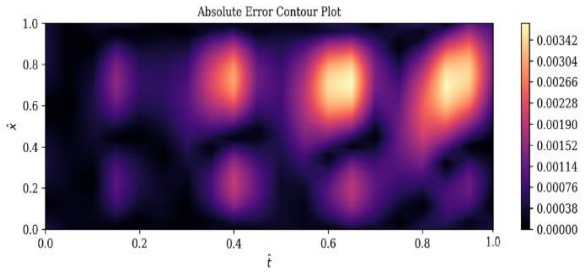


Figure 15. Absolute error in prediction PINNs & FEM $|w_{Pred} - w_{FEM}|$ for beam with variable cross section area ($\delta_2 = 0.5$)

The consequences of narrowing the beam with $\delta_3 = -0.5$ focus on describing the deflection of the beam during various bending stages, as shown in Figure 16. Observations indicate that the deflection shifts from the center towards the right, resulting in increased stiffness on one side of the beam. Consequently, the maximum deflection occurs on the stiffer side.

Furthermore, this configuration leads to an increase in the intensity of vibrations compared to beams with $\delta_1 = 0$. This increase is clearly visible in Figure 17, and when compared to the finite element method in Figure 18, it becomes evident that pins serve as a faster and more reliable solution, as the absolute error calculated in this case in Figure 19 is approximately 0.5%.

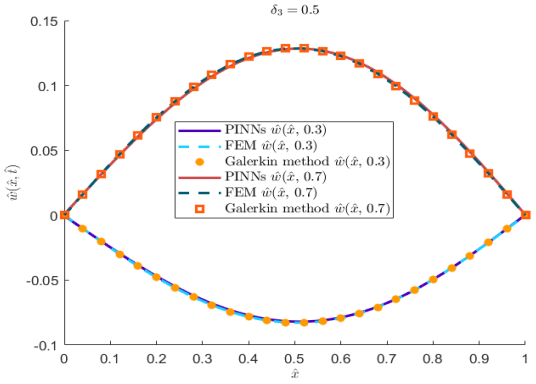


Figure 16. Forced vibration of beam with variable cross section area ($\delta_3 = -0.5$). comparison of the Galerkin, FEM and PINNs

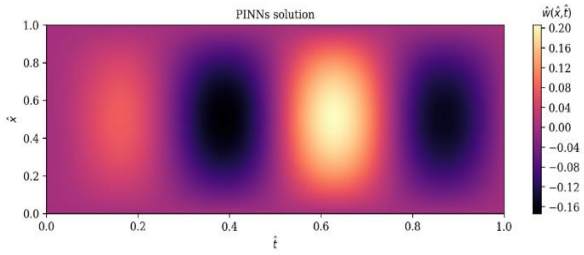


Figure 17. Transverse displacement of beam with constant cross section area ($\delta_3 = -0.5$). color bar represents PINNs solution

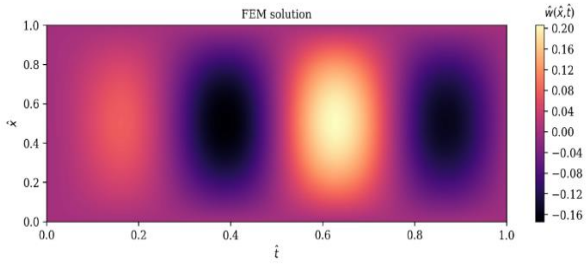


Figure 18. Transverse displacement of beam with constant cross section area ($\delta_3 = -0.5$). color bar represents FEM solution

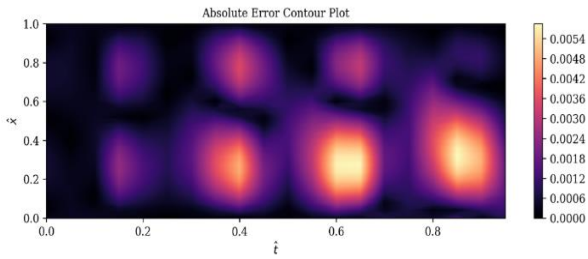


Figure 19. Absolute error in prediction PINNs & FEM $|w_{Pred} - w_{FEM}|$ for beam with variable cross section area ($\delta_3 = -0.5$)

6. CONCLUSIONS

In this study, physics-informed neural networks (PINNs) are used to formulate the equations of motion based on Euler-Bernoulli beam theory and Hamilton's principle to simulate the bending behavior of a beam with a variable cross-section. By applying boundary and initial conditions, a loss function is formulated, which allows for the smooth solution of partial differential equations (PDEs). The efficiency of physics-informed neural networks (PINNs) is evaluated by comparing its prediction accuracy with the results of data collected from the Galerkin and finite element methods. In all cases, the results obtained from the proposed physics-informed neural network (PINN) approach closely match the results obtained from the Galerkin and finite element methods. However, achieving the same level of accuracy as the Galerkin method requires a larger number of basics functions and the calculation of additional frequencies and finding mode shapes. For example, in this study, this problem was solved with a 6th-order Galerkin method. Finite element methods must first obtain the weak form of the equation after obtaining the differential equation, then obtain the stiffness, mass, and force matrices, and find a suitable mesh, which in this case consisted of 22 elements, and finally solve the system of equations with numerical methods such as the fourth-order Runge-Kutta method, which was used in this study, which significantly increases the computation time and the solution of a problem. In contrast, the time required for the proposed physics-informed neural network (PINN) framework is significantly shorter and provides results that match the accuracy of the Galerkin and finite element methods, even with a large number of basic functions. Therefore, expanding the application of physics-informed neural networks (PINNs) is essential to address more complex real-world problems. Furthermore, once trained, neural networks can quickly provide solutions at any given point with minimal additional computational effort. The most important results are summarized below:

1. It has been established that the PINNs Architecture, utilizing the specified Hyperparameters hidden layers configured as Neuron layers $N_h = 3$, $N_n = 100$, N_{pde}= 1000, N_{ic} = 200 and N_{bc}= 400, l_r = 0.1, Max- evaluations = $50k$, epochs = $15k$ Results obtained from the 6th-order Galerkin method and finite element method with 22 elements provide accurate and reliable predictions.
2. In forced vibrations, an increase in the delta value within the variable width causes one end of the beam to widen. This change creates a condition where one end of the beam has greater stiffness than the opposite end, resulting in a significant deviation from the center of the beam. Consequently, the overall distribution of forces within the beam changes, further amplifying the response to the occurrence of vibrations. In forced vibrations, an increase in the delta value in the variable width causes the beam end to widen and reduce the amplitude of vibrations.
3. In forced vibrations, decreasing delta in the beam width causes one end to become narrower than the other. This change in geometry increases the vibration amplitudes compared to a wider beam, leading to a non-uniform stress distribution that ultimately creates more pronounced oscillations under dynamic loads and results in greater stiffness in a portion of the system, with greater stiffness at one end of the beam.
4. Physics-informed neural networks (PINNs) effectively capture forced vibrations in beams with variable cross-sections. They provide solutions 45 times faster than the Galerkin method and 50 times faster than the finite element method, offering a highly efficient approach for analyzing complex beam behavior and facilitating faster design iterations.

7. REFERENCES

1. Elshafei MA. FE modeling and analysis of isotropic and orthotropic beams using first order shear deformation theory. Materials Sciences and Applications. 2013;4(1):77-102. https://doi.org/10.4236/msa.2013.41010
2. Thomsen DK, Søe-Knudsen R, Balling O, Zhang X. Vibration control of industrial robot arms by multi-mode time-varying input shaping. Mechanism and Machine Theory. 2021;155:104072. https://doi.org/10.1016/j.mechmachtheory.2020.104072
3. Li S, Shen L. Seismic optimization design and application of civil engineering structures integrated with building robot system technology. High-tech And Innovation Journal. 2024;5(4):1118-34. https://doi.org/10.28991/HIJ-2024-05-04-017
4. Makrup L, Pawirodikromo W, Muntafi Y. Comparison of structural response utilizing probabilistic seismic hazard analysis and design spectral ground motion. Civil Engineering Journal. 2024;10:235-51. https://doi.org/10.28991/CEJ-SP2024-010-012
5. Canales F, Mantari J. Free vibration of thick isotropic and laminated beams with arbitrary boundary conditions via unified formulation and Ritz method. Applied Mathematical Modelling. 2018;61:693-708. https://doi.org/10.1016/j.apm.2018.05.005
6. Moreno-García P, dos Santos JVA, Lopes H. A review and study on Ritz method admissible functions with emphasis on buckling and free vibration of isotropic and anisotropic beams and plates. Archives of Computational Methods in Engineering. 2018;25(3):785-815. https://doi.org/10.1007/s11831-017-9214-7
7. Azzara R, Carrera E, Filippi M, Pagani A. Vibration analysis of thermally loaded isotropic and composite beam and plate structures. Journal of Thermal Stresses. 2023;46(5):369-86. https://doi.org/10.1080/01495739.2023.2188399
8. Ghugal YM, Sharma R. A hyperbolic shear deformation theory for flexure and vibration of thick isotropic beams. International Journal of Computational Methods. 2009;6(04):585-604. https://doi.org/10.1142/S0219876209002017

9. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational physics. 2019;378:686-707. https://doi.org/10.1016/j.jcp.2018.10.045
10. Namdeo J, Dubey S, Dorji L. Dynamic analysis of isotropic homogeneous beams using the method of initial functions and comparison with classical beam theories and FEM. Complexity. 2023;2023(1):6636975. https://doi.org/10.1155/2023/6636975
11. Zienkiewicz O, Taylor R, Ziaie-Moayed R, Kamalzare M, Safavian M, Zhou S, et al. Numerical Study of Floating Stone (static). Comput Geotech. 2013;2(2).
12. Kumar R, Tiwari R, Prasad R. Numerical solution of partial differential equations: finite difference method. Mathematics and Computer Science Volume 1. 2023:353-72. https://doi.org/10.2307/2007849
13. Lipnikov K, Manzini G, Shashkov M. Mimetic finite difference method. Journal of Computational Physics. 2014;257:1163-227. https://doi.org/10.1016/j.jcp.2013.07.031
14. Wang J, Wu R. The extended Galerkin method for approximate solutions of nonlinear vibration equations. Applied sciences. 2022;12(6):2979. https://doi.org/10.3390/app12062979
15. Polycarpou AC. Introduction to the finite element method in electromagnetics: Springer Nature; 2022.
16. Chen Z, Lai S-K, Yang Z. AT-PINN: Advanced time-marching physics-informed neural network for structural vibration analysis. Thin-Walled Structures. 2024;196:111423. https://doi.org/10.1016/j.tws.2023.111423
17. Bazmara M, Silani M, Mianroodi M, editors. Physics-informed neural networks for nonlinear bending of 3D functionally graded beam. Structures; 2023: Elsevier.
18. Bazmara M, Mianroodi M, Silani M. Application of physics-informed neural networks for nonlinear buckling analysis of beams. Acta Mechanica Sinica. 2023;39(6):422438. https://doi.org/10.1007/s10409-023-22438-x
19. Kapoor T, Wang H, Núñez A, Dollevoet R. Transfer learning for improved generalizability in causal physics-informed neural networks for beam simulations. Engineering Applications of Artificial Intelligence. 2024;133:108085. https://doi.org/10.1016/j.engappai.2024.108085
20. Kapoor T, Wang H, Núñez A, Dollevoet R. Physics-informed neural networks for solving forward and inverse problems in complex beam systems. IEEE Transactions on Neural Networks and Learning Systems. 2023;35(5):5981-95. https://doi.org/10.1109/TNNLS.2023.3310585
21. Söyleyici C, Ünver HÖ. A Physics-Informed Deep Neural Network based beam vibration framework for simulation and parameter identification. Engineering Applications of Artificial Intelligence. 2025;141:109804. https://doi.org/10.1016/j.engappai.2024.109804
22. Teloli RdO, Tittarelli R, Bigot M, Coelho L, Ramasso E, Le Moal P, et al. A physics-informed neural networks framework for model parameter identification of beam-like structures. Mechanical Systems and Signal Processing. 2025;224:112189. https://doi.org/10.1016/j.ymssp.2024.112189
23. Barillaro L. Deep learning platforms: PyTorch. 2025. https://doi.org/10.1016/B978-0-323-95502-7.00093-2
24. Vadyala SR, Betgeri SN, Betgeri NP. Physics-informed neural network method for solving one-dimensional advection equation using PyTorch. Array. 2022;13:100110. https://doi.org/10.1016/j.array.2021.100110
25. Chaudhary A, Chouhan KS, Gajrani J, Sharma B. Deep learning with PyTorch. Machine learning and deep learning in real-time applications: IGI Global; 2020. p. 61-95.
26. Hou Q, Li Y, Singh VP, Sun Z, Wei J. Physics-informed neural network for solution of forward and inverse kinematic wave problems. Journal of Hydrology. 2024;633:130934. https://doi.org/10.1016/j.jhydrol.2024.130934
27. Mehmood F, Ahmad S, Whangbo TK. An efficient optimization technique for training deep neural networks. Mathematics. 2023;11(6):1360. https://doi.org/10.3390/math11061360
28. Tong Q, Liang G, Cai X, Zhu C, Bi J. Asynchronous parallel stochastic Quasi-Newton methods. Parallel computing. 2021;101:102721. https://doi.org/10.1016/j.parco.2020.102721
29. Jagtap AD, Kawaguchi K, Karniadakis GE. Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. Journal of Computational Physics. 2020;404:109136. https://doi.org/10.1016/j.jcp.2019.109136
30. Jagtap AD, Kawaguchi K, Em Karniadakis G. Locally adaptive activation functions with slope recovery for deep and physics-informed neural networks. Proceedings of the Royal Society A. 2020;476(2239):20200334. https://doi.org/10.1098/rspa.2020.0334

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

این مطالعه به بررسی ارتعاشات اجباری تیرهای ایزوتروپیک با مقاطع متغیر می‌پردازد که بر مبنای نظریه تیر اویلر- برنولی مدلسازی شده است. با بهره‌گیری از اصل هامیلتون، معادلات دیفرانسیل حاکم بر دینامیک تیر استخراج شده و رفتارهای پیچیده ارتعاشی تحلیل می‌شوند.در این تحقیق، شبکه‌های عصبی مبتنی بر فیزیک (PINNs) به عنوان روشی نوآورانه و بدون نیاز به مش‌بندی، معرفی شده‌اند. این فناوری قادر است نتایج سریع و دقیقی را با تلفیق قوانین فیزیکی در فریمورک یادگیری ماشین ارائه کند. مزیت اصلی PINNs نسبت به روش‌های سنتی مانند اجزا محدود (FEM) یا روش‌های تفاضلی، ساده‌سازی فرآیند محاسباتی و کاهش زمان و منابع مورد نیاز است، زیرا دیگر نیاز به ساخت مش و تنظیم شبکه‌های پیچیده ندارد.نتایج به‌دست‌آمده . با استفاده از روش گالرکین مرتبه ششم واجزا محدود ، صحت و کارایی بالای این رویکرد را در تحلیل ارتعاشات تیرهای با مقاطع متغیر تأیید می‌کند. در مجموع، این پژوهش، کاربردهای PINNs در ارزیابی ارتعاشات را توسعه می‌دهد و در بهینه‌سازی طراحی و عملکرد در حوزه‌های مختلف مهندسی، از جمله سازه‌ها و سیستم‌های مکانیکی، نقش مؤثری ایفا می‌کند.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1063-1076

International Journal of Engineering

Journal Homepage: www.ije.ir

Experimental and Numerical Investigations for Improvement of Building Ventilation Performance by Using Solar Chimney Integrated with Different Phase Change Materials and Nanoparticles

W. Younus Abeid Al-Zubaidi, A. A. Ranjbar*

Faculty of Mechanical Engineering, Babol Noshirvani University of Technology, Babol, Iran

P A P E R I N F O

Paper history:
Received 26 February 2025
Received in revised form 30 June 2025
Accepted 19 July 2025

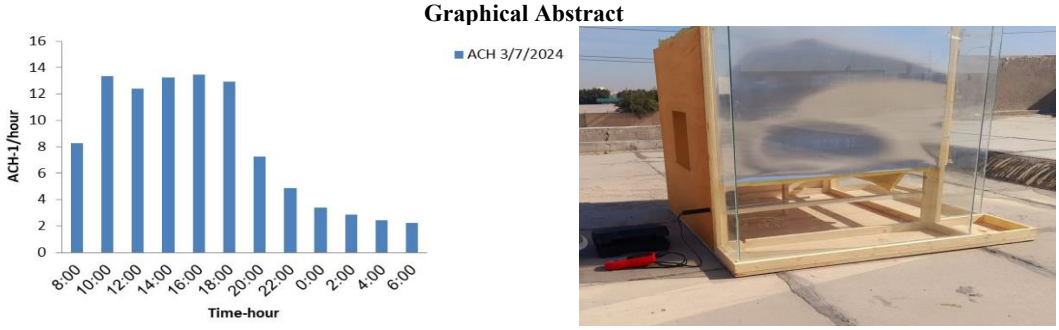
Keywords:
Natural Ventilation
Solar Chimney
Thermal Performance
Phase Change Materials

A B S T R A C T

A detailed experimental and numerical investigation of a constructed solar chimney (SC) model and its thermal performance is presented, conducted under the climatic conditions of Al-Kut City, Iraq. This study focuses on the thermal performance of two vertical solar chimneys oriented towards the south and west directions. The first configuration consists of a building with two vertical solar chimneys, incorporating two types of phase change materials (PCMs): Al-Dura paraffin wax and block-shaped paraffin wax (PCM1) and (PCM2), respectively. These PCMs are placed in the same aluminum container, separated by partitions, and equipped with fins to enhance the thermal contact area. In the second configuration, conductive mono nanoparticles (NP), specifically aluminum oxide (Al_2O_3), were introduced into both (PCM1) and (PCM2) at a volume fraction of approximately 1.6% to improve the thermal characteristics of the paraffin and overall heat storage performance. The study investigates the role of paraffin wax as an energy storage medium along with nanoparticles, where the PCM stores thermal solar energy during the day and discharges it at night through the phase change process. Numerical results demonstrate that the PCMs significantly affect air change per hour (ACH) during the day and extend ventilation hours post-sunset. The second configuration showed an ACH improvement of approximately 5.35% and a reduction in room temperature by 4.7°C compared to the first case. Experimental results indicated a 5.2% improvement in ACH and a decrease of 3.9°C in room temperature. Both experimental and numerical results confirm that integrating nanofluids with PCM significantly enhances the overall efficiency of the solar chimney. The novelty of this study lies in the innovative integration of phase change materials with conductive nanoparticles, which not only enhances the thermal performance but also significantly extends ventilation periods, especially under the high solar radiation conditions typical of AL-Kut city, Iraq. Finally, due to the intense solar radiation in AL-Kut, the passive ventilation rates achieved by the solar chimney were found to be highly satisfactory.

***doi**: 10.5829/ije.2026.39.05b.02*

Graphical Abstract



*Corresponding author email: ranjbar@nit.ac.ir (A. A. Ranjbar)

Please cite this article as: Younus Abeid Al-Zubaidi W, Ranjbar AA. Experimental and Numerical Investigations for Improvement of Building Ventilation Performance by Using Solar Chimney Integrated with Different Phase Change Materials and Nanoparticles. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1063-76.

NOMENCLATURE

ACH	Air change per hour	PCM	Phase change material
Sc	Solar chimney	NP	Nanoparticles
H.F	Heat of fusion (KJ/kg)	ES	Energy storage
M.t	Melting temperature (K)	*Cp*	Specific heat (J/kg.K)
Slr	solar intensity (w/m^2)	T.room Ave	Average room temperature
k	Thermal conductivity(W/m.K)	**Greek Symbols**	
v	Viscosity (kg/m.s)	ρ	Density (kg/m^3)

1. INTRODUCTION

The solar chimney is one of the applications of negative ventilation effectiveness in providing healthy and comfortable thermal conditions, which is an air channel based on the principle of buoyancy, where the solar energy is absorbed after the energy is released to the air to create currents indoor air. The importance of the solar chimney has increased in recent years as a result of the great benefits compared to mechanical ventilation in terms of energy saving and economic cost in addition to environmental effect. The utilization of suitable energy storage is one of the options for improving the efficiency of the solar based power plants, which is as important as developing new sources of energy (1, 2). Several kinds of research and findings were conducted to study and enhance these ventilation systems arranged in this paper. Tolabi et al. (3) emphasize the importance of estimating solar radiation in solar energy systems, which requires precise and often costly equipment. The Bees Algorithm, a population-based heuristic search method implemented in MATLAB, is a new approach based on the Angstrom model. This method predicts the monthly average daily global solar radiation on a horizontal surface. The Angstrom model's experimental coefficients are calculated for Iran's six climate zones. The results demonstrate the effectiveness of this technique in providing more accurate solar radiation estimates. Ajay and Kundan (4) used a 40:60 ethylene-glycol-water mixture as the base fluid, dispersing α-Al_2O_3 nanoparticles with an average size of 20 nm to create nanofluids with four distinct volumetric concentrations (0.05%, 0.075%, 0.1%, and 0.125%). The use of nanofluid instead of a water-ethylene glycol mixture has been shown to enhance the overall efficiency of solar collectors. Bassiouny and Koura (5) employed both numerical and analytical methods to investigate the effect of vertical chimney width on space ventilation. They found that, while keeping the chimney entrance size constant, increasing the chimney width from 0.1 to 0.3 m led to a nearly 25% increase in volume flow rate. Bassiouny and Koura (6) expanded their research on inclined solar chimney performance by analyzing, both analytically and numerically, the impact of varying chimney inclination angles on indoor flow patterns and air change per hour at a latitude of 28.4°. Both analytical and numerical results showed that optimal flow rates were achieved for inclination angles between 45° and 70°. Karimipour-Fard and Beheshti (7) improved the performance of the solar chimney power plant (SCPP) by enhancing the geometrical characteristics of the collector and chimney. In the absence of solar radiation, thermal energy storage was used to generate power in the new optimal configurations. The results showed that the output power of the SCPP increased by nearly 139%, and the overall performance of the power plant improved by approximately 68.1%. Majeed and Ali (8) conducted an experimental investigation on the use of solar chimneys as heating sources in Iraqi settings. The experimental model consisted of a room with a 40° southward slope and a solar collector of specific dimensions (2.5×1.29×1.07 m). Data were collected at various points within the test area during January and February. The results showed that sun-facing walls could enhance the stored energy. Increased heat energy also promoted higher air flow rates inside the building. Jubear and Ghareer (9) used the ANSYS simulation program to perform a numerical study on how chimney inclination angles affect building ventilation. Three angles 30°, 45°, and 60° were employed, where 45° indicates the best performance. A building with two solar chimneys was later designed by the researchers. Results indicated a 31.9% improvement in ventilation efficiency and a (1.34°C) drop in temperature. Kaneko et al. (10) investigated how well solar chimneys with built-in Sulfate Dehydrate ($Na_2SO_4.10H_2O$) latent heat storage performed in terms of ventilation. According to a study done in Osaka, Japan, PCM storage inside the solar chimney was accessible for natural ventilation at night and in the evening in the event that PCM melted completely during the day. Amori and Mohammed (11) assessed the amount of energy stored and released by paraffin when added to a solar chimney, using both experimental and numerical methods. The results indicated that 5016 W of energy was stored during the 5 hour charging period, while 4954.4 W was released during the 9-hour discharging period. Imran et al. (12) conducted both experimental and numerical analyses to determine the optimal performance of a solar chimney under various geometrical conditions. Natural convection was achieved through a turbulent flux, with three air gap sizes in the solar chimney (50 mm, 100 mm, and 150 mm). The experiments were conducted at

varying angles (15°–60°) and solar radiation levels (150–750 W/m²). The results showed that the maximum ventilation rate occurred at the ideal inclination angle of 60°, with a solar radiation intensity of 750 W/m² and a 50 mm air gap. Chung et al. (13) provided the optimum values for the parameters that affect solar chimney performance. The results show that the optimum air gap is between 0.6 and 1.0 m, the chimney length is between 1.5 and 2 m, and the induced air velocity is between 0.04 and 0.22 m/s. When the air gap is 10 cm and the angle of inclination rises from 15 to 45 degrees, the ventilation rate increases by 24%. Li and Liu (14) conducted an experimental study to examine the thermal performance of a solar chimney using PCM at three different heat flux levels: 500, 600, and 700 W/m². The results showed that the PCM was completely melted only at 700 W/m², with the maximum air flow rate reaching 0.04 kg/s, and thermal efficiency exceeding 80%. This study aims to assess the thermal performance of a solar chimney and improve ventilation rates through two vertical solar chimneys. The model includes a room connected to two vertical solar chimneys, with PCM and mono nanoparticles added. The use of mechanical air cooling, especially in hot climates during the summer and sometimes in autumn and spring, extends the hours of active operation to provide optimal thermal conditions. Assari et al. (15) conducted experiments with a phase change material (PCM) triplex-tube heat exchanger, using two types of sinusoidal inner tubes: full-width and half-width. The inner, outer, and both tubes were utilized to carry out the charging and discharging processes. The results showed that the PCM melting and solidification times were shorter for the full-width sinusoidal inner tube compared to the half-width sinusoidal inner tube during both charging and discharging operations, whether from the inner tube, outer tube, or both sides. The volume and duration of PCM melting were significantly improved when comparing these processes to those of a straight inner tube. Krishna et al. (16) conducted both experimental and numerical studies to investigate the effects of using phase change material (PCM) in the adiabatic section of heat pipes for electronic cooling applications. Tricosane (100 ml) was chosen as the PCM, and Al_2O_3 nanoparticles with volume fractions of 0.5%, 1.0%, and 2.0% were dispersed throughout the PCM using an ultrasound technique. The system was tested under heating powers of 13W, 18W, and 23W. At 1.0% and 2.0% nanoparticle concentrations, the system's performance with Tricosane and nanoparticles improved and declined, respectively, suggesting an optimal nanoparticle concentration. Jadidi and Jadidi (17) highlighted that many engineering problems, such as the Stefan problem involving melting and solidification in phase change materials (PCMs), are common real-world challenges. They proposed an algorithm to solve the one-dimensional Stefan problem under various boundary conditions. Additionally, the algorithm can be easily extended to 2D and 3D Stefan problems using the finite difference approach. Kumar et al. (18) examined the literature on TiO_2 nanoparticles, specifically those that had been suspended in a 60:40 mixture of ethylene glycol and water. The results showed that the rate of heat transfer is higher when the water-ethylene glycol mixture is used at this ratio than when water is used alone. Ebrahimnataj-Tiji et al. (19) examined a PCM-enhanced solar chimney both with and without fins attached to the absorber plate. The results showed that using PCM as a storage medium improves the temperature homogeneity of the room. However, the mean of the room temperature achieved was 14.68°C, which is significantly below the thermal comfort level. Therefore, the addition of fins to the absorber was investigated to raise the room's temperature. The findings indicate that the presence of fins significantly enhances the PCM's energy storage capacity. More importantly, compared to the non-finned case, the use of fins in the PCM-based solar chimney system increases the room's mean temperature by 20%. Additionally, the finned absorber with PCM leads to more non-uniform airflow compared to the non-finned case. Selvan et al. (20) expanded motor cooling by adding two cooling channels. PCM serves as a secondary coolant, while the liquid acts as the primary coolant. Even when the liquid cooling operates at its lowest working point, this innovative cooling technique helps to maintain the bracket temperature within the permissible limit. To thoroughly examine the PCM, the main focus was on the PCM coolant channel, with the liquid coolant kept under specific working conditions. The motor cooling performance was influenced by the PCM thickness. Three different PCM channel thicknesses6 mm, 8 mm, and 10 mm were examined, with the best results observed for a PCM thickness of 6 mm, which improves heat transfer. However, aside from the articles mentioned above, similar studies conducted by other researchers can also be implicitly referred (21-26).

The goal of this study is to assess the effectiveness of natural ventilation using two solar chimneys oriented south and west, incorporating phase change materials (PCMs) as energy storage (ES) and enhancing their thermal properties with aluminum oxide (Al_2O_3) nanoparticles at a 1.6% volume fraction. The novelty of this study lies in the integration of two types of PCMs with aluminum oxide nanoparticles, which significantly enhance the thermal characteristics of paraffin wax and optimize heat storage. This approach aims to extend natural ventilation after sunset in a single-story building in the hot climate of Al-Kut, Iraq. Passive ventilation, based on the buoyancy principle, is proven effective in maintaining thermal comfort in hot climates with high

solar radiation. By using conductive nanoparticles, this study reduces reliance on active energy systems and improves overall thermal performance. The solar radiation on the system falls within the optimal range of 2000–2500 kWh/m² [1], ensuring efficiency under local conditions.

2. MODEL DESCRIPTION

The model consists of a room made of wood connected to two solar chimneys oriented in different directions (south and west). Each vertical solar chimney has dimensions of (1.0 m length, 1.0 m width, and 0.3 m depth). The solar chimneys are connected to the room through ventilation openings measuring (1.0×0.3) m² on each chimney. The room is cubical, with dimensions of (1.0×1.0×1.0) m³. The solar chimneys are made of 1.0 mm aluminum absorber panels, and 5.0 cm glass wool insulation is placed between the absorber panel and the wood to prevent heat transfer from the panel to the room. In addition, 2.5 cm of insulation is used to cover the roof of the room. The absorber panel is wrapped on three sides with a 4.0 mm-thick glass cover. The room is equipped with a window measuring (0.3×0.3) m², which is placed in the center of the northern wall to allow air to enter during the operation of the solar chimney. Figure 1 illustrates the experimental model with two vertical solar chimneys, and the schematic of the experimental setup is shown in Figure 2. The two vertical solar chimneys use two types of phase change materials (PCMs): Al-Dura paraffin wax and block-shaped paraffin wax, placed in the same aluminum container, separated by partitions. Fins are added to the PCM container to enhance heat transfer. In another case, conductive mono nanoparticles of aluminum oxide (Al_2O_3) are added to the PCMs at a volume fraction of 1.6%.

3. AIR CHANGE PER HOUR

Air change per hour (ACH) stands for Air Changes per Hour and is commonly referred to as “air exchange rate” or “air change rate”. It is a measurement of how many times a volume of air within a room will be added, removed, or exchanged with outdoor air. It was calculated according to the following equation (27):

$$ACH = \frac{3600\,A\,V}{Vr} \quad (1)$$

where A represents the area of solar chimney air entrance which equals 0.3 m², V represents the average velocity of air in the chimney entrances measured every two hours by using the hot wire anemometer, and Vr represents the volume of room that equals 1.0 m³.

4. DESCRIPTION OF ENERGY STORAGE

Two types of wax were used as an energy storage. The thermo-physical properties of the two types of wax are illustrated in the Table 1. The wax is one of the phase change materials (PCM) that charge and discharge heat during the phase change process. The dimensions of the wax container were (0.7 m in length × 0.9 m in width). Two types of PCMs namely: Al-Dura paraffin wax and Block-shaped paraffin wax.

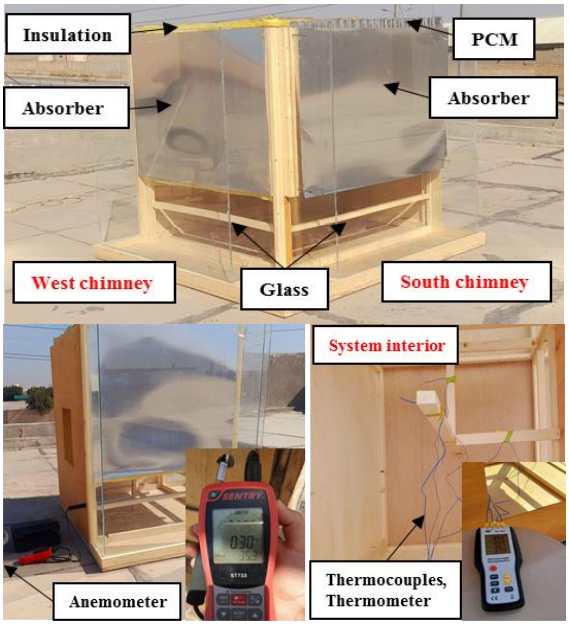


Figure 1. Experimental model with instrumnts

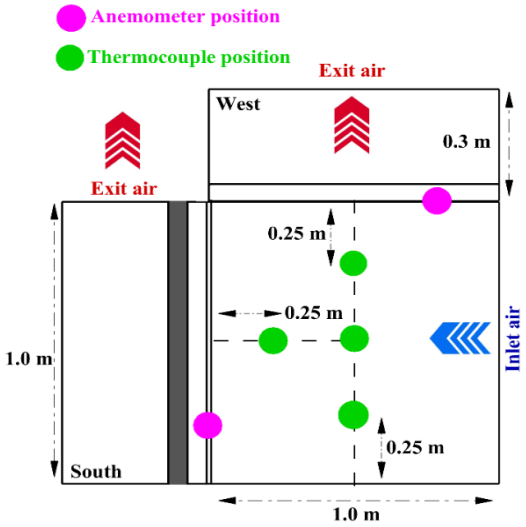


Figure 2. Schematic of experimental set up

[1] https://commons.wikimedia.org/wiki/File:483897main_GlobalPM2.5-map, https://www.pinterest.com/pin/1788070476839406962T

TABLE 1. Thermo-physical properties of wax (28, 29)

PCM	ρ	k	v	M.t	Cp	H.f
Al-Dura	786.5	0.329	0.01	326	2900	270
Block-shaped	802	0.260	0.017	329.5	2600	200
Al-Dura & Al_2O_3	836.15	0.344	0.0104	-	2737.81	250.15
Block-shaped & Al_2O_3	851.12	0.269	0.0166	-	2463.37	185.07

Thermo-physical properties of Al_2O_3 are listed in Table 2 (30, 31):

5. EXPERIMENTAL CASE STUDY

Two cases were studied over two days (3/7/2024 and 7/7/2024) in the summer climate of Al-Kut City, Iraq. In the first case, two vertical solar chimneys were used, incorporating two types of phase change materials (PCMs): Al-Dura paraffin wax and block-shaped paraffin wax, placed simultaneously in an aluminum container, separated by partitions. Fins were added to the PCM container to enhance heat transfer. The study was conducted from 8:00 AM to 6:00 PM on 3/7/2024. In the second case, conductive mono nanoparticles of aluminum oxide (Al_2O_3) were added to the PCMs at a volume fraction of 1.6% for the day of 7/7/2024, as shown in Table 3. The outcomes of the two cases were compared, and the impact of the conductive mono nanoparticles on the passive cooling system was evaluated.

5. 1. Experimental Study Procedure and Tools Measurements

The site study was conducted in Al-Kut City, Iraq, located at a longitude of (45.78E) and a latitude of (32.54N). Readings were taken every two hours, starting at 8:00 AM and finishing at 6:00 AM the following day. The measurements include solar radiation, temperature, and air velocity at the chimney air entrance. A solar power meter was used to measure the solar radiation intensity, while K-type thermocouples were employed to measure the temperatures within the building. The Sentry ST733 anemometer was used to measure air velocity at the chimney openings to calculate the air change rate (ACH) value. Tables 4 and 5 show the names, accuracy, and errors of the instruments.

TABLE 2. Thermo-physical properties of Nano-particles

Properties	ρ	k	Cp
Al_2O_3	3970	40	765

TABLE 3. Expermntal Conditions

Cases 1	3/7/2024	PCM1+ PCM2
Cases 2	7/7/2024	(PCM1+NP) + (PCM2+NP)

5. 2. Overall Uncertainty

The purpose of uncertainty calculation is to determine a confidence interval within which the true value is likely to fall. The smaller the uncertainty of the results, the higher the accuracy and precision of the parameters influencing the main data. In this section, the uncertainty of the experimental results has been calculated. According to the error distribution theory using the Taylor series, the following equation is used to calculate the uncertainty of a multivariate parameter with 95% confidence (14).

$$U_{95} = \left[\sum_{i=1}^{1} \left(\frac{\partial r}{\partial X_i}\right)^2 U_i^2\right]^{\frac{1}{2}} \quad (2)$$

Accordingly, the maximum calculation error in the graphs is 2.63%.

6. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section the experimental results are discussed. The results show that PCMs affects on ACH during the daylight period and elongates the hours of ventilation after sunset and average indoor temperature. In the first case, Table 5 depicts the thermal conditions for the day hours in Kut City, Iraq, on 3/7/2024. The peak solar intensity was 825 W/m^2 at 12:00 p.m., then progressively fell to 75 W/m^2 at 6:00 a.m. Moreover, the ambient temperature increases from 8:00 a.m. and reaches a maximum value 43 ^{o} C at 2:00 p.m. and then gradually decreases, especially in the evening and early

TABLE 4. Experimental insruments and accuracy

No.	Measuring Tools	Accuracy
1	Anemometer	±(0.03+3%)m/s
2	Solar power meter	±5 w/m^2
3	Thermocouples	±(0.1% of reading + 0.7°C)

TABLE 5. Experimental insruments and error

No.	Measuring Tools	Example	Error
1	Anemometer	measuring air velocity of 10 m/s,	±0.33 m/s
2	Solar power meter	measuring solar radiation intensity	±5% (typical)
3	Thermocouples	measuring 300°C	0.3°C + 0.7°C = ±1.0°C

morning hours reached lowest value 30 ° C at 6:00 a.m. The maximum ACH was noted around 4:00 p.m., which was equal (13.49), while the average minimum ACH was approximately (2.5) from 2:00 a.m. to 6:00 a.m. The relation between air change per hour (ACH) and test hours is illustrated in Figure 3. The experimental data shows that (ACH) begins to rise gradually after 8:00 a.m., peaks at 4:00 p.m. and drops to the lowest value from 12:00 a.m. to 6:00 a.m. due to solar intensity and sunlight orientation effects, but presences of PCMs the ventilation was continued after sunset. To specify the temperature of the building interior, three thermocouples were placed in each room at different levels. The readings were registered every two hours.

The average of the room temperatures reached the maximum values 46.4 ° C and 44.5 ° C at 2:00 p.m. and 4:00 p.m. respectively , while the minimum values were 33.5 ° C and 31.7 ° C at 4:00 a.m. and 6:00 a.m. respectively. The temperatures reading according to test periods are shown in Figure 4. The results also demonstrated that the ambient air temperature are the

TABLE 5. Experimental climate data for day (3/7/2024)

NO.	Time	T.Amb.	RH%	Slr. w/m²
1	00:00	34	15	0
2	02:00	32.6	18	0
3	04:00	31	22	0
4	06:00	30	21	75
5	08:00	34.5	17	370
6	10:00	40	14	675
7	12:00	42	12	825
8	14:00	43	11	775
9	16:00	42.4	11	520
10	18:00	41.5	12	120
11	20:00	39	13	0
12	22:00	37.6	14	0

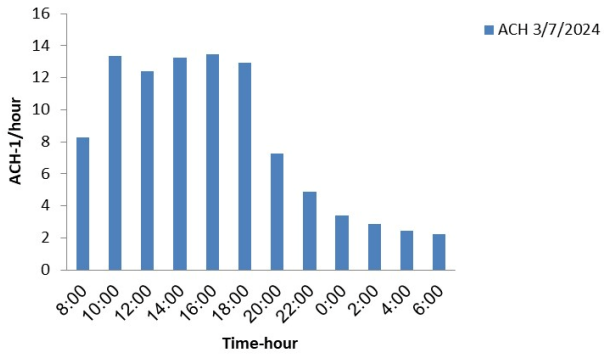


Figure 3. ACH values when using 2PCM

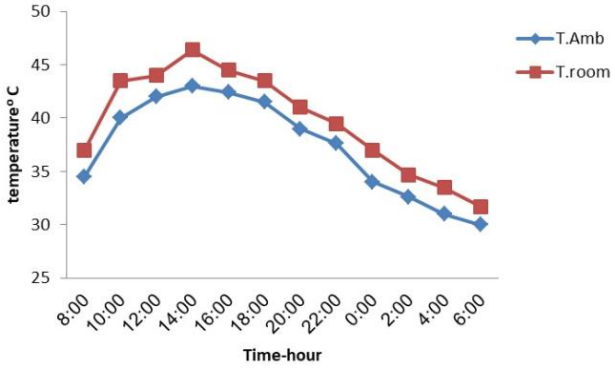


Figure 4. Room and ambient temperatures when using 2PCM

key element that affect the temperature within the building. There is a clear correlation between the temperature inside the building and the temperatures outside , the greater ambient temperature, the higher temperature inside the building.

In the second case conductive mono nanoparticles Aluminum oxide (Al_2O_3) added to Al-Dura paraffin wax & Block-shaped paraffin wax with volume fraction (1.6%). Table 6 shows the intensity of solar radiation for the day 7/7/2024, where it was the lowest value at 6:00 a.m. and amounted to 80 w/m^2, while the highest value was at noon, where it recorded 835 and 785 w/m^2 at 12:00 and 14:00, respectively, as in the sun's radiation, The highest temperature was at the afternoon hours, and it was recorded at 44 °C for 2:00 p.m., while the lowest temperature was at 6:00 a.m. and recorded at 29 °C. Table 6 also shows the average relative humidity.

The maximum ACH was noted around 4:00 p.m., which was equal (13.95) while the average minimum ACH was approximately (2.6) from 4:00 a.m. to 6:00 a.m. The ACH values are shown in Figure 5 as a

TABLE 6. Experimental climate data for day (7/7/2024)

NO.	Time	T.Amb.	RH%	Slr. w/m²
1	00:00	33	15	0
2	02:00	31	17	0
3	04:00	30.5	21	0
4	06:00	29	23	80
5	08:00	34	19	380
6	10:00	42	15	680
7	12:00	43	12	835
8	14:00	44	10	785
9	16:00	43.6	10	525
10	18:00	43	11	130
11	20:00	40.5	14	0
12	22:00	36	14	0

function of time. The results show the presences (Al_2O_3) with PCMs the ventilation was continued to 2:00 a.m. The average room temperatures reached the maximum values 46. 5 ° C and 46 ° C at 2:00 p.m. and 4:00 p.m. respectively , while the minimum values were 32 ° C and 30.5 ° C at 4:00 a.m. and 6:00 a.m. respectively. The temperatures reading according to test periods are shown in Figure 6.

6. 1. Comparison Between First and Second Experimental Cases The comparative analysis of the two experimental cases conducted on 3/7/2024 and 7/7/2024 highlights the significant improvements achieved by incorporating nanoparticle-enhanced phase change materials (2PCMs) in building thermal regulation. In the first case, conventional PCMs were utilized, and the results indicated notable effects on indoor temperature moderation and air change rates. The maximum indoor temperature reached 46.4 °C at 2:00 p.m., while the lowest was 31.7 °C at 4:00 a.m. The solar intensity peaked at 825 W/m² at noon and gradually decreased during the evening, influencing both indoor temperature and ventilation rates. The Air Changes per Hour (ACH) values demonstrated a pronounced increase after 8:00 a.m., peaking at 13.49 by 4:00 p.m., then dropping to a minimum of 2.5 during the early morning hours. In contrast, the second case incorporated mono-nanoparticles of aluminum oxide (Al_2O_3) into paraffin wax at a concentration of 1.6%. This modification led to a more favorable thermal environment. The maximum room temperature was slightly lower at 44 °C, and the minimum was 29 °C, suggesting better thermal control. Despite a slightly higher solar radiation intensity, reaching 835 W/m² at noon, the 2PCM with (Al_2O_3) system managed to stabilize indoor conditions more effectively than the traditional PCM setup. Importantly, the second case displayed a more uniform and sustained ventilation pattern. The ACH values followed a similar trend to the first case but exhibited a smoother distribution, indicating less thermal stress and more consistent air exchange throughout the day. This implies enhanced indoor air quality and prolonged natural ventilation effectiveness, even during hours when solar radiation diminished. The improved thermal conductivity provided by the nanoparticles allowed the 2PCM material to absorb and release heat more efficiently, thereby reducing peak indoor temperatures and extending the cooling effect beyond sunset.This suggests strong potential for implementing nanoparticle-enhanced PCMs in sustainable building applications, especially in hot climates where thermal regulation is critical.

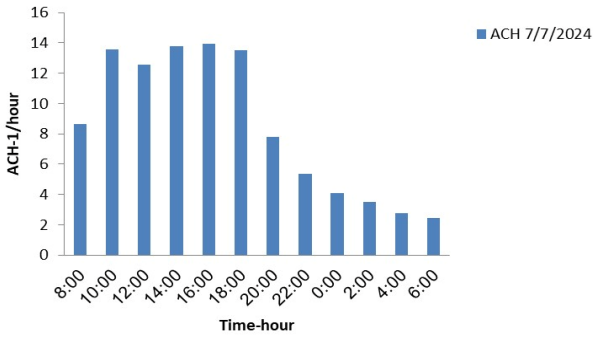


Figure 5. ACH values when using 2 PCM & NP volume fraction 1.6%

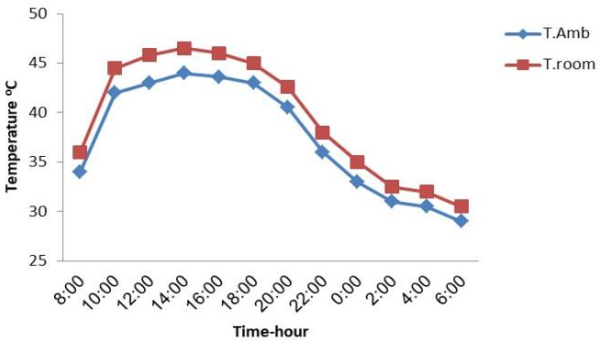


Figure 6. Room and ambient temperatures when using 2 PCM & NP volume fraction 1.6%

7. NUMERICAL SIMULATION

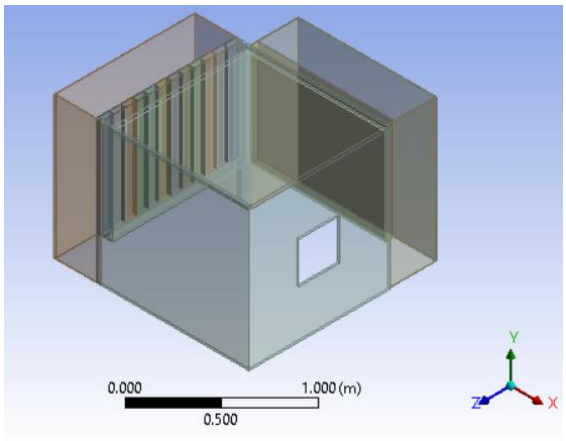
CFD analysis of the absorber tube of the solar chimney is done through ANSYS FLUENT 2022 R1. To obtain the optimal solution for implementing the numerical simulation in the ANSYS Fluent program, the following steps are followed:

- Create the numerical models using SolidWorks.
- Configure the numerical model in ANSYS Fluent and create a mesh appropriate for the model, ensuring that the software functions correctly.
- Select the appropriate solution by activating the energy equations, the turbulent flow (K-epsilon) model, and the radiation equation.
- Introduce the properties of the following materials: Air, Glass, Wood, Absorber panel (Aluminum), Insulation materials and Phase change materials.
- Define the boundary conditions of the model, which include ambient temperature, atmospheric pressure, and solar radiation, as well as the emissivity, absorptivity, and reflectivity of the materials.
- Initialize the flow procedure.
- Run the simulation and perform calculations.
- Save the results.

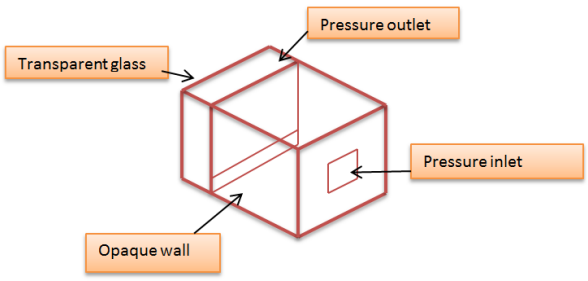
7. 1. Boundary Conditions Figure 7a-b shows the room connected to the solar chimney consists of

cubical wooden with dimensions (1×1×1) m^3 and the roof was covered with insulation. All walls of the model are opaque, except in the solar chimney has three semi-transparent wall made of glass. The solar chimney consisted from channel has a glass facade is used to allow the solar radiation to penetrate the glass surface of chimney to heat the aluminum absorber panel. In both cases of CFD simulation the type of boundary at the inlet and outlet of the model are pressure inlet and pressure outlet (atmospheric pressure) and no-slip condition between the fluid and wall. A standard k-ε model of turbulent flow was associated with the laws of the wall along solid boundaries is employed, where are depended on the Rayleigh number (Ra) > 10^8 and type of turbulent flow. Rayleigh number is a dimensionless number used in fluid dynamics and heat transfer to determine the occurrence of convection or buoyancy-driven flow in a fluid system. The solar radiation was simulated by the Discrete Ordinates (DO) model. The simple algorithm is use for coupling of continuity and pressure. The walls are non-adiabatic heat transfer by conduction, convection and radiation consideration.

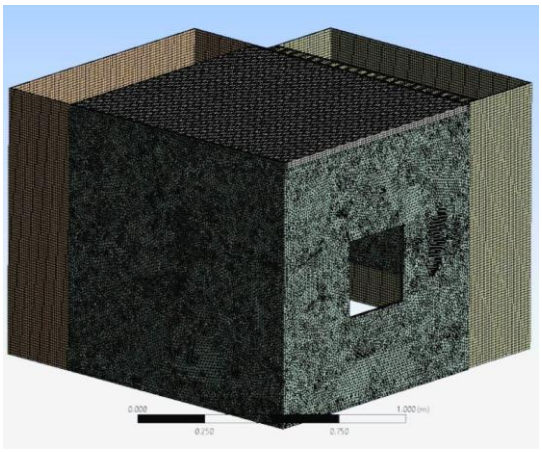
7. 2. MESH Independence Figure 7c shows the computational mesh. To examine the computational mesh resolution, five different meshes were created for the 3D model of the solar chimney, and the simulation was performed using ANSYS Fluent software, as shown



(a) Two vertical solar chimneys



(b) Boundary condition of model



(c) The computational mesh.

Figure 7. Experimental solar chimney studied model

in Table 7. Based on the CFD simulation results of the ventilation system, two parameters—room temperature and air changes per hour (ACH)—were considered to assess the mesh accuracy.The optimum mesh selected is (9929188) elements in model 4, through which can reach accurate results and short time in solution.

8. NUMERICAL RESULTS AND DISCUSSIONS

In this section we will discuss the numerical results on 3/7/2024 and analysis the numerical results of thermal performance of solar chimney ventilation (ACH, room average temperature, and hours of ventilation). The maximum ACH was noted around 4:00 p.m., which was equal (14.1), while the minimum ACH was approximately (2.8) from 2:00 a.m. to 6:00 a.m. The ACH values are shown in Figure 8 as a function of time. The average room temperatures reached the maximum values 45.2 ° C and 44 ° C at 2:00 p.m. and 4:00 p.m. respectively , while the minimum values were 33 ° C and 31 ° C at 4:00 a.m. and 6:00 a.m. respectively.

The temperatures reading according to test periods are in Figure 9. The ventilation took place between 8:00 a.m. and 12:00 a.m. on the next day depending on the values of the ACH, which change during these hours and then approximately stabilize at fixed values after

TABLE 7. Choosing the optimum mesh

No.	Elements	Nodes	Total ACH	Total T.Ave.Room
1	4378215	932559.795	81.6	471.2
2	7390644	1574207.172	92.73	471.6
3	8376886	1784276.718	95.46	470.87
4	9929188	2114917.044	96.82	470.54
5	12582098	2679986.874	97.79	471.13

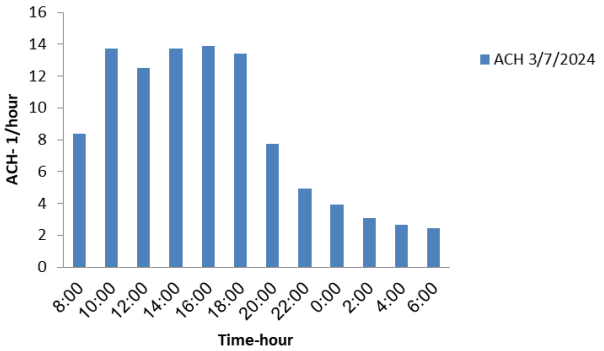


Figure 8. ACH values when using 2PCM

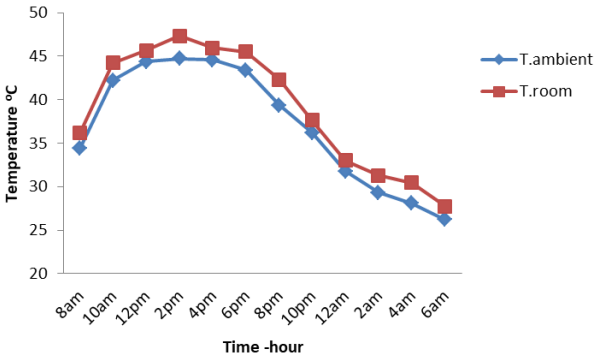


Figure 9. Room and ambient temperatures when using 2 PCM

12:00 a.m. To demonstrate the behavior of air velocity and temperature in this case through the day when using these two types of wax, Figures 10 and 11 illustrate the velocity and temperature contours at 12:00 p.m.

The process of melting and solidification of wax had an impact on the ventilation rate because the temperature of the wax affects the temperature of the absorbent wall. Figures 12 and 13 illustrate the liquid fraction at 12:00 p.m. and 12:00 a.m., respectively.

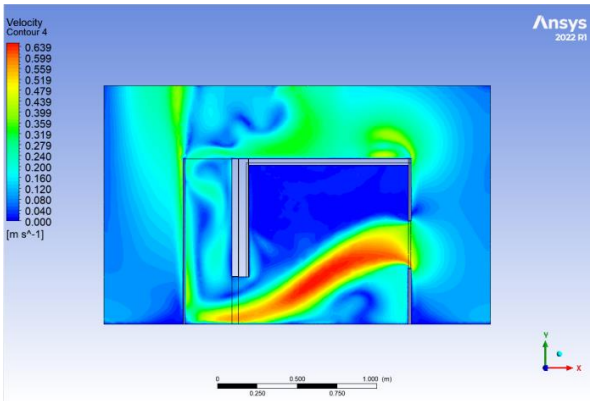


Figure 10. Velocity distribution when using 2PCM at 12:00 p.m.

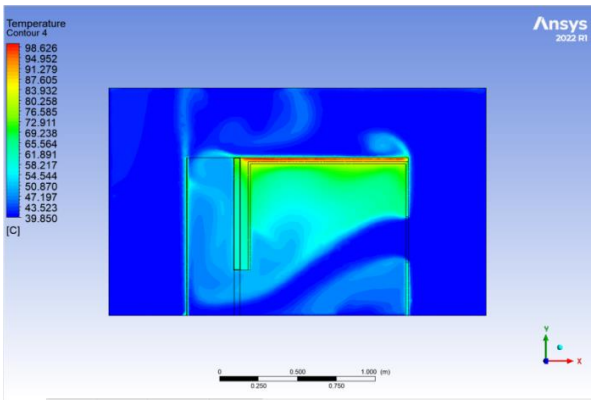


Figure 11. Temperature distribution when using 2PCM at 12:00 p.m.

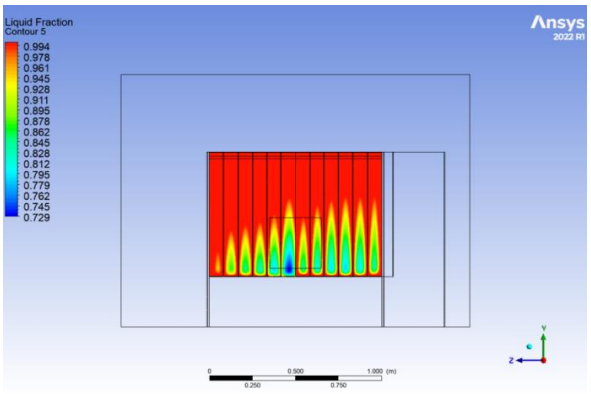


Figure 12. Liquid fraction when using 2PCM at 12:00 p.m.

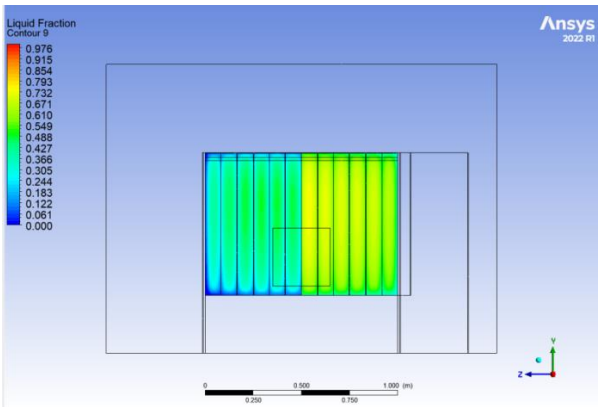


Figure 13. Liquid fraction when using 2PCM at 12:00 a.m.

Second case was on 7/7/2024 , the results showed that the best case added conductive mono nanoparticles Aluminum oxide to Al-Dura paraffin wax & Block-shaped paraffin wax with volume fraction (1.6%). The maximum ACH was noted around 4:00 p.m., which was equal (14.3) while the minimum ACH was approximately (2.9) from 4:00 a.m. to 6:00 a.m. The ACH values are shown in Figure 14 as a function of

time. The average room temperatures reached the maximum values 45 ° C and 44.7 ° C at 2:00 p.m. and 4:00 p.m. respectively , while the minimum values were 30.26 °C and 27.23 ° C at 4:00 a.m. and 6:00 a.m. respectively. The temperatures reading according to test periods are in Figure 15.

The ventilation took place between 8:00 a.m. and 2:00 a.m. on the next day depending on the values of the ACH, which changed during these hours and then approximately stabilized at fixed values after 2:00 a.m. To demonstrate the behavior of air velocity and temperature in this case through the day when using these two types of wax and NP volume fraction 1.6% , Figures 16 and 17 illustrate the velocity and temperature contours at 12:00 p.m.

The process of melting and solidification of wax had an impact on the ventilation rate because the temperature of the wax affects the temperature of the absorbent wall which affects the temperature of the air adjacent to the absorber wall. Figures 18 and 19 illustrate the liquid fraction at 12:00 p.m. and 12:00 a.m., respectively.

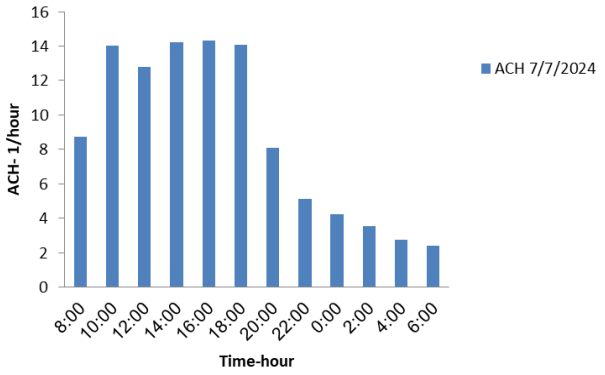


Figure 14. ACH values when using 2 PCM & NP volume fraction 1.6%

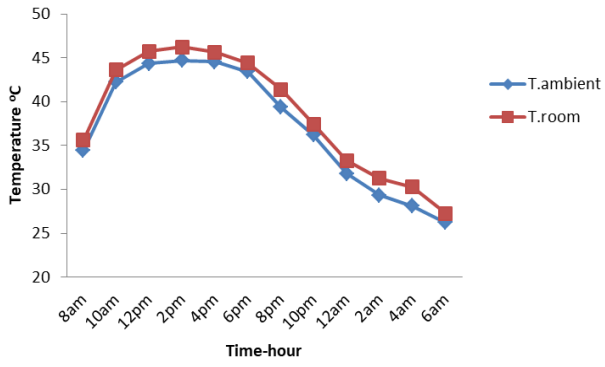


Figure 15. T.room and T. ambient when using 2 PCM & NP volume fraction 1.6%

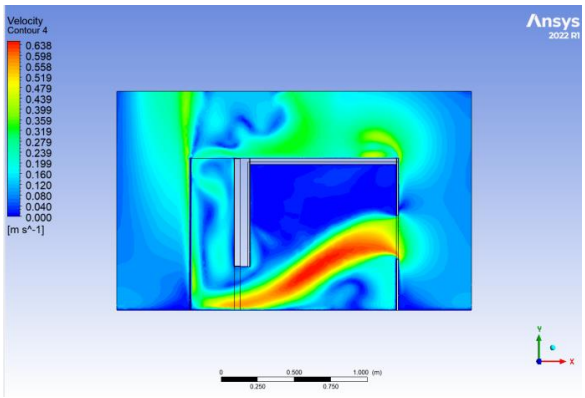


Figure 16. cVelocity distribution when using 2 PCM & NP volume fraction 1.6% at 12:00 p.m.

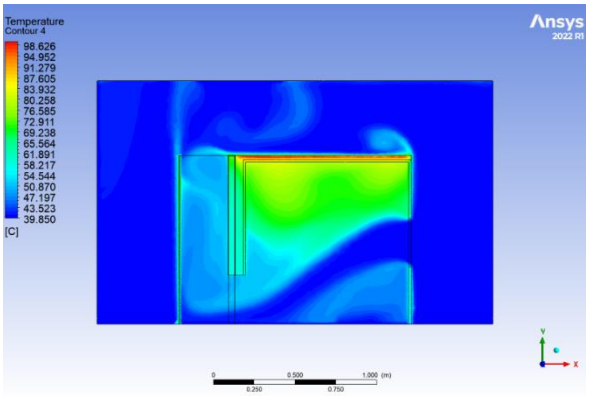


Figure 17. Temperature distribution when using 2 PCM & NP volume fraction 1.6% at 12:00 p.m.

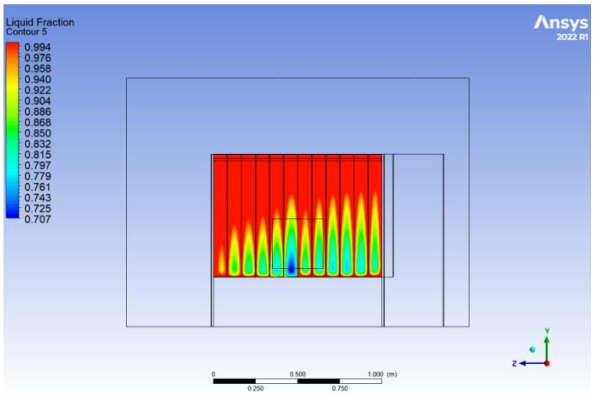


Figure 18. Liquid fraction when using 2 PCM & NP volume fraction 1.6% at 12:00 p.m.

8. 1. Comparison between First and Second Numerical Cases The numerical simulations conducted for both experimental cases offer critical insights into the thermal behavior, ventilation efficiency, and dynamic responses of the solar chimney-integrated systems under varying PCM configurations.

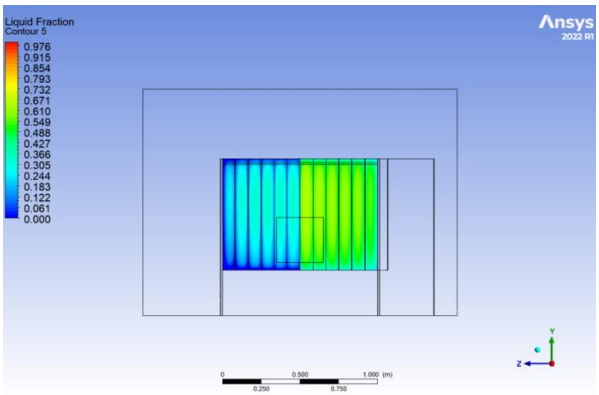


Figure 19. Liquid fraction when using 2 PCM & NP volume fraction 1.6% at 2:00 a.m.

In the first case (3/7/2024), which employed standard PCMs, the maximum ACH (Air Changes per Hour) reached 14.1 at 4:00 p.m., while the minimum dropped to 2.8 during the early morning hours (2:00–6:00 a.m.). The room temperature peaked at 46.4 °C at 2:00 p.m. and 4:00 p.m., and decreased to 33 °C and 31 °C at 4:00 a.m. and 6:00 a.m., respectively. The ventilation period extended from 8:00 a.m. to 12:00 a.m., based on ACH values, and showed a distinct correlation with solar radiation and ambient temperature fluctuations. Conversely, the second case (7/7/2024) incorporated 2PCMs enhanced with 1.6% aluminum oxide (Al_2O_3) nanoparticles, and the simulation demonstrated notable improvements in performance. The maximum ACH slightly increased to 14.3, with a more consistent and sustained ventilation pattern lasting from 8:00 a.m. to 2:00 a.m. The room temperature fluctuated between 27.23 °C and 44.7 °C, significantly lower than in the first case, especially during nighttime, confirming enhanced thermal regulation. This can be attributed to the improved thermal conductivity of the nano-enhanced PCM, which allows for faster and more effective heat absorption and dissipation. Visual analyses, including Figures 10 through 17, further validated these findings. The velocity and temperature distributions revealed stronger and more uniform airflow patterns in the second case, particularly around the absorber and exit zones of the solar chimney. The liquid fraction contours at 12:00 p.m. (Figure 12) and 12:00 a.m. (Figure 13) confirmed superior melting characteristics in the second case, with a wider and more active heat transfer zone. These improvements positively impacted the buoyancy-driven airflow, supporting higher ACH values and more effective cooling.

8. 2. Comparison between Numerical and Experimental Results A comparison between the numerical and experimental results reveals a close agreement in both air changes per hour (ACH) and average room temperature, with minor deviations attributed to physical factors such as heat losses from the solar chimney and the influence of wind speed. In the first case (3/7/2024), the experimental results showed a maximum ACH of 13.49 at 4:00 p.m., while the numerical simulation predicted a slightly higher value of 14.1 at the same time. The average deviation in ACH between the experimental and numerical results was calculated at 6.6%. Regarding the temperature, the experimental measurements recorded the maximum room temperatures of 46.4°C and 44.5°C at 2:00 p.m. and 4:00 p.m. respectively, while the numerical results showed corresponding values of 45.2°C and 44°C. The deviation in average room temperature between the two methods was 2.1%,as illustrated in Figures 20 and 22. In the second case (7/7/2024), where conductive mono nanoparticles (Al_2O_3) were added to the paraffin wax, the experimental ACH peaked at 14.3 at 4:00 p.m., while the numerical model reported a deviation of 5.9%. The experimental maximum room temperatures reached 45°C and 44.7°C at peak times, with minimum values of 30.26°C and 27.23°C recorded during early morning hours. The average temperature deviation between numerical and experimental results in this case was 3.1%,as seen in Figures 21 and 23. These results

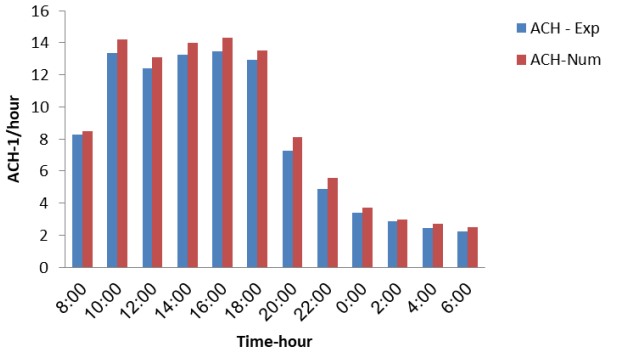


Figure 20. Comparison ACH for experimental and numerical results of case 1

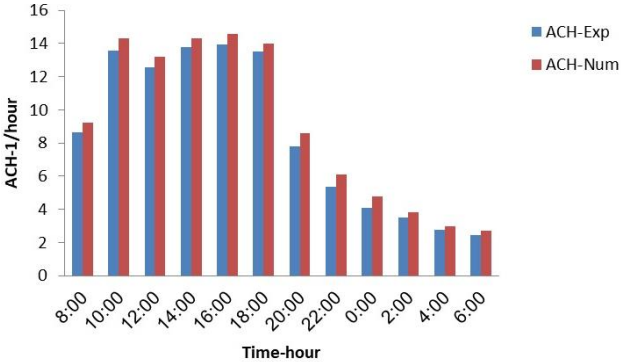


Figure 21. Comparison ACH for experimental and numerical results of case 2

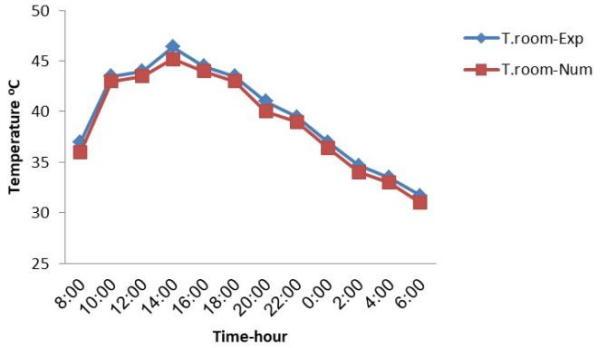


Figure 22. Comparison T.room for experimental and numerical results of case 1

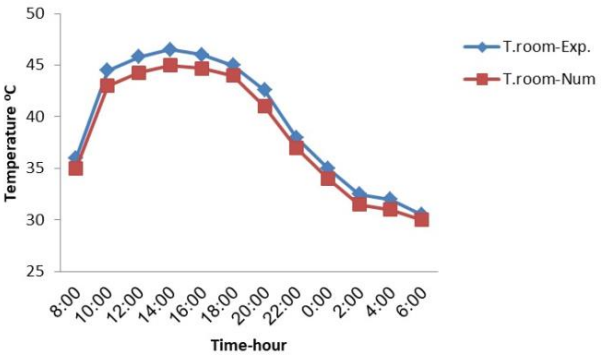


Figure 23. Comparison T.room for experimental and numerical results of case 2

confirm the reliability of the numerical model, as it successfully captured the trends and performance of the experimental setup with reasonable accuracy. The small discrepancies can be attributed to environmental conditions and assumptions within the simulation model.

9. CONCLUSIONS

This study presents a comprehensive investigation of the thermal performance of solar chimney systems integrated with phase change materials (PCMs) and nanoparticle-enhanced phase change materials under the climatic conditions of Al-Kut City, Iraq. The results from both experimental and numerical simulations show that combining conductive nanoparticles, such as aluminum oxide (Al_2O_3), with PCM-based paraffin significantly improves the thermal performance and ventilation of solar chimney systems. The key findings are as follows:

- The incorporation of PCMs significantly increases the air changes per hour (ACH) during daylight hours and extends ventilation time after sunset. Specifically, in the second configuration, which utilized nanoparticle-enhanced PCMs, there was an improvement in ACH by approximately 5.35% and a reduction in room temperature by 4.7°C compared to the first configuration. This highlights the effective role of PCMs in regulating indoor temperature and maintaining airflow after sunset.
- The addition of 1.6% aluminum oxide nanoparticles to paraffin wax improved heat absorption and release properties, resulting in a more uniform and continuous ventilation pattern. This change led to a slight increase in the maximum ACH from 14.1 to 14.3 and a reduction in internal temperature fluctuations, which were clearly reduced throughout the day and night.
- In the second configuration, which used nanoparticle-enhanced PCMs, better thermal regulation was observed. The maximum room temperature decreased by 2.4°C, and indoor temperatures during the night were significantly lower, enhancing the comfort of the indoor environment.
- The numerical simulations closely matched the experimental results, with a slight deviation of 6.6% for ACH and 2.1% for room temperature in the first case, and 5.9% for ACH and 3.1% for room temperature in the second case. This indicates the reliability of the numerical model in predicting the performance of solar chimney systems with various configurations.
- The results of this study indicate the potential of using nanoparticle-enhanced PCMs to improve the performance of passive solar ventilation systems, especially in hot regions such as Al-Kut. This innovation not only enhances the thermal performance of solar chimneys but also contributes to energy savings and improved indoor air quality, ultimately promoting sustainable building practices.

Future research should explore the long-term impacts of these systems under various climatic conditions, as well as the scalability of these systems for larger buildings and urban applications, expanding the use of these advanced materials as a crucial step in developing sustainable building technologies.

10. REFERENCES

1. Rezaie B, Reddy BV, Rosen MA. Exergy analysis of thermal energy storage in a district energy application. Renewable Energy. 2015;74:848-54. https://doi.org/10.1016/j.renene.2014.09.014
2. Sadafi M, Hosseini R, Safikhani H, Bagheri A, Mahmoodabadi M. Multi-objective optimization of solar thermal energy storage using hybrid of particle swarm optimization and multiple crossover and mutation operator. International Journal of Engineering Transactions B: Applications. 2011;24:366-76. https://doi.org/10.5829/idosi.ije.2011.24.04b.07

3. BAGHERI TH, Moradi M, BAGHERI TF. New technique for global solar radiation forecast using bees algorithm. International Journal of Engineering Transactions B: Applications. 2013;26:1385-92. https://doi.org/10.5829/idosi.ije.2013.26.11b.14
4. Ajay K, Kundan L. Performance evaluation of nanofluid (al2o3/h2o-c2h6o2) based parabolic solar collector using both experimental and cfd techniques. International Journal of Engineering-Transactions A: Basics. 2016;29(4):572-80. https://doi.org/10.5829/idosi.ije.2016.29.04a.17
5. Bassiouny R, Koura NS. An analytical and numerical study of solar chimney use for room natural ventilation. Energy and buildings. 2008;40(5):865-73. https://doi.org/10.1016/j.enbuild.2007.06.005
6. Bassiouny R, Korah NS. Effect of solar chimney inclination angle on space flow pattern and ventilation rate. Energy and Buildings. 2009;41(2):190-6. https://doi.org/10.1016/j.enbuild.2008.08.009
7. Karimipour-Fard P, Beheshti H. Performance enhancement and environmental impact analysis of a solar chimney power plant: Twenty-four-hour simulation in climate condition of isfahan province, iran. International Journal of Engineering Transactions B: Applications. 2017;30(8):1260-9. https://doi.org/10.5829/ije.2017.30.08b.20
8. Majeed MH, Ali A. Passive solar heating of a space. IOSR Journal of Mechanical and Civil Engineering. 2016;13(6):101-7. https://doi.org/10.9790/1684-130605101107
9. Jubear AJ, Ghareer AD. Numerical investigations on solar chimney inclination angle for room ventilation. International Journal of Engineering & Technology. 2018;7(4.17):1-9.
10. Kaneko Y, Sagara K, Yamanaka T, Kotani H, Sharma S, editors. Ventilation performance of solar chimney with built-in latent heat storage. Proceedings of 10th international conference of thermal energy conference (ECOSTOCK); 2006.
11. Amori KE, Mohammed SW. Experimental and numerical studies of solar chimney for natural ventilation in Iraq. Energy and Buildings. 2012;47:450-7. https://doi.org/10.1016/j.enbuild.2011.12.014
12. Imran AA, Jalil JM, Ahmed ST. Induced flow for ventilation and cooling by a solar chimney. Renewable energy. 2015;78:236-44. https://doi.org/10.1016/j.renene.2015.01.019
13. Chung LP, Ahmad MH, Ossen DR, Hamid M. Effective solar chimney cross section ventilation performance in Malaysia terraced house. Procedia-Social and Behavioral Sciences. 2015;179:276-89. https://doi.org/10.1016/j.sbspro.2015.02.431
14. Li Y, Liu S. Experimental study on thermal performance of a solar chimney combined with PCM. Applied Energy. 2014;114:172-8. https://doi.org/10.1016/j.apenergy.2013.09.022
15. Assari M, Basirat Tabrizi H, Parvar M, Alkasir Farhani M. Experimental investigation of sinusoidal tube in triplex-tube heat exchanger during charging and discharging processes using phase change materials. International Journal of Engineering Transactions A: Basics, . 2019;32(7):999-1009. https://doi.org/10.5829/IJE.2019.32.07A.13
16. Krishna J, Kishorea P, Solomonb AB, Sharmac VK. Experimental and Numerical Investigations on Al2O3–Tricosane Based Heat Pipe Thermal Energy Storage. International Journal of Engineering Transactions C: Aspects. 2018;31:980-5. https://doi.org/10.5829/IJE.2018.31.06C.16
17. Jadidi A, Jadidi M. An algorithm based on predicting the interface in phase change materials. International Journal of Engineering Transactions B: Applications. 2018;31(5):799-804. https://doi.org/10.5829/ije.2018.31.05b.15
18. Vandrangi S, Hassan S, Sharma K, Baheta A. Heat Transfer Coefficients Investigation for TiO2 Based Nanofluids. International Journal of Engineering, Transactions A: Basics. 2019;32:1491-6. https://doi.org/10.5829/ije.2019.32.10a.19
19. Tiji ME, Eisapour M, Yousefzadeh R, Azadian M, Talebizadehsardari P. A numerical study of a PCM-based passive solar chimney with a finned absorber. Journal of Building Engineering. 2020;32:101516. https://doi.org/10.1016/j.jobe.2020.101516
20. Selvana J, Manavalla S. An Innovative and Reliable Hybrid Cooling Method for Electric Vehicle Motors. International Journal of Engineering Transactions A: Basics 2024;37(07):1263. https://doi.org/10.5829/IJE.2024.37.07A.06
21. Ali MH, Mawlood MK, Jalal RE. Investigation of the performance of a newly designed solar chimney for enhancing natural ventilation and mitigation of summer overheating. Journal of Building Engineering. 2024;95:110310. https://doi.org/10.1016/j.jobe.2024.110310
22. Huang S, Li W, Lu J, Li Y, Wang Z, Zhu S. Experimental study on thermal performances of a solar chimney with and without PCM under different system inclination angles. Energy. 2024;290:130154. https://doi.org/10.1016/j.energy.2023.130154
23. Saleh MJ, Ahmed OK, Atallah FS. Comparison assessment for the effect of porous media and phase change material on the performance of photovoltaic solar chimney. Journal of Energy Storage. 2024;98:113158. https://doi.org/10.1016/j.est.2024.113158
24. Nia ES, Ghazikhani M. Enhancing reliability and efficiency of solar chimney by phase change material Integration: An Experimental study. Thermal Science and Engineering Progress. 2024;51:102600. https://doi.org/10.1016/j.tsep.2024.102600
25. Sornek K, Figaj R, Papis-Frączek K. Development and tests of the novel configuration of the solar chimney with sensible heat storage. Applied Thermal Engineering. 2025;258:124515. https://doi.org/10.1016/j.applthermaleng.2024.124515
26. Hassan HZ. Performance enhancement of the basic solar chimney power plant integrated with an adsorption cooling system with heat recovery from the condenser. Energies. 2023;17(1):136. https://doi.org/10.3390/en17010136
27. Qasim HH, Alshara AK, Abood FA. Simulation and design of a solar chimney integrated with phase change material layer for building ventilation. International Journal of Thermofluids. 2024;24:100853. https://doi.org/10.1016/j.ijft.2024.100853
28. ALubaidi AB, Abas FO, Abass RA, Abdulredhi NJ. Thermal Behaviour of Paraffin Wax/Poly Vinyl Alcohol Composite Material. Al-Khwarizmi Engineering Journal. 2016;12(2):18-26.
29. Li C, Li Q, Ge R. Fabrication and investigation of paraffin based shape-stable composite phase change materials with styrene ethylene butylene styrene as supporting skeleton for low temperature thermal energy storage and management. Journal of Energy Storage. 2023;68:107802. https://doi.org/10.1016/j.est.2023.107802
30. Alawi OA, Sidik NAC, Xian HW, Kean TH, Kazi SN. Thermal conductivity and viscosity models of metallic oxides nanofluids. International Journal of Heat and Mass Transfer. 2018;116:1314-25. https://doi.org/10.1016/j.ijheatmasstransfer.2017.09.133
31. Daood IY. Effects of nano-fluids types, volume fraction of nano-particles, and aspect ratios on natural convection heat transfer in right-angle triangular enclosure. Eng And Tech Journal. 2010;28(16):5365-88.

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

یک تحقیق تجربی و عددی دقیق از مدل ساخته شده دودکش خورشیدی (SC) و عملکرد حرارتی آن ارائه شده است که تحت شرایط اقلیمی شهر ال-کوت، عراق انجام شده است. این مطالعه بر روی عملکرد حرارتی دو دودکش خورشیدی عمودی با جهت‌گیری به سمت جنوب و غرب تمرکز دارد. پیکربندی اول شامل یک ساختمان با دو دودکش خورشیدی عمودی است که دو نوع ماده تغییر فاز (PCM) را در بر می‌گیرد: واکس پارافین آل-دورا و واکس پارافین بلوک‌شکل (PCM1) و (PCM2)، به ترتیب. این مواد PCM در همان ظرف آلومینیومی قرار داده شده‌اند، که توسط partitions از یکدیگر جدا شده‌اند و با فین‌ها تجهیز شده‌اند تا سطح تماس حرارتی را افزایش دهند. در پیکربندی دوم، نانوذرات تک‌عنصری (NP)، به‌ویژه اکسید آلومینیوم (Al_2O_3)، به هر دو ماده PCM1 و PCM2 در یک حجم تقریبی ۱/۶٪ اضافه شده‌اند تا ویژگی‌های حرارتی پارافین و عملکرد ذخیره‌سازی گرما را بهبود بخشند. این مطالعه نقش واکس پارافین را به‌عنوان یک ماده ذخیره‌سازی انرژی همراه با نانوذرات بررسی می‌کند، به‌طوری که PCM انرژی خورشیدی حرارتی را در طول روز ذخیره کرده و در شب از طریق فرآیند تغییر فاز آن را آزاد می‌کند. نتایج عددی نشان می‌دهند که مواد PCM تأثیر قابل توجهی بر میزان تغییر هوای در ساعت (ACH) در طول روز دارند و ساعات تهویه پس از غروب آفتاب را تمدید می‌کنند. پیکربندی دوم نشان داد که ACH حدود ۵/۳۵٪ بهبود یافته و دمای اتاق به میزان ۴/۷ درجه سانتی‌گراد کاهش یافته است در مقایسه با مورد اول. نتایج تجربی نشان داد که ACH به میزان ۵.۲٪ بهبود یافته و دمای اتاق ۳/۹ درجه سانتی‌گراد کاهش یافته است. نتایج تجربی و عددی هر دو تأیید می‌کنند که ادغام نانو سیالات با PCM به طور قابل توجهی کارآیی کلی دودکش خورشیدی را بهبود می‌بخشد. نوآوری این مطالعه در ادغام نوآورانه مواد تغییر فاز با نانوذرات هدایت‌کننده است که نه تنها عملکرد حرارتی را بهبود می‌بخشد بلکه به‌طور چشمگیری دوره‌های تهویه را تمدید می‌کند، به‌ویژه تحت شرایط تابش خورشیدی شدید که در شهر ال-کوت، عراق معمول است. در نهایت، به‌دلیل تابش خورشیدی شدید در ال-کوت، نرخ‌های تهویه غیرفعال که توسط دودکش خورشیدی به‌دست آمد، بسیار رضایت‌بخش گزارش شد.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1077-1087

International Journal of Engineering

Journal Homepage: www.ije.ir

Mathematical Modeling of a Hydro-mechanical System with a Positive Displacement Motor and Hydraulic Thruster to Optimize the Drilling Process for Extended-reach Drilling Wells

G. V. Buslaev[a], A. V. Konoplyannikov*[b]

[a] *Laboratory of Thermodynamic, Gas Chemical and Energy Processes of Oil and Gas Production of the "Arctic" Research Center, Empress Catherine II Saint Petersburg Mining University, Russia*

[b] *Department of Well Drilling, Empress Catherine II Saint Petersburg Mining University, Russia*

PAPER INFO

Paper history:
Received 18 April 2025
Received in revised form 27 May 2025
Accepted 28 June 2025

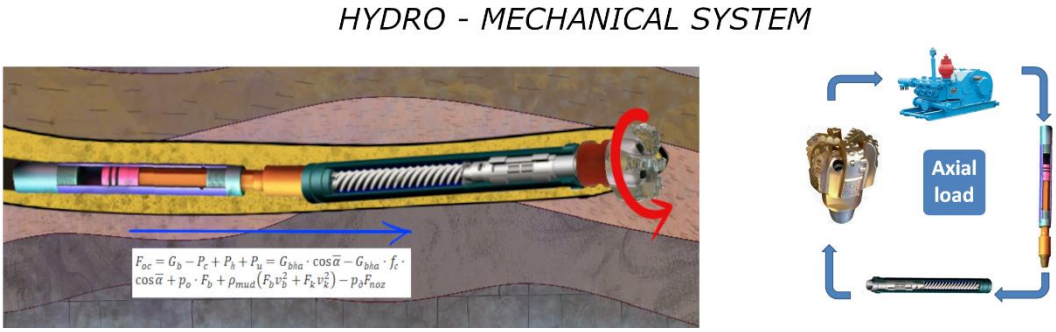
Keywords:
Drilling
Hydraulic Thruster
Positive Displacement Motor
Mathematical Model
Axial Force
Extended-Reach Drilling Well
Drilling Optimization

ABSTRACT

This study addresses key challenges in Extended-Reach Drilling operations, focusing on limited axial load transfer to the bit and increased hydraulic pressure losses. These issues are critical as they directly impact drilling efficiency and well construction capabilities. A technological solution involving the integration of a hydraulic thruster into the bottomhole assembly was proposed and investigated. A comprehensive mathematical model of the hydromechanical system comprising the bit, Positive Displacement Motor (PDM), hydraulic thruster and mud pump was developed. An iterative computational algorithm in Python was created for equipment selection and drilling regime optimization. The novelty of the work lies in the development of a comprehensive modeling approach that integrates the operation of a hydraulic thruster with a PDM within the bottom hole assembly and mud pump. This research provides a scientifically grounded methodology for designing bottomhole assembly (BHA) with integrated hydraulic thrusters, filling a gap in existing literature. The study confirms the effectiveness of hydraulic thruster applications in drilling deep, deviated, and horizontal wells.

***doi**: 10.5829/ije.2026.39.05b.03*

Graphical Abstract



1. INTRODUCTION

In the early stages of oil and gas production, drilling was limited exclusively to vertical wells. The primary objective of drilling operations was to reach target depths and expose hydrocarbon-bearing formations. Precise spatial positioning of the wellbore trajectory was not considered a significant concern (1).

Advances in drilling technology enabled the construction of deviated directional wells. The ability to

*Corresponding Author Email: konoplyannikovaleksandr@gmail.com (A.V. Konoplyannikov)

Please cite this article as: Buslaev GV, Konoplyannikov AV. Mathematical Modeling of a Hydro-mechanical System with a Positive Displacement Motor and Hydraulic Thruster to Optimize the Drilling Process for Extended-reach Drilling Wells. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1077-87.

drill complex well profiles—with multiple angle build-up, drop-off, and stabilization sections—emerged thanks to downhole surveying tools capable of taking measurements while drilling (MWD) (2). This process requires precise tools, real-time data acquisition, and control systems to monitor and adjust the well trajectory promptly. The development of directional drilling technology has set records for extended horizontal sections, making it possible to construct wells exceeding 15,000 meters in drilled depth, both onshore and offshore (3).

Although Extended-Reach Drilling (ERD) offers numerous operational advantages over conventional directional wells (4), it presents several technical challenges during well construction, including poor cuttings removal, insufficient weight transfer to the bit, increased hydraulic resistance and torque, and elevated equivalent circulating density (ECD) (5).

The PDM have been widely used for drilling directional and deep wells since the mid-century. The PDM design incorporates a “rotor-stator” pair. The rotor has a helical screw-shaped geometry. The stator’s inner surface is coated with a vulcanized elastomer that matches the rotor’s geometry and transmits reactive torque to the rotor through the flow of drilling mud.

PDMs are suitable for high-rate drilling with various bit types, including roller-cone, diamond, and PDC bits. However, PDMs are sensitive to hot, abrasive, and corrosive environments (6) due to the elastomer lining of the stator.

Drilling extended deviated and horizontal wells poses challenges related to the inability to maintain consistent bit weight and instability caused by oscillations and poor wellbore trajectory control. Modern engineering solutions include: (a) devices generating induced pulses or resonators that create vibrations in the BHA system (7, 8); and (b) BHA components combined with technological measures to enhance well trajectory accuracy (9).

To improve drilling quality in wells with extended horizontal sections, it is proposed to incorporate Rotary Steerable Systems (RSS) into the BHA (10, 11). The RSS may be used with or without a PDM section (12). Although RSS is used to drill most ERD wells, several wells with measured depths exceeding 11,000 meters have been successfully drilled using PDMs (13).

It should be noted that directional and horizontal well drilling with PDMs involves alternating between rotary mode (with drillstring rotation) and sliding mode (without rotation) to maintain the planned well trajectory and correct local deviations caused by rock anisotropy (14, 15). However, as the horizontal section length increases, substantial frictional resistance develops due to contact between the non-rotating drillstring and the wellbore wall (16-18).

At well depths and curvatures where frictional losses and drag forces approach the axial load exerted by the BHA on the bit (19), transferring weight from the surface becomes inefficient or even impossible (20).

In such cases, downhole bit wieght control systems integrated into the BHA offer a significant advantage, eliminating the need for weight transfer through the drillstring (21). This approach allows for shortening the BHA by reducing the number of heavy-weight drill pipes (HWDP), which in turn positively impacts the equivalent circulating density (ECD).

ECD can be controlled by regulating the drilling mud flow rate in the drillpipe annulus and minimizing tool joint friction losses. However, this approach complicates cuttings transport and can cause ECD fluctuations, potentially resulting in lost circulation and wellbore breathing (22), while also negatively impacting the near-wellbore formation (23-25). Cuttings accumulation on the low side of the wellbore reduces the effective flow area, increasing pressure differentials. Pressure losses in the annular space have been observed to rise with increasing drillstring rotation speeds (26, 27). Notably, pressure differentials within the drillstring assembly have little effect on annular ECD (28, 29).

The technologies employed in ERD well drilling—such as lubricant additives in drilling fluids, RSS, and oscillators—cannot fully eliminate drag forces. Consequently, as the wellbore deviation from vertical increases, there will always be a limitation in delivering sufficient weight on bit (WOB) for effective rock destruction.

This paper proposes addressing the aforementioned challenges through integration of a hydraulic thruster into the BHA. By optimizing mud pump flow rates, selecting appropriate PDM specifications, and adjusting bit nozzle diameters, significant improvements in bit performance can be achieved. Additionally, separating BHA components using a hydraulic thruster helps to reduce bit-induced vibrations (30).

Analysis of bit dynamics under various drilling methods has demonstrated that vibrations in the BHA and drillstring originate from tool rotation and its interaction with both formation rock and wellbore walls. These vibrations manifest as axial, lateral, and torsional oscillations, which arise under certain drilling conditions—such as excessive bit weight or specific rotational speeds (31). These effects can lead to tool instability, bit balling of PDC cutters, and other performance-degrading phenomena (32-34).

To optimize PDC bit performance and mitigate vibrations, the scientific literature extensively explores cutting structure control systems and methods for determining weight on bit (WOB). These developments aim to regulate the stress-dynamic environment in both the drillstring and BHA, thereby improving overall drilling efficiency (35).

Researchers from Freiberg University of Mining and Technology developed a hydraulic thruster with a diaphragm and choke (Shockspear) featuring a replaceable sleeve for pressure regulation. The methodology for calculating its parameters was successfully applied to drilling Emmen-9 well in a fractured carbonate reservoir, achieving a 4400 m borehole with 2000 m deviation.

In 2006, researchers at Ukhta State Technical University, Buslaev and Belkin (36) developed a multifunctional downhole feed system (DFS) that combines the functions of a shock absorber, oscillation compensator, and hydraulic thruster. Prototype testing confirmed reduced vibration levels and improved bit weight transfer.

Field trials of the "ZUPD-195" in hard rock formations at depths greater than 3000 m increased bit run length from 7.5 m to 29.2 m, while reducing bit wear and non-productive time (NPT) by a factor of four. The drillstring weight was reduced from 300 kN to 100 kN, resulting in a time saving of 49.42 hours. Although further testing is needed for comprehensive evaluation, the results confirm the system's effectiveness (36).

Research was patented a hydraulic thruster connected to a PDM via a hollow rod shaped as a multi-start helical rotor. The rotor engages with a helically profiled stator mounted on the rotor, both sharing identical planar and spatial geometry.

The effectiveness of hydraulic thrusters was confirmed by 2020 field trials in China, which showed a 50% increase in rate of penetration (ROP) compared to offset wells and data reported by Zhao and Cui (37).

The use of hydraulic thrusters—downhole feed systems (DFS) integrated into the BHA—addresses several challenges encountered during ERD well construction. As these downhole tools convert hydraulic energy into axial force, optimizing drilling mud flow rate and bit nozzle diameter enables a significant increase in weight-on-bit (WOB) without the need to lengthen the BHA for mass-based loading.

In "pressure relief" mode, the hydraulic thruster isolates the lower BHA from the rest of the drillstring, maintaining a constant, controlled WOB while reducing axial vibrations and shock loads. The lack of a rigid connection with the upper BHA minimizes transmission of vibrational loads. Additionally, the DFS helps prevent stuck pipe incidents by generating vertical impacts and inducing rapid pressure surges in the system.

DFS systems are compatible with both rotary steerable systems (RSS) and positive displacement motors (PDM).

Table 1 summarized the analysis hydraulic thruster components which are well-studied and field tests confirm their operational reliability, there is a lack of methodological frameworks for modeling their interaction with other BHA elements. The absence of scientifically grounded design principles for BHAs with integrated hydraulic thrusters remains a key limitation to the widespread adoption of these systems.

This paper presents a study focused on integrating the operation of a hydraulic thruster and a positive displacement motor (PDM) within the BHA.

2. METHODOLOGY

When using a hydraulic thruster in the well drilling process, the principle of drilling control differs from conventional drilling technology. The key distinction lies in the ability to regulate the bit weight by adjusting the flow rate of the drilling mud. This imposes specific constraints on the operational regime of the downhole motor (PDM).

The study introduces a boundary condition in which the conventional drill string is modeled as a suspended assembly, divided by a hydroulic thruster device into two distinct sections:

The upper section of the drill string, suspended from the hoisting system during drilling, features a free lower end equipped with the device housing. To increase applied load via gravitational force contribution, the housing may alternatively rest on the Lower BHA (LBHA). LBHA mobile drill section, with its bottom end (bit) in contact with the borehole bottom during drilling.

TABLE 1. Testing Results of Hydraulic Thruster Devices

Technology	Formations	Drilling Interval (m)	Effect
"ZUPD -195"+rotory (36)	Severniy Ugid (Russia)	3321 - 3250	increasing the drill bit footage and reducing the number HWDP in the BHA.
"Hydra-Thrust" + PMD (38)	Mahogany (Gulf of Mexico)	120 - 4265	average improvement of penetration rates of 35-5% while both sliding and rotating respectively.
"Thruster Device" + PMD (39)	Emmen-9 (Netherland)	4212 -4505	average increased ROP from 6.7 m/h to 10 m/h
"165mm hydraulic thruster" + PMD (37)	Tao-2 (China)	2146 - 2600	the average mechanical drilling rate wasincreased by about 50%; the footage of a single bit has been increased from 251 meters to 454 meters.

The opposing end incorporates a telescoping shaft capable of axial displacement within the device stroke length along the wellbore axis.

The development of a computational mathematical model for the hydro-mechanical system "bit – downhole motor – hydraulic thruster – mud pump" to determine the final values of axial WOB, rotary speed, and torque on the bit is a relevant research objective.

In Figure 1 A, C – in vertical wells; B, D – in horizontal wells; A, C – with manual load application; B, D – with automated load application; 1 – Bit; 2 – Stabilizer; 3 – Downhole motor (PDM); 4 – Centralizer; 5 – Heavy-weight drill pipe (HWDP); 6 – Drill pipes; 7 – Kelly; 8– Extendable shaft of the hydraulic thruster; 9 – Hydraulic thruster body.

In general form, accounting for all factors except torque and drill string dynamics. This simplification is justified by the altered operational mode of the drill string and the capability to compensate for axial oscillations through displacement of the lower mobile drill section (LBHA) relative to the drill pipe column. This approach effectively eliminates the influence of dynamic processes in the drill string on the axial component of bit load. The axial weight on bit (WOB) is given by (kN):

$$P_{oc} = G_b - P_c + P_h + P_u \quad (1)$$

G_b – axial component of the buoyed weight of the bottomhole assembly (BHA), including heavy-weight drill pipes (HWDP), extendable shaft of the hydraulic thruster, downhole motor (PDM), bit, stabilizers, and other elements in the drilling mud (kN):

$$G_b = G_{bha} \cdot cos\,\overline{\alpha}, \quad (2)$$

G_{bha} – weight of the BHA section used to apply WOB in conventional drilling or the buoyed weight of the movable BHA when using a downhole bit feed device (kN).

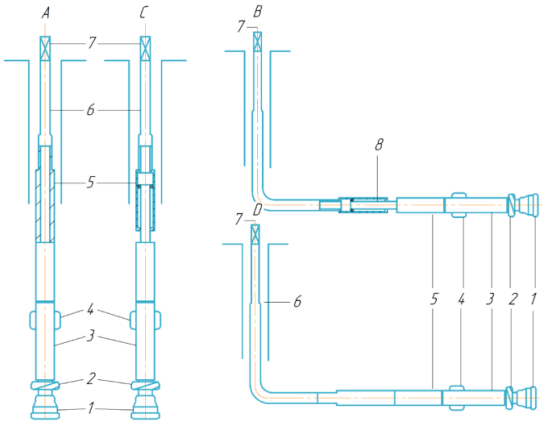


Figure 1. Schematic of axial WOB generation in horizontal wells

$\overline{\alpha}$ – average zenith angle over the BHA section (°).

P_c – resistance force opposing WOB transfer to the bit (kN):

$$P_c = f_c \cdot N = f_c \cdot G_{bha} \cdot sin\,\overline{\alpha}, \quad (3)$$

N – wellbore wall reaction force to the perpendicular component of the BHA weight (kN).

f_c – generalized coefficient of resistance forces, accounting for friction, tool joint wall contact, differential sticking, and other factors.

P_h – hydraulic component of WOB. Excluded in conventional drilling (as it induces internal tension) but included for downhole feed devices, where it becomes the primary WOB component (kN).

F_{noz} – total flow area of bit nozzles (m²).

$$P_h = p_o \cdot F_b - p_{bit}F_{noz} - p_{pdm}F_b \quad (4)$$

p_o – pressure differential between the extendable shaft cavity and annular pressure at the piston level (MPa):

$$p_o = p_{bit} + p_{pdm} + p_y + p_{ht} + p_{an} \quad (5)$$

$p_{bit}, p_{pdm}, p_y, p_{ht}, p_{an}$– hydraulic losses in the bit, downhole motor (PDM), downhole feed device, pipes, and annular space below the device (MPa).

F_b– inner cross-sectional area of the PDM body (or HWDP in rotary drilling); for feed devices, this equals the piston area (m²).

P_u– inertial force component due to drilling mud momentum (kN):

$$P_u = \rho_{mud}(F_b v_b^2 + F_k v_k^2), \quad (6)$$

ρ_{mud} – drilling mud density (kg/m³); v_b, v_k – fluid velocity inside the drill string and annular space, respectively (m/s); F_b, F_k – cross-sectional area of the pipe inner diameter and annular space, respectively (m²).

2. 1. Determination of Bit Torque Under Given WOB and Rotary Speed The torque at the bit depends on multiple factors, the primary ones being the bit type, cutting structure aggressiveness, and rotary speed. According to Plotnikov V.M.:

$$M_{bit} = m \cdot G \quad (7)$$

M_{bit}– Torque at the bit (N·m); m – Specific torque per unit WOB (min⁻¹); G – axial weight on bit (WOB) (N). The specific torque can be estimated using a simplified empirical formula proposed by Simonyants:

$$m = a_{dem}\left[\frac{28}{n} + 0{,}14\right]D_{bit}^2 \quad (8)$$

a_{dem}– Dimensionless coefficient dependent on bit type; n – Bit rotary speed (rpm); D_{bit} – Bit diameter (m).

For research purposes, the dimensionless coefficient can be approximated within the following ranges for different bit types and formations:

2. 2. Determination of Pressure Drop in a Positive Displacement Motor (PDM) **A**n analysis of bench test results established that the pressure loss (p_{pdm}) in a PDM depends on the bit torque (m), average drilling mud flow velocity between the rotor and stator, PDM displacement volume (V_{pdm}), drilling mud density, and radial interference. The relationship is expressed as follows:

$$p_{pdm} = \frac{\rho Q_{pdm}^2}{f^2}\left[a_1 + b_1\left(\frac{mf^2}{Q^2 V_{pdm}\rho}\right)^{c_4}\right] \quad (9)$$

Q_{pdm} – drilling mud flow rate through the PDM (m³/s); ρ – drilling mud density (kg/m³); f – Cross-sectional flow area between the rotor and stator (m²); V_{pdm} – PDM displacement volume (m³); m – Output shaft torque (N·m).

a_p, b_p, c_p – Empirical coefficients for pressure drop calculation:

$$a_p = 5{,}689 + 0{,}831103 \cdot 10^3 \cdot \frac{\delta}{V_{\text{pdm}}^{1/3}} + 0{,}262323 \cdot 10^6 \cdot \left(\frac{\delta}{V_{\text{pdm}}^{1/3}}\right)^2 \quad (10)$$

$$b_p = 9{,}02793 - 0{,}280917 \cdot 10^3 \cdot \frac{\delta}{V_{\text{pdm}}^{1/3}} + 0{,}028979 \cdot 10^6 \cdot \left(\frac{\delta}{V_{\text{pdm}}^{1/3}}\right)^2 \quad (11)$$

$$c_p = 1{,}1958 + 0{,}042242 \cdot 10^3 \cdot \frac{\delta}{V_{\text{pdm}}^{1/3}} - 0{,}0043379 \cdot 10^6 \cdot \left(\frac{\delta}{V_{\text{pdm}}^{1/3}}\right)^2 \quad (12)$$

Here, δ represents the radial interference (positive values) or clearance (negative values), which depends on the output shaft torque and the wear condition of the helical rotor-stator pair (m).

2. 3. Determination of PDM Rotary Speed The rotational speed of the PDM output shaft depends on the bit torque (m), similar to how the pressure drop $p_{\text{ВЗД}} = f(m)$ characterizes PDM performance. The shaft speed is influenced by drilling mud flow rate, density, displacement volume, and radial interference:

$$n = \frac{Q_{pdm}}{V_{pdm}^{1/3} f}\left[a_n + b_n\left(\frac{mf^2}{Q^2 V_{pdm}\rho}\right)^{c_4}\right] \quad (13)$$

TABLE 2. Dependence of Coefficient a_{dem} on Bit Type and Rock Formation

Rock Formation	Roller-cone bit	Diamond bit	PDC bit
Soft to medium	0,9…1,0	-	3,4…3,5
Medium to hard	0,7…0,8	-	3,0…3,3
Hard to abrasive	0,5…0,6	1,8…2,5	-

TABLE 3. Example of Calculated Pressure Drops for Different Radial Interference Values

M, *N·m*	Radial interference/clearance δ, *m*				
	-0,6·10⁻³	**-0,3·10⁻³**	**0**	**0,3·10⁻³**	**0,6·10⁻³**
0	0,82	0,85	0,90	1,10	1,50
500	1,16	1,18	1,20	1,37	1,77
1000	1,54	1,50	1,50	1,72	2,15
1500	1,88	1,90	1,83	2,05	2,46
2000	2,25	2,26	2,23	2,40	2,86
2500	2,69	2,68	2,68	2,84	3,40
3000	3,10	3,06	3,12	3,33	3,63
3500	3,50	3,55	3,62	3,88	4,12
4000	4,00	3,90	4,15	4,40	4,65
4500	-	4,40	4,68	5,00	5,15
5000	-	4,95	5,30	5,60	5,80
5500	-	-	-	-	6,50

a_n, b_n, c_n – empirical coefficients for shaft speed calculation:

$$a_n = 3{,}402 \cdot 10^{-2} + 0{,}000749 \cdot 10^3 \cdot \frac{\delta}{V_{pdm}^{1/3}} - 0{,}000149 \cdot 10^6 \cdot \left(\frac{\delta}{V_{pdm}^{1/3}}\right)^2 \quad (14)$$

$$b_n = -0{,}89119 \cdot 10^{-2} + 1{,}53166 \cdot \frac{\delta}{V_{pdm}^{1/3}} + 0{,}01911751 \cdot 10^4 \cdot \left(\frac{\delta}{V_{pdm}^{1/3}}\right)^2 \quad (15)$$

$$c_n = 1{,}32618 + 0{,}0756845 \cdot 10^3 \cdot \frac{\delta}{V_{pdm}^{1/3}} + 0{,}013533 \cdot 10^6 \cdot \left(\frac{\delta}{V_{pdm}^{1/3}}\right)^2 \quad (16)$$

2. 4. Algorithm of the Mathematical Model for the Drilling System When drilling with a hydraulic thruster, the axial weight on bit (WOB) is functionally dependent on multiple parameters. The general interaction scheme of the hydro-mechanical system components is shown in Figure 2.

The interactions between system elements create complex functional dependencies.

nitial Conditions of the Hydro-mechanical System:

PDM₀: $p_{pdm} = p_{0pdm}$

$n_{pdm} = n_{0pdm}$

Hydraulic Thruster₀: $F_{oc} = (p_{bit} + p_{pdm} + p_{an}) * S$

Drill Bit₀: $p_{bit} = f(Q)$

$m_0 = f(F_{oc}, n)$

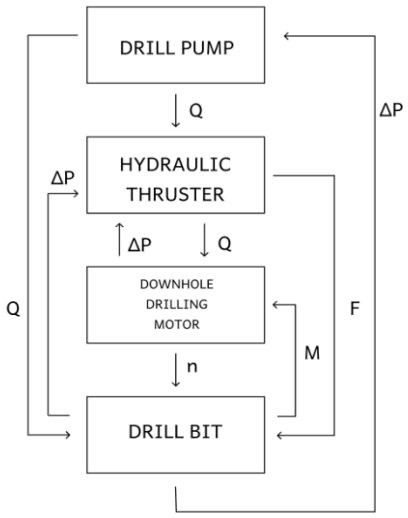


Figure 2. Schematic of the mathematical model

IThe pump flow rate (Q) is selected to ensure the required axial load, enhance wellbore cleaning efficiency, and maintain continuous operation of the PDM. Pressure losses in the drillstring and annular space (p_{an}) were calculated using a viscoplastic fluid model. The Shvedov–Bingham model was used to describe the behavior of the viscoplastic drilling fluid.

The system efficiency coefficient is defined as the ratio of useful hydraulic energy to mechanical energy, expressed as a percentage (40). The Shvedov – Bingham model is widely used to describe the rheological behavior of drilling fluids. The model assumes that the fluid begins to flow only after overcoming a critical shear stress (yield point, τ_0). The model is described by two parameters: plastic viscosity (μ_p) – resistance to flow after the onset of motion; yield point (τ_0) – the minimum stress required to initiate flow. The model accurately predicts hydraulic losses in annular spaces and pipes compared to Newtonian models. Despite its advantages, the model does not account for: thixotropy (time-dependent viscosity changes); complex effects at high shear rates (e.g., turbulence).

System performance analysis identified three key interdependent parameters that govern overall productivity. A critical feature of the system is the initial pressure drop of the drilling fluid in the hydraulic thruster, which generates additional axial load at the bit.

The most sensitive component to axial load fluctuations is the PDC bit, which remains in direct contact with the bottomhole Due to the bit's shearing–crushing action, increased axial load results in deeper cutter penetration into the rock. This significantly complicates the mechanical rock-breaking process and leads to a considerable increase in reactive torque at the bit.

The rising reactive torque is transferred to the motor shaft, hindering its rotation and increasing pressure drop within the motor's working chamber. This elevated system pressure drop subsequently impacts the hydraulic thruster, further increasing axial load at the bit. As a result, the system enters a self-sustaining feedback loop.

It is important to note that this sequence of system interactions forms a continuous feedback chain, in which any change in one parameter inevitably requires adjustments to other operational characteristics of the system.

3. RESULTS CALCULATION

The following initial data were used for the calculations (see Table 4). The computational model was implemented in Python using an iterative algorithm based on the mathematical formulations described in the Methodology section. The iterative algorithm analyzes key parameters: bit torque, PDM spindle speed, and axial load generated by the hydraulic thruster.

As an example, the iterative calculation results for Q = 0.028 m³/s are presented Table 5. The iteration process was deemed convergent upon satisfying the following criterion: the relative change in the load applied by the hydraulic actuator and the rotational speed of the downhole motor. The convergence threshold was set at 1e-3, with a maximum iteration count of 100.

The relationship between rotational speed and applied torque is attributed to seal failure in helical channels due to working fluid leakage. This phenomenon causes speed variations depending on torque magnitude. The system parameters stabilize, reaching a steady-state operating regime.

The study of bit load dependence on flow rate is crucial for analyzing the hydraulic thruster's performance. Using the axial load graph, we examine the system's initial and equilibrium states obtained from simulations within the flow rate range Q = 0.042 – 0.001 m³/s.

The dependence of rotational speed on applied torque is caused by loss of sealing in the helical channels due to working fluid leakage. This phenomenon leads to speed variation as a function of torque magnitude. The system parameters reach equilibrium, resulting in stable operating conditions.

The investigation of bit load dependence on drilling mud flow rate represents a critical stage in analyzing the hydraulic thruster's performance. Using the axial load graph Figure 3, we examine the system's initial and equilibrium states obtained through modeling within the flow rate range Q = 0.042-0.001 m³/s:

The red curve indicates initial load at bit-bottomhole contact, while the blue curve shows final load after hydro-mechanical system stabilization.

TABLE 4. Initial Calculation Data

Input Parameter	Value	Input Parameter	Value
Bit Parameters		***Drill Pipe Parametrs***	
Dimensionless bit-type coefficient	3.0	Drill pipe length, m	12
Bit diameter, m	0.1905	Outer diameter of drill pipe, m	0.127
Number of nozzle jets	6	Inner diameter of drill pipe, m	0.1140
Nozzle diameter, m	0.012	Tool joint outer diameter, m	0.166
Bit weight, kg	40	Total drill string length, m	2500
Rotary speed, rpm	30	***Hydraulic Thruster Parameters***	
Drilling mud Parameters		HT inner diameter, m	0.152
Density, kg/m³	1200	Piston diameter of HT, m	0.110
Plastic viscosity	0.01	HT shaft weight, kg	500
Dynamic viscosity	5	***Well Parameters***	
PDM Parameters		Well inclination angle, °	0
Rotor lobe count	9	Friction coefficient	0.3
Stator lobe count	10	Casing inner diameter, m	0.2699
PDM outer diameter, m	0.172	Cased interval length, m	2000
Thread pitch coefficient	4	Open hole diameter, m	0.210
PDM weight, kg	500	Open hole interval length, m	500

The results of mathematical modeling demonstrate that this relationship increases in direct proportion to the drilling fluid flow rate.

The red curve (Figure 3) indicates initial load at bit-bottomhole contact, while the blue curve shows final load after hydro-mechanical system stabilization.

Mathematical modeling results demonstrate that this relationship follows a directly proportional function and decreases monotonically.

Figure 4 presents calculated dependencies of well inclination angle versus both bit load (generated by the hydraulic thruster) and bit rotational speed at a fixed pump flow rate Q = 0.028 m³/s.

Figure 5 displays the hydro-mechanical system's torque (N•m) and rotational speed (rpm) as functions of working fluid flow rate (m³/s). Analysis reveals the correlation between these key parameters. The intersection point corresponds to maximum efficiency.

Figure 6 shows overall system efficiency (%) as a function of fluid flow rate (m³/s). The vertical axisrepresents efficiency values (14% to 19%), while the horizontal axis indicates flow rates ranging from 0.0225 m³/s to 0.0425 m³/s. The data demonstrate an inverse relationship between flow rate and system efficiency. Peak efficiency (19%) occurs at 0.023 m³/s, decreasing to 14% at maximum flow (0.0425 m³/s). This trend suggests either energy losses or suboptimal equipment configuration.

TABLE 5. Iterative Calculation Results

Iteration No	PDM Speed, rpm	Axial Load from Hydraulic Thruster, N	Bit Torque, N•m	PDM Pressure Drop, KPa
1	126	32769	1137	1323.5
2	120	38594	1345	1875.9
3	116	40842	1417	2092.2
4	115	41194	1485	2129.3
5	115	41216	1489	2131.7

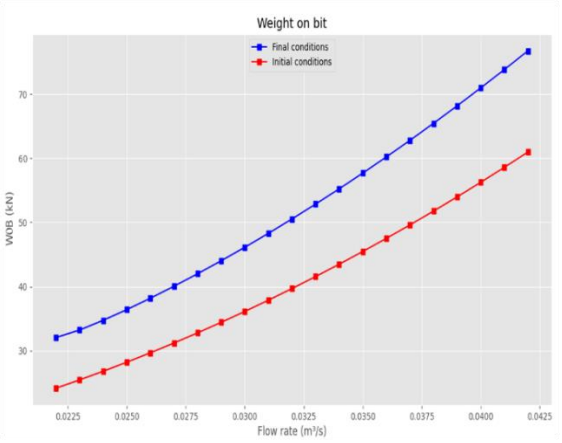


Figure 3. Thruster load as a function of mud pump flow rate

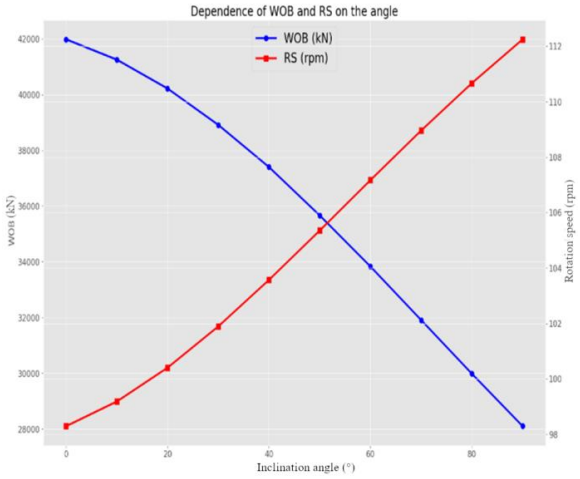


Figure 4. Relationship between rotational speed/load and well inclination angle

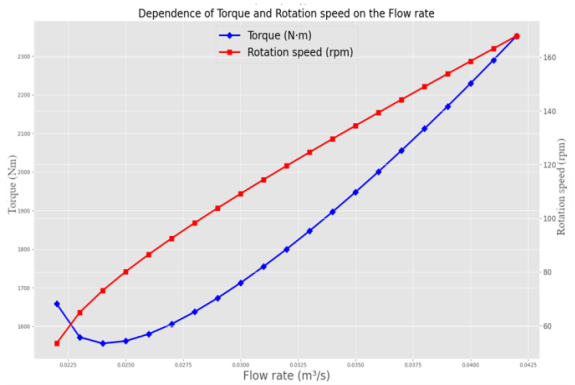


Figure 5. Torque and rotational speed versus drilling mud flow rate

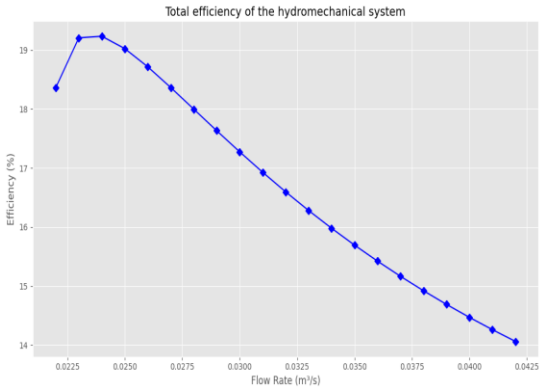


Figure 6. Hydro-mechanical system efficiency versus drilling mud flow rate

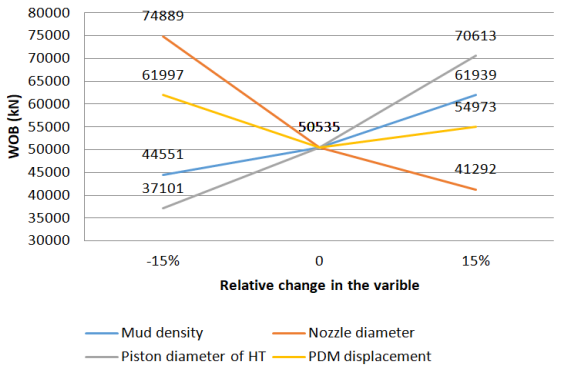


Figure 7. Sensitivity analysis

TABLE 6. Parameters for sensitivity analysis

Flow rate Q = 0.032 (m^3/s)			
Parameters	**-15%**	**0**	**15%**
Mud density (kg/m)	1020	1200	1380
Nozzle diameter (m)	0,01	0,012	0,014
Piston diameter of HT (m)	0,094	0,11	0,126
PDM displacement ($m^3 10^{-3}/s$)	11,56	13,6	15,6

4. DISCUSSION

Drilling with a hydraulic thruster and a Positive Displacement Motor (PDM) is a complex, yet technologically viable process. The complexity arises from the fact that both BHA components depend on well flushing parameters. Specifically, the bit torque depends on the axial load, which in turn affects the pressure drop across the PDM. This pressure drop defines the hydraulic component of the axial force generated by the downhole thruster. The incremental load resulting from the incremental torque is transmitted back to the bit, thereby closing the feedback loop.

This problem can be solved analytically, but the study employed an iterative computational algorithm, which allows for model refinement and increased complexity without requiring analytical solutions.

A comprehensive sensitivity analysis was conducted to evaluate the influence of four critical operational parameters on the hydraulic system's performance under a constant flow rate of 0.032 m³/s (Figure 7 and Table 6). The study employed a parametric variation approach (±15% deviation from baseline values) to systematically assess each parameter's individual and synergistic effects on WOB. The increase in density elevates hydraulic resistance, potentially influencing pressure losses and load distribution. During drilling operations, bit nozzles are subject to wear-induced diameter enlargement, which may reduce system WOB due to decreased pressure differential. The piston diameter and PDM displacement volume represent non-adjustable parameters during drilling operations, yet remain critical for load optimization during equipment selection and exert significant influence on system performance.

Accordingly, the model in this study was enhanced to account for the effect of bit rotational speed on bit torque. There is a correlation between the linear cutter velocity and the axial reaction force of the rock on the cutter face during cutting. Figure 8 demonstrates that as the cutter's relative velocity increases, the axial reaction force at the bottomhole also increases, leading to a reduction in single-cutter penetration depth.

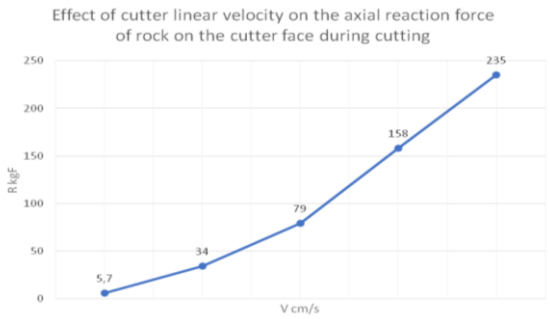


Figure 8. Effect of cutter linear velocity on the axial reaction force of rock on the cutter face during cutting

From a physical process perspective, the key aspect lies in the interaction between hydrodynamic and mechanical system components. During hydraulic thruster operation, drilling fluid flow energy is converted into axial mechanical load through a complex system of hydrodynamic resistances. Crucially, a feedback mechanism exists between WOB and flow characteristics: increased axial load generates (see Equation 8) higher reactive torque, which subsequently affects the pressure differential across PDM and thruster performance parameters. Importantly, increased PDM torque reduces rotational speed, thereby altering the relationship between bit torque and generated axial load.

In turn, the bit rotational speed at a given pump flow rate depends on the PDM torque (see Equation 13).

The use of a numerical solution method allowed for the incorporation of this new influencing factor into the system's behavior, and the computational results are presented in Figures 3 to 6.

Optimization of BHA parameters with a hydraulic thruster requires selection of nozzle sizes and piston diameters capable of generating the designed WOB. For standard 215.9 mm bits, typical WOB ranges between 100-120 kN for roller cone bits and 50-70 kN for PDC bits. While field tests (Table 1) have demonstrated successful hydroulic thruster performance across various regions, these trials lacked downhole WOB measurement capabilities. Load estimation relied solely on hook load reduction - an indirect parameter that cannot accurately reflect actual bit force generation.

5. CONCLUSIONS

Based on a comprehensive analysis of theoretical and experimental data, and as part of research conducted under the state assignment of the Ministry of Science and Higher Education of the Russian Federation (FSRW-2023-0002), the following scientifically substantiated conclusions were obtained:

The effectiveness of hydraulic thruster applications in drilling deep, deviated, and horizontal wells has been theoretically proven and experimentally confirmed. The proposed modeling approach for the hydro-mechanical system "Bit – PDM – Hydraulic Thruster – Mud Pump" enables equipment selection and optimization of the well drilling regime. This solution is particularly significant when drilling wells with high deviations from vertical, where bit weight-on-bit (WOB) becomes a limiting factor for the achievable interval length. Calculations performed within the study demonstrated that the hydro-mechanical system can provide a WOB exceeding 72 kN during vertical drilling and over 32 kN in the horizontal section, excluding the weight of heavyweight drill pipes.

A mathematical model has been developed and successfully implemented, along with an iterative calculation algorithm coded in Python, which accounts for and integrates the key systems of the drilling assembly. The model allows for simulation and optimization of system and equipment parameters by varying input data and visualizing calculation results in the form of graphs and nomograms. The simulation results demonstrate high accuracy and convergence of the calculated data, confirming the reliability of the proposed model.

The modeling showed that when a hydraulic thruster is included in the BHA, the mud flow rate becomes the primary control parameter for regulating bit WOB and rotation speed.

The integration of a hydraulic thruster with a PDM enhances the operational scope of sliding-mode technology for trajectory adjustment in ERD wells. By enabling drill string reciprocation while maintaining toolface orientation, the thruster mitigates sticking risks during directional operations. Furthermore, in tight hole conditions or differential sticking scenarios, the thruster generates controlled upward/downward impulse loads, providing an alternative to conventional jarring tools, which exhibit limited efficacy in ERD applications.

The proposed solutions demonstrate significant practical value, as they not only enhance the efficiency of deviated and horizontal well drilling but also, under certain conditions, enable the long-term feasibility of drilling highly deviated wells using PDMs or RSS.The developed methodologies can be applied in the design of hydraulic thrusters and used for selecting PDM and bit parameters to optimize the drilling process.

The use of hydraulic thrusters opens new possibilities for advancing drilling technology in wells with high deviations from vertical. Further research may focus on optimizing the device design, expanding its operational range, and integrating it with intelligent drilling control systems.

6. REFERENCES

1. Wang H, Huang H, Bi W, Ji G, Zhou B, Zhuo L. Deep and ultra-deep oil and gas well drilling technologies: Progress and prospect. Natural Gas Industry B. 2022;9(2):141-57. 10.1016/j.ngib.2021.08.019
2. Zhdaneev OV, Zaytsev AV, Prodan TT. Possibilities for creating Russian high-tech bottomhole assembly. Записки Горного института. 2021;252:872-84. 10.31897/PMI.2021.6.9
3. Alzaki H, Rahmani N, Carr M, editors. Breaking ERD Records with Optimized Engineering and Practices: Making the Impossible Possible. SPE Middle East Oil and Gas Show and Conference; 2021: SPE. 10.2118/204550-MS
4. Panga M, Al Hammadi F, Marsh J, Mohamed A, Pace D, Wills J, editors. Ultra ERD Well Completions: Extending the Limits to Drive Profitability. Abu Dhabi International Petroleum Exhibition and Conference; 2023: SPE. 10.2118/216085-MS
5. Litvinenko V, Bowbrick I, Naumov I, Zaitseva Z. Global guidelines and requirements for professional competencies of

natural resource extraction engineers: Implications for ESG principles and sustainable development goals. Journal of Cleaner Production. 2022;338:130530. 10.1016/j.jclepro.2022.130530

6. Liagov I, Liagov A, Liagova A. Optimization of the Configuration of the Power Sections of Special Small-Sized Positive Displacement Motors for Deep-Penetrating Perforation Using the Technical System “Perfobore”. Applied Sciences. 2021;11(11):4977. 10.3390/app11114977
7. Zhong Z, Li Y, Zhao Y, Ju P. A Method of Evaluating the Effectiveness of a Hydraulic Oscillator in Horizontal Wells. Sound & Vibration. 2023;57(1):15-27. 10.32604/sv.2023.041954
8. Shi X, Huang W, Gao D, Zhu N, Li W. Extension limit analysis of drillstring with drag reduction oscillators in horizontal drilling. Geoenergy Science and Engineering. 2023;228:211996. 10.1016/j.geoen.2023.211996
9. Tian J, Mao L, Yang Y, Song H, Song J. Study on directional drilling coupling dynamics based on drill string rotary controller. Journal of Computational and Nonlinear Dynamics. 2023;18(9):091008. 10.1115/1.4062911
10. Li Y, Yu Y, Sun Y, Li B. Dynamics and fatigue life investigation of push-the-bit rotary steerable bottom hole assembly with whirling finite element model. Engineering Failure Analysis. 2023;150:107300. 10.1016/j.engfailanal.2023.107300
11. Никишин ВВ, Блинов ПА, Терехин ВА. Effectiveness of the method for selecting rotary-steerable systems based on the machine learning algorithm Random Forest Classifier. Bulletin of the Tomsk Polytechnic University Geo Assets Engineering. 2024;335(4):185-99. 10.18799/24131830/2024/4/4129
12. Litvinenko VS, Dvoynikov MV. Methodology for determining the parameters of drilling mode for directional straight sections of well using screw downhole motors. Записки Горного института. 2020;241:105-12. 10.31897/PMI.2020.1.105
13. El Sabeh K, Gaurina-Međimurec N, Mijić P, Medved I, Pašić B. Extended-reach drilling (ERD)—the main problems and current achievements. Applied Sciences. 2023;13(7):4112. 10.3390/app13074112
14. Wanpeng C, Danyang M, Yongmei G. Deep learning approach to prediction of drill-bit torque in directional drilling sliding mode: Energy saving. Measurement. 2025;250:117144. 10.1016/j.measurement.2025.117144
15. Barjini AH, Khoshnazar M, Moradi H. Design of a sliding mode controller for suppressing coupled axial & torsional vibrations in horizontal drill strings using Extended Kalman Filter. Journal of Sound and Vibration. 2024;586:118477. 10.1016/j.jsv.2024.118477
16. Zhang H, Ashok P, van Oort E, Shor R. Investigation of pipe rocking and agitator effectiveness on friction reduction during slide drilling. Journal of Petroleum Science and Engineering. 2021;204:108720. 10.1016/j.petrol.2021.108720
17. Shi X, Huang W, Gao D. Optimal design of drag reduction parameters with the rotary reciprocating control system and hydraulic oscillators in sliding drilling. Geoenergy Science and Engineering. 2024;243:212856. 10.1016/j.geoen.2024.212856.
18. El Sabeh K, Pašić B, Mijić P, Medved I. Reducing torque and drag in extended-reach wells using thermoplastic polymers for protective sliding rings. Applied Sciences. 2024;14(14):6161. 10.3390/app14146161
19. Sadeghi AN, Arıkan KB, Özbek ME. Modelling and controlling of drill string stick slip vibrations in an oil well drilling rig. Journal of Petroleum Science and Engineering. 2022;216:110759. 10.1016/j.petrol.2022.110759
20. Long Y, Wang X, Wang P, Zhang F. A method of reducing friction and improving the penetration rate by safely vibrating the drill-string at surface. Processes. 2023;11(4):1242. 10.3390/pr11041242
21. Li B, Zhu X, Dong L. Axial transfer characteristics of a drillstring in conjunction with bit penetration behavior in horizontal wells. Geoenergy Science and Engineering. 2024;240:212989. 10.1016/j.geoen.2024.212989
22. Islamov S, Islamov R, Shelukhov G, Sharifov A, Sultanbekov R, Ismakov R, et al. Fluid-loss control technology: From laboratory to well field. Processes. 2024;12(1):114. 10.3390/pr12010114
23. Raupov I, Rogachev M, Sytnik J. Overview of Modern Methods and Technologies for the Well Production of High-and Extra-High-Viscous Oil. Energies. 2025;18(6):1498. 10.3390/en18061498
24. Leushevaa E, Alikhanov N, Moradib ST. Experimental Evaluation of Influence of Physico-chemical Properties of Surfactants on Drilling Process in Pay-zones. International Journal of Engineering, Transactions A: Basics. 2025;38(4):819-29. 10.5829/ije.2025.38.04a.13
25. Nikitin V, Agrelkina M. Justification for the Selection of a Relative Permeability Model in the Task of Predicting Drilling Fluid Filtrate Invasion into the Formation. International Journal of Engineering Transactions A: Basics. 2025;38(10):2312-20. 10.5829/ije.2025.38.10a.08
26. Badrouchi F, Rasouli V, Badrouchi N. Impact of hole cleaning and drilling performance on the equivalent circulating density. Journal of Petroleum Science and Engineering. 2022;211:110150. 10.1016/j.petrol.2022.110150
27. Liu T, Leusheva E, Morenov V, Li L, Jiang G, Fang C, et al. Influence of polymer reagents in the drilling fluids on the efficiency of deviated and horizontal wells drilling. Energies. 2020;13(18):4704. 10.3390/en13184704
28. Dvoynikov MV, Minaev YD, Minibaev VV, Kambulov EY, Lamosov ME. Technology for killing gas wells at managed pressure. Известия Томского политехнического университета Инжиниринг георесурсов. 2024;335(1):7-18. 10.18799/24131830/2024/1/4315
29. Nikitin VI. Problem solution analysis on finding the velocity distribution for laminar flow of a non-linear viscous flushing fluid in the annular space of a well. Записки Горного института. 2022;258:964-75. 10.31897/PMI.2022.93
30. Tikhonov VS, Baldenko FD, Bukashkina OS, Liapidevskii VY. Effect of hydrodynamics on axial and torsional oscillations of a drillstring with using a positive displacement motor. Journal of Petroleum Science and Engineering. 2019;183:106423. 10.1016/j.petrol.2019.106423
31. Song J, Liu S, He Y, Jiang S, Zhou S, Zhu H. The state-of-the-art review on the drill pipe vibration. Geoenergy Science and Engineering. 2024;243:213337. 10.1016/j.geoen.2024.213337
32. Li X, Li M, Guo X, Agila W, Zhang H, Yang J. Influential mechanism of operational parameters on internal resonance of drill strings by numerical simulations. Engineering Research Express. 2024;6(4):045502. 10.1088/2631-8695/ad7cc9
33. Lalehparvar M, Vaziri V, Aphale SS, Wiercigroch M. Time delay effect in suppressing drill-string stick-slip vibration. IFAC-PapersOnLine. 2024;58(5):164-9. 10.1016/j.ifacol.2024.07.080
34. Volpi LP, Lobo DM, Ritto TG. A stochastic analysis of the coupled lateral–torsional drill string vibration. Nonlinear Dynamics. 2021;103(1):49-62. 10.1007/s11071-020-06099-z
35. Dvoynikov M, Kutuzov P. Analysis of efficiency of communication channels for monitoring and operational control of oil and gas wells drilling process. International Journal of Engineering, Transactions A: Basics. 2025;38(1):120-31. 10.5829/ije.2025.38.01a.12
36. Buslaev GV, Belkin AS, editors. Development and field testing of the downhole multi-purpose thrusting device for drilling of deep vertical and directional wells. SPE Russian Petroleum Technology Conference; 2015: SPE. 10.2118/176539-MS

37. Zhao J, Cui X, editors. Design and field test of hydraulic thruster. IOP Conference Series: Earth and Environmental Science; 2020: IOP Publishing. 10.1088/1755-1315/558/2/022067
38. Giles C, Seesahai T, Brooks J. Hydraulic thrusting devices improve drilling efficiencies. World Oil. 2001;222(9):59-64.
39. Reich M, Hoving P, Makohl F, editors. Drilling performance improvements using downhole thrusters. SPE/IADC Drilling Conference and Exhibition; 1995: SPE. 10.2118/29420-MS
40. Moradi SST. A probabilistic study on hole cleaning optimization. Записки Горного института. 2022;258:956-63. 10.31897/PMI.2022.67

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

این مطالعه به بررسی چالش‌های کلیدی در عملیات حفاری با دسترسی گسترده (ERD) می‌پردازد و تمرکز آن بر انتقال محدود نیروی محوری به مته، افزایش افت فشار هیدرولیکی، و ارتعاشات رشته حفاری است. راه‌حل فناورانه پیشنهادی شامل ادغام یک تراستر (تقویت‌کننده) هیدرولیکی در مجموعه ابزار کف چاه (BHA) می‌باشد .یک مدل ریاضی از سامانه هیدرومکانیکی شامل مته، موتور جابجایی مثبت(PDM) ، تراستر هیدرولیکی، و پمپ گل طراحی شده است که در آن، ارتباطات بین اجزا و حالت‌های مختلف حفاری در نظر گرفته شده‌اند. الگوریتمی تکراری برای محاسبات عددی در محیط پایتون توسعه یافته است تا بهینه‌سازی پارامترهای حفاری را ممکن سازد .نتایج نشان می‌دهد که تراستر هیدرولیکی می‌تواند تا ۷۲ کیلونیوتن نیروی محوری در چاه‌های عمودی و ۳۲ کیلونیوتن در بخش‌های انحرافی تأمین کند، که این امر موجب بهبود کارایی فرآیند حفاری می‌گردد. این مدل می‌تواند به عنوان ابزاری جهت پشتیبانی از طراحی BHA و بهینه‌سازی پارامترهای حفاری مورد استفاده قرار گیرد و به‌ویژه در ساخت چاه‌های ERD از ارزش بالایی برخوردار است.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1088-1100

International Journal of Engineering

Journal Homepage: www.ije.ir

Distribution Substation Allocation Considering Load Uncertainty based on Information Gap Decision Theory

M. Ramezani*, H. Falaghi, A. Azimi

Faculty of Electrical and Computer Engineering, University of Birjand, Birjand, Iran

PAPER INFO

Paper history:
Received 10 March 2025
Received in revised form 24 May 2025
Accepted 11 June 2025

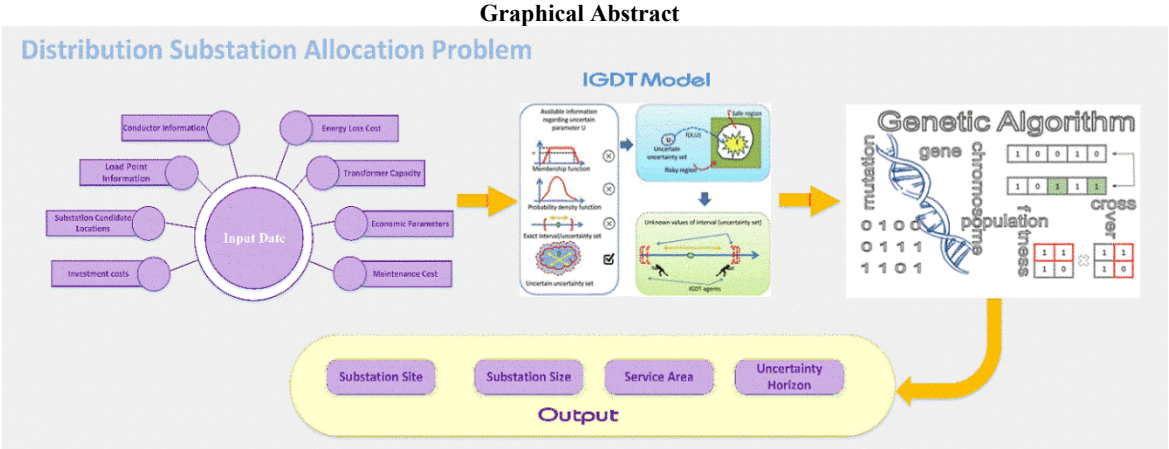
Keywords:
Distribution Substation
Load Uncertainty
Information Gap Decision Theory
Genetic Algorithm

ABSTRACT

This paper presents a novel methodology for optimizing the location, sizing, and service area of distribution substations, explicitly incorporating load uncertainty. The escalating integration of renewable distributed generation and electric vehicles has introduced substantial uncertainties into distribution loads, exceeding conventional prediction errors. To mitigate this challenge, a robust optimization framework is employed, enabling risk modeling and ensuring secure operation under extreme uncertainty scenarios. Information Gap Decision Theory (IGDT) is utilized to effectively model load uncertainty, with a risk-averse strategy adopted to enhance robustness. The objective is to minimize the total cost of distribution system planning, encompassing investment, maintenance, and loss-related expenditures, while adhering to technical constraints. Initially, the distribution substation problem is solved using predicted load point values. Subsequently, a risk-averse IGDT approach is applied to identify robust solutions under load uncertainty. The efficacy of the proposed methodology is demonstrated using a test system. Results indicate that an average load increment of approximately 42.88% can be accommodated with a corresponding 50% increase in permissible cost. In such scenarios, if potential load uncertainty does not materialize, the average substation load would be approximately 55% of their nominal capacities.

doi: *10.5829/ije.2026.39.05b.04*

Graphical Abstract



NOMENCLATURE

Sets			
S	Set of selected substations	Cap_{n_i}	Rated capacity of substation i (MVA)
L	Set of load points	u	uncertain variable
DL	Set of downstream feeders	$\tilde{u}$	Predicted value of the uncertain variable u
		α	Uncertainty horizon

*Corresponding Author Email: mramezani@birjand.ac.ir (M. Ramezani)

Please cite this article as: Ramezani M, Falaghi H, Azimi A. Distribution Substation Allocation Considering Load Uncertainty based on Information Gap Decision Theory. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1088-100.

Variables			
C_{total}	Total cost ($)	β	Maximum allowable of objective function increase
C_{IC}	Investment cost ($)	Z_{base}	Base value of the objective function
C_{MC}	Maintenance cost ($)	**Parameters**	
C_{loss}^{S}	Substation energy loss cost ($)	$Infr$	Inflation rate
C_{loss}^{LL}	Downstream network energy loss cost ($)	$Intr$	Interest rate
C_{CS_i}	Fixed cost of substation i	n_y	Equipment life (y)
C_{FL}	Cost of downstream network ($/km)	K_E	Cost of energy loss ($/kWh)
$C_{MC_i}^{S}$	Annual maintenance cost of substation i ($)	LSF	Loss factor
$C_{MC_j}^{LL}$	Annual maintenance cost of downstream network ($)	R	Conductor resistance (Ω/km)
D_{ij}	Distance between substation i and load point j (km)	V	Nominal voltage magnitude (v)
x_i, y_i	Coordinates of node i	$S_{j,y}$	Predicted load of load point j at year y (VA)
β_{ij}	Decision variable (0: load point has not been connected to Substation, 1: load point has been connected to Substation)	$S_j, \bar{S}_j, \tilde{S}_j$	Load value j at horizon year /mean load/load variable at second level
P_{nl_i}	Copper losses of substation i at the rated loading (kW)	n_s	Number of existing and potential substations
P_{l_i}	Iron losses in substation *i* at the rated loading (kW)	n_l	Number of load points
$Cap_{l_i,y}$	Total power provided by substation *i* at year y (MVA)	Z_{cap}	Maximum permissible loading percentage
		ΔV_{max}	Maximum allowed voltage drop (V)

1. INTRODUCTION

The planning and operation of distribution networks face increasing challenges due to uncertainties associated with system participants. Deploying distributed generation, particularly with high penetration levels, and the growing adoption of electric vehicles introduce significant uncertainties into distribution networks, adversely impacting their efficient planning and operation (1).

The primary objective of distribution network planning is to meet future demand with the most cost-effective solution while adhering to technical constraints. The sub-transmission system can be divided into two main components: substations and primary feeders (2). These components require multiple variables to be considered during the planning stage. Various distribution planning models exist, which can be broadly categorized into subsystem and total system approaches. Subsystem approaches focus on planning either substations or feeders (2, 3), whereas total system techniques address both substations and feeders as a single, integrated problem (4, 5).

The complexity of distribution system planning is further compounded by its inherently nonlinear nature and the uncertain environment in which it operates. Consequently, many existing methods do not simultaneously account for both subsystems during the planning stage. A review of relevant studies reveals that most focus has been on allocating of distribution substations, transformer placement, feeder routing, and conductor sizing.

1. 1. Literature Review Numerous papers have been published on optimal distribution substaccction placement, typically modeling the problem in deterministic settings to determine optimal substation siting, sizing, and timing, as well as service area definition.

Given the similarities between substation allocation and transformer placement, significant literature on transformer allocation has also been examined. Relevant literatures have discussed the subject from different points of view. It is worth to categorize them dependent to their innovations or bold features.

- **Optimization method**

Existing researches in this field can be categorized into three major groups:

1. Mathematical optimization (5-7)
2. Evolutionary optimization (1, 8, 9)
3. Heuristic optimization (3, 8, 10-12)

Mathematical optimization approaches often provide optimal solutions but are highly complex and impractical for large-scale case studies. Evolutionary optimization algorithms are more straightforward to implement across various problem types. However, these methods may struggle with local optima and can be time-consuming. Heuristic methods, while simple and easy to apply, lack guarantees for achieving globally optimal solutions.

Mazhari et al. (3) employed a modified fuzzy membership matrix and a memorable cost index vector to determine the substation service areas. Also, a hybrid heuristic and learning automata (LA)-based algorithm is

used for substations sizing and defining the associated service area, simultaneously.

- **Study environment**

The substation allocation problem is solved considering vertical power system structure or deregulated power system. Some research has been conducted in traditional power system environments (8, 13, 14), while others address market-oriented frameworks (1, 15). Both of these environment are applicable and can be considered in power system studies depending on the environment of the case under study.

- **Scope of study**

The studies differ in their scope and context. Some researches just focus on distribution substation problem [1-3] and in some instances, substation allocation is integrated with the planning of other components, such as distributed generation units (4, 5, 15, 16) or overall distribution expansion planning (13, 14).

A dynamic bacterial foraging algorithm determines the optimal location, capacity, service area, and construction timing of substations under technical constraints, aiming to maximize profits from energy sales while minimizing associated costs (1).

Behzadi and Bagheri (5) addressed the siting and sizing of substations, feeder routing, and renewable-based distributed generation (DG) planning to create resilient and sustainable distribution systems. First, all problem relations are convexfied. Then, a mixed-integer quadratically constrained programming (MIQCP) algorithm is employed to minimize investment, operational, emission, and resiliency costs while accounting for renewable DG uncertainty through scenario-based modeling. Jalai et al. (15) Shayeghi and Bagheri (16) focussed on the expansion planning of distributed generation units and substations, incorporating energy prices for electricity received via substations or produced by distributed generators.

- **Static or dynamic modelling**

Determining the time of substation installation is one of the remarkable features that can be considered while categorizing studies. Distribution substation problem can be solved taking the whole planning period in optimization process (1, 17) or can be modeled considering different time intervals (2, 7, 8). It is obvious that solving the problem in dynamic procedure adds more accuracy to the results, on the other hand, makes the problem more complicated.

Temraz and Salama (2) used the General Interactive Optimizer (GINO) to solve the non-linear planning of problem. The model includes a continuous function that accurately simulates various cost components of a distribution substation. In this study besides size, location, and service area, time of substation installation is determined using pseudo dynamic procedure.

The Imperialist Competitive Algorithm (ICA) was utilized by Najafi et al. (8) to optimize the sizing, siting, and timing of medium-voltage substations (distribution transformers). This study introduces a loss characteristic matrix as a novel concept for the substation placement problem.

Lastly, a Mixed Integer Linear Programming (MILP) method is employed by El-Fouly et al. (7) to address the siting, sizing, and timing of distribution substations.

- **Incorporating geographical information system**

Several studies attempt to address real-world planning challenges, ranging from parameter uncertainty (18) to geographical characteristics (11, 19, 20). For instance, Yu et al. (11) proposed a method to determine the optimal location and size of substations based on GIS and semi-supervised learning. This method highlights the significance of land costs in substation allocation. By integrating actual land cost into the objective function, GIS tools enable precise substation location determination after load clustering using the k-means method.

Vahedi et al. (19), substation expansion planning is applied to a large-scale real network using GIS raster maps for analysis. To manage the complexity of large-scale networks, the planning area is divided into irregular mini-regions. A "system of systems" approach is introduced to facilitate interactions between these mini-regions, ensuring the overall integrity of the distribution network.

- **Reliability modelling**

One of the features that makes the proposed approaches more applicable to real distribution network is incorporating reliability modeling into the problem. Some researches consider the possibility of components failures in the study. A deterministic heuristic algorithm is used by Mazhari et al. (10) to approximate substation service areas, employing an expert selection strategy to enhance the likelihood of achieving a globally optimal solution. This study also examines supply interruptions caused by substation unreliability and feeder failures.

A related problem, distribution transformer planning, was addressed by Esmaeeli et al. (21), where reliability costs and load growth are considered. The paper allows transformer loading to exceed the nameplate rating during parts of the duty cycle, demonstrating that this does not significantly impact the relative thermal aging rate of transformers.

Smart substation allocation, aimed at restoring power to interrupted customers, was explored by Sun et al. (22). Interruption costs and reliability criteria are key factors in determining the number and location of smart

substations. The problem is solved using Mixed Integer Linear Programming (MILP).

- **Uncertainty modelling**

Accounting for system parameter uncertainties is now essential to ensure practical solutions for real-world systems. Haghifam and Shahabi (23) proposed a pseudo-dynamic methodology for substation planning under load uncertainty modeled with LR fuzzy numbers.

Esmaeelia et al. (24), a Conditional Value at Risk (CVaR)-based method is applied for optimal planning of LV distribution networks and distribution transformer allocation. System uncertainties are modeled using scenarios, and the problem is solved with Mixed Integer Nonlinear Programming (MINLP).

Franco et al. (25) employed a robust model to address load uncertainty, using chance constraints to ensure substation capacity limits are not violated within a given robustness probability. Monte Carlo simulations validate constraint satisfaction under uncertain conditions.

Abedi et al. (18) used a genetic algorithm and heuristic clustering method to assign uncertain load centers to existing or candidate substations. The approach focuses on uncertain load centers with variability in both load values and locations. Genetic algorithms generate different load scenarios, while clustering provides potential sizes and locations for new substations. Fuzzy membership function also was used to model load uncertainty in distribution transformer allocation (17).

1. 2. Proposed Method This paper addresses the challenges of substation siting and sizing under high uncertainty, considering the potential presence of electric vehicles and the high penetration of distributed renewable generation. These factors increase the uncertainty of electric load demand, surpassing traditional load forecasting errors. The random behavior of renewable generation and electric vehicle charging patterns makes it difficult to model the probability distribution function of loads at different load centers with sufficient accuracy.

To tackle this issue, the Information-Gap Decision Theory (IGDT) is applied as a robust solution for substation siting and sizing under uncertainty. IGDT is particularly advantageous for decision-making in uncertain environments as it enables risk-based decisions without requiring knowledge of the probability distribution functions of uncertain parameters (26). IGDT has been successfully applied to various power system problems, including self-scheduling of generation companies (27), day-ahead achedualing of micro grids (28), and bidding strategy problems (29).

The main contributions of this paper are as follows:

- Incorporation of significant uncertainty in electric load demand, considering the potential impact of electric vehicles and renewable energy sources.
- Application of IGDT to effectively manage and model uncertainties.
- Development of a robust solution framework that accommodates load points with varying levels of uncertainty.
- Economic analysis of the impact of prediction errors in uncertainty estimations on the final solution and overall system costs.

In the following, after introducing the principles of the problem in section 2, the objective function and constraints are discussed in section 3. After that, section 4 presents the IGDT-based decision-making methodology. Section 5 states the problem modeling using IGDT. The optimization algorithm is expressed in section 6. Finally, the numerical studies are explained in Section 7, showcasing the applicability and effectiveness of the proposed method.

2. SUBSTATION ALLOCATION PRINCIPLES

In distribution network planning and expansion, it is essential to determine the optimal location, size, and service area of distribution substations in order to meet demand at the minimum cost. To achieve this, the area under study is divided into smaller sub-areas, and load prediction is performed for each sub-area for the target year. Each sub-area is then modeled as a load point located at the center of gravity of that specific area. During the study period, annual load values (from the first year to the horizon year) are calculated using the load growth factor. Based on the load point values and their locations, as well as the candidate locations of distribution substations, the problem is formulated to obtain the following decision variables:

- The number of distribution substations
- The optimal location of distribution substations
- The size of distribution substations
- The service area of distribution substations

3. PROBLEM FORMULATION WITHOUT UNCERTAINTY

The substation planning problem is first formulated in a deterministic environment with the following objective function and constraints.

3. 1. Objective Function As mentioned in the introduction, various studies have used different objective functions and constraints in distribution substation planning. These problems are often solved within a traditional, deterministic environment. Commonly, the objective functions include the investment cost of substations (including equipment and land cost) and feeders, maintenance costs, and energy loss costs. The problem formulation is as follows:

$$Min\ C_{total} = C_{IC} + C_{MC} + C^{S}_{loss} + C^{LL}_{loss} \quad (1)$$

$$C_{IC} = \sum_{i\in S} C_{CS_i} + \sum_{i\in S}\sum_{j\in L} C_{FL} \times D_{ij} \times \beta_{ij} \quad (2)$$

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3)$$

$$C_{MC} = \sum_{y=1}^{n_y} ((\frac{1+Infr}{1+Intr})^{y} \times (\sum_{i\in S} C^{S}_{MC_i} + \sum_{i\in DL} C^{LL}_{MC_i})) \quad (4)$$

$$C^{s}_{loss} = \sum_{y=1}^{n_y}\sum_{i\in S} (\frac{1+Infr}{1+Intr})^{y} \times K_E \times 8760 \times (P_{nl_i} + P_{l_i} \times (\frac{Cap_{l_i,y}}{Cap_{n_i}})^2 \times LSF) \quad (5)$$

$$C^{LL}_{loss} = \sum_{y=1}^{n_y}\sum_{i\in S}\sum_{j\in L} (\frac{1+Infr}{1+Intr})^{y} \times K_E \times LSF \times D_{ij} \times \frac{8760 \times R}{V^2} \times S^2_{j,y} \times \beta_{ij} \quad (6)$$

$$Cap_{l_i,y} = \sum_{j=1}^{n_l} \beta_{ij} \times S_{j,y} \quad (7)$$

The objective function is presented in Equation 1, where the first term represents the investment costs of the selected substations within the planning horizon. The second term accounts for the total maintenance costs of substations and feeders. The third term reflects the total cost associated with energy losses in substations, including both no-load losses and loading losses. The costs imposed by feeder losses are modeled in the fourth term of Equation 1.

Considering load growth during the planning period, an annual load growth rate of 2% is used. Transformer losses, feeder losses, and substation loads are updated based on this growth rate. Since the load center points are considered hypothetical, this study uses Euclidean distance to calculate the distance between the load points and candidate substations.

3. 2. Constraints To operate the distribution network within its allowable component limits, several constraints must be imposed during system planning. The problem constraints are as follows:

3. 2. 1. Load Supply Constraints The primary purpose of a power system is to meet all electricity demands. In this work, each load point must be connected to one selected substation.

3. 2. 2. Capacity Constraints To ensure sufficient reserve capacity for future load growth, and to extend the lifetime of transformers and other components, the maximum load for each distribution substation should be limited to a predefined value.

$$\sum_{j=1}^{n_l} \beta_{ij} \times S_j \le Z_{cap} \times Cap_{n_i} \qquad i=1, \ldots \quad (8)$$

3. 2. 3. Voltage Constraints As with other planning problems, permissible voltage magnitudes at load points are considered. A fixed voltage range is applied as follows:

$$\left| \frac{D_{ij} \times S_j}{V^2} \times R \right| \le \Delta V_{\max} \quad (9)$$

3. 2. 4. Radial-flow Constraints Each load point should be supplied by only one substation to reflect the practical operation of radial distribution networks.

4. INFORMATION GAP DECISION THEORY

IGDT is a risk control method that can be used for decision-making in the presence of uncertainties. Unlike other probabilistic methods, IGDT does not require probability density functions for uncertain parameters. Instead, IGDT only uses the predicted values of uncertain parameters (30). This method is particularly useful when initial information is limited, or when changes cannot be predicted with certainty.

IGDT creates a confidence interval by considering a prediction horizon for uncertain parameters, ensuring that if the uncertain parameter falls within this interval, the value of the objective function will not exceed the bounds determined by the planner (29).

Several models within the IGDT framework can represent parameter uncertainty. In this study, uncertainty is modeled using the envelope bound model. Equation 10 illustrates the uncertainty modeling function, where $\tilde{u}$ is the predicted value of the uncertain variable u, and α is the uncertainty horizon. Both parameters, u and α, are uncertain.

$$U(\alpha, u) = \{u : | \frac{u - \tilde{u}}{\tilde{u}} | \le \alpha \ , \alpha \ge 0\} \quad (10)$$

As previously mentioned, IGDT performs risk-based decision-making. Depending on company policies, either a risk-averse or risk-seeking approach can be adopted. In a risk-averse policy, the IGDT method seeks to maximize the uncertainty interval and provide a solution that guarantees a certain magnitude of the objective function. In this case, the decision maker minimizes risk by focusing on the most unfavorable deviations of uncertain parameters (such as reduced output from distributed generators or increased network load). By considering the worst-case scenario of uncertainty, the best decision is made. Thus, resilience against unfavorable deviations from the forecasted values of uncertain parameters is modeled using a robustness function (26). This robust function ensures that the objective function value never exceeds a predetermined critical value; thus, providing a safeguard against excessive risk. Equation 11 states the robustness function:

$$\tilde{\alpha} = max\{\alpha: \text{maximum} C_{\text{Total}} \text{ which is less than a target value.}\} \quad (11)$$

5. PROBLEM MODELING USING The IGDT METHOD

Distribution substations are installed at appropriate locations to meet customer's demands in the distribution network. Today, despite the strong emphasis on installing distributed renewable generation sources to supply part of the distribution network's load, there is a possibility of load reduction at certain nodes. Therefore, in addition to the inherent uncertainty of electric load, the uncertainty introduced by renewable energy sources also affects distribution substation planning.

Furthermore, in recent years, distribution networks have encountered new loads with uncertainty in location and magnitude. The increasing penetration of electric vehicles (EVs) has significantly altered the conditions and requirements of distribution networks. Therefore, special attention must be given to load uncertainty and its impact on the location, capacity, and service area of distribution substations.

The IGDT method is applied in this study to provide a solution that is robust against load uncertainty. The objective is to determine the maximum horizon of load variation while ensuring that the additional cost of substation placement does not exceed an allowable limit.

The base value of the objective function, denoted as Z_{base}, is first calculated in a deterministic framework using the forecasted values of load points. The maximum allowable load variations are then determined using Equations 12 to 17. A risk-averse policy is adopted to ensure that the solution remains valid despite unfavorable deviations in load.

$$\max \alpha \quad \text{Subject to:} \quad (12)$$

$$0 \le \alpha \le 1 \quad (13)$$

$$(1-\alpha)\bar{S}_j \le \tilde{S}_j \le (1+\alpha)\bar{S}_j \quad (14)$$

$$\{\min_{\max \alpha} C_{total}\} \le (1+\beta) \times Z_{base} \quad (15)$$

$$\sum_{j=1}^{n_i} \beta_{ij} \times \tilde{S}_j \le Z_{cap} \times Cap_{n_i} \quad (16)$$

$$0 \le \frac{D_{ij} \times \tilde{S}_j}{V^2} \times R \le \Delta V_{\max} \quad (17)$$

As seen in Equation 15, the objective function of the deterministic problem (Z_{base}) is incorporated as a limiting condition in the risk-averse strategy. Specifically, the highest level of load points is determined while ensuring that the total cost does not exceed β% more than Z_{base}.

The above problem is a bi-level optimization problem that can be transformed into a single-level problem. Since the maximum risk level and the objective function value occur at the highest level of load uncertainty, the values of load points can be defined using Equation 18.

$$\tilde{S}_j = (1+\alpha)\bar{S}_j \quad (18)$$

This transformation simplifies the problem by considering the worst-case scenario, eliminating the second level, and reducing it to a single-level problem as formulated in Equations 19 to 23.

$$\max \alpha \quad \text{Subject to:} \quad (19)$$

$$0 \le \alpha \le 1 \quad (20)$$

$$\{\min_{\max \alpha} C_{total}\} \le (1+\beta) \times Z_{base} \quad (21)$$

$$0 \le \frac{D_{ij} \times ((1+\alpha)\bar{S}_j)}{V^2} \times R \le \Delta V_{\max} \quad (22)$$

$$\sum_{j=1}^{n_i} \beta_{ij} \times ((1+\alpha)\bar{S}_j) \le Z_{cap} \times Cap_{n_i} \quad (23)$$

6. OPTIMIZATION ALGORITHM

Genetic algorithm is an evolutionary algorithm that has been used in many optimization studies (31-33). Sabattin et al. (31) introduced a parallel genetic algorithm to optimize a tree-based topology for large-scale electric power distribution networks. Mahmoodabadi et al. (32) solved AC/DC hybrid distribution system planning using genetic algorithm and Esmaeilzadeh et al. (33) introduced a new optimization algorithm based on particle swarm optimization, genetic algorithm, and sliding surfaces.

In this study, genetic algorithm has been used to solve the optimization problem. A two-dimensional chromosome structure is considered. In the base case study (deterministic problem), the number of columns corresponds to the number of candidate locations for distribution substations, while the number of rows represents the number of load points. Figure 1 illustrates the structure of the two-dimensional chromosome used in this study. Each element (gene) in the matrix can take values of either 0 or 1.

If the element in the i^{th} row and j^{th} column is equal to 1, it indicates that the i^{th} load point is supplied by the j^{th} substation. Given the radial structure of the system, each load center can only be supplied by one substation, ensuring that each row contains only one element with a value of 1.

For the IGDT-based problem, an additional column has been added to the chromosome (highlighted in gray

in Figure 1) to represent the uncertainty horizon (α) for each load point.

Using this chromosome representation, the location, size, and service area of distribution substations can be determined. Any column with at least one element equal to 1 indicates a selected candidate location for substation installation. The service area is identified by determining which load points are connected to the selected substation. The required capacity is then chosen as the closest standard substation capacity to accommodate the total connected load. The substation load is calculated as the sum of the loads from all connected load points.

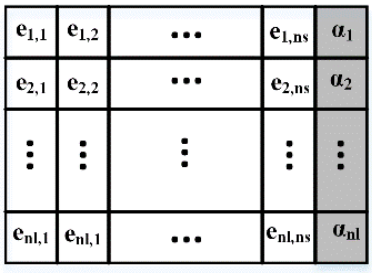


Figure 1. Proposed chromosome structure

7. TEST SYSTEM AND SIMULATION RESULTS

To evaluate the proposed model and verify its performance, a test system has been designed. This system consists of 36 load points and 14 candidate substation locations. The study area is assumed to be divided into 36 sub-areas, and load forecasting results for the horizon year are provided. Coordinates of the candidate locations for the distribution substation are provided in Table 1. The information regarding load points is available in Table 2, while the remaining required data can be found in Table 3. Three standard substation capacities of 30, 45, and 60 MVA are considered for all candidate locations. The construction costs for these substations are \$2 million, \$2.5 million, and \$3 million, respectively, covering the costs of substation infrastructure, buildings, and other necessary components.

In this study, the maintenance costs of equipment are assumed to be 3% of the initial investment cost. The simulations are based on a static planning approach, and the long-term planning period is set to 10 years.

The accuracy of the proposed method has been investigated using a test system introduced by Haghifam and Shahabi [24]. The study of substation allocation is performed for the following cases:

- **Case 1**: Deterministic substation allocation
- **Case 2**: Probabilistic substation allocation with a uniform uncertainty horizon
- **Case 3**: Probabilistic substation allocation with a variable uncertainty horizon

7. 1. Deterministic Substation Allocation-Case 1

In this case, it is assumed that there is no uncertainty in the load point values, and the study is conducted using predicted values. The loading factor of substations is set at 0.85. Figure 2 presents the results. The total planned cost is \$50.7 million, and eight locations have been selected for substation installation. Table 4 presents the selected substations' locations, sizes, and estimated loadings for the horizon year.

TABLE 1. Load point coordinates information

Load point	Load (MVA)	Coordinates		Load point	Load (MVA)	Coordinates		Load point	Load (MVA)	Coordinates	
		X(km)	Y(km)			X(km)	Y(km)			X(km)	Y(km)
1	10	2	2	13	10	3	13	25	10	4	24
2	10	7	3	14	6	7	12	26	10	۶6	21
3	10	12	4	15	2	13.5	12.5	27	5	14	22
4	5	17	1	16	6	19	11	28	5	17.5	23
5	5	21	3	17	2	22	14	29	6	22	21
6	5	28	2	18	5	28.5	13.5	30	5	26	23
7	10	2.5	7	19	10	2	16	31	5	2	26
8	6	6	9	20	10	7	17	32	10	8	27
9	6	11	8	21	1	14	17.5	33	10	12	28
10	1	16	9	22	6	18	19	34	10	18	29
11	5	24	6	23	5	21	16	35	5	24	26
12	5	27	6	24	5	27	18	36	6	28	27

TABLE 2. Coordinates information of Candidate points

Candidate point	Coordinates		Candidate point	Coordinates	
	X(km)	Y(km)		X(km)	Y(km)
1	5	6	8	28	22
2	4	11	9	3	36
3	11	10	10	9	20
4	12	2	11	11	29
5	16	5	12	21	23
6	20	12	13	16	16
7	25	8	14	27	19

TABLE 3. Required parameters information (2, 34)

Life time (Yr)	Power Loss Cost ($/kW)	Energy Loss Cost ($/kWh)
30	300	0.17
Inflation Rate	**Interest Rate**	**Loss Factor**
0.15	0.07	0.36
Maximum allowed voltage drop (%)	**Feeder construction cost ($/Km)**	**Resistance (Ω/Km)**
5	3000	0.0695

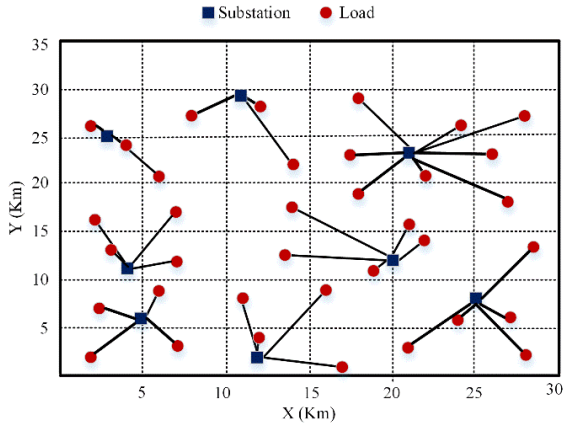


Figure 2. Substation allocation results for loading factor of 0.85 without uncertainty

7. 2. Probabilistic Substation Allocation-Common Uncertainty Horizon- Case 2 In this scenario, the uncertainty is assumed to be the same across all load points. The problem is solved to determine the maximum uncertainty horizon (α) for three different values of the β coefficient: 0.2, 0.3, and 0.5. In all cases, the maximum allowable substation loading factor is set to 85%.

When β=0.2, the maximum cost incurred by the planner to account for uncertainty is 20% higher than the base case value (Z_{base}). Numerical results indicate that a confidence interval of 12% is formed for each load point value. This means that by accepting a 20% increase in total cost, the results remain valid for up to a 12% increase in load values. The selected substations and their service areas are shown in Figure 3. Additionally, substation characteristics are provided in Tables 5.

TABLE 4. Selected substation features in Case 1

Selected substations	Capacity (MVA)	Loaded capacity (MVA)	Loading percentage (%)
1	45	36	80
2	45	36	80
4	30	22	73.3
6	30	16	53.3
7	30	25	83.3
9	30	25	83.3
11	30	25	83.3
12	60	48	80

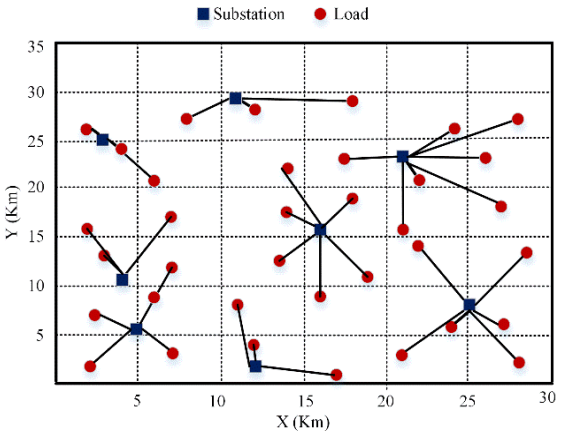


Figure 3. Substation allocation results for Case 2, β=0.2

TABLE 5. Selected substations features in Case 2, β=0.2

Selected substations	Capacity (MVA)	Loaded capacity (MVA)	Loading percentage (%)
1	60	47.04	78.40
2	45	33.60	74.67
4	30	23.52	78.40
7	45	30.24	67.20
9	45	28.00	62.20
11	45	33.60	74.70
12	60	41.44	69.07
13	30	23.52	78.40

In this case, 99.03% of the allowable budget covers the maximum possible uncertainty. When β=0.3, the uncertainty horizon (α) increases to 23%, utilizing 99.99% of the allowable budget. When β=0.5, α reaches 41%, and 98.06% of the allowable budget is used.

Figures 4 and 5 illustrate the selected substations and their service areas. Tables 6 and 7 present the characteristics of the selected substations for the cases where β equals 0.3 and 0.5, respectively. Cost details of case2 expriments are provided in Table 8.

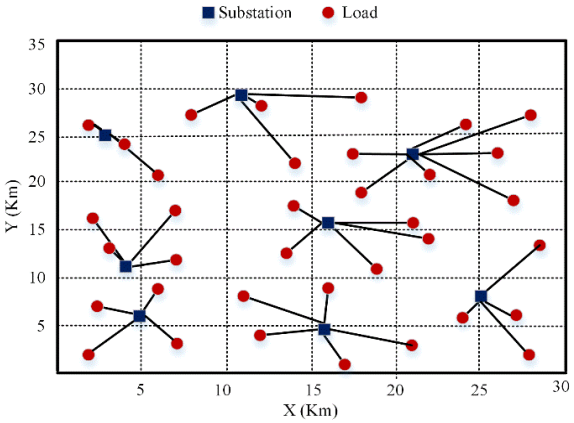


Figure 4. Substation allocation results for Case. 2, β=0.3

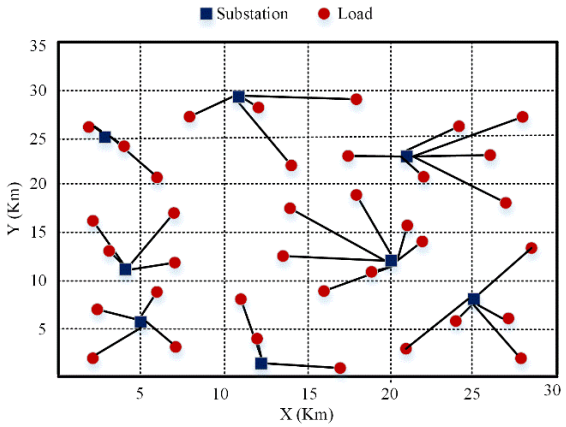


Figure 5. Substation allocation results for Case. 2, β=0.5

TABLE 6. Selected substation features in Case 2, β=0.3

Selected substations	Capacity (MVA)	Loaded capacity (MVA)	Loading percentage (%)
1	60	44.28	73.8
2	60	44.28	73.8
5	45	33.21	73.8
7	30	24.60	82.0
9	45	30.75	68.3
11	60	43.05	71.8
12	60	46.74	77.9
13	30	19.68	65.6

TABLE 7. Selected substation features in Case 2, β=0.5

Selected substations	Capacity (MVA)	Loaded capacity (MVA)	Loading percentage (%)
1	60	44.28	84.6
2	60	44.28	84.6
4	45	33.21	65.8
6	45	24.60	72.1
7	45	30.75	78.3
9	45	43.05	78.3
11	60	46.74	82.3
12	60	19.68	75.2

TABLE 8. Planning cost details in case 2

β	0.2	0.3	0.5
Investment cost (M$)	20.501	21.483	22.490
Maintenance cost (M$)	10.654	11.165	11.688
Loss cost (M$)	29.092	33.266	40.396
Total cost (M$)	60.247	65.914	74.575

7. 3. Probabilistic Substation Allocation-Variable Uncertainty Horizon- Case 3 In this scenario, the substation allocation problem is solved under a risk-averse approach. The uncertainty horizons of the load points vary and are determined through the optimization process. The aim is to determine the maximum allowable increase in load points, allowing the scheme's cost to be up to β% higher than the base case (Z_{base}). The study is conducted for β values of 0.2, 0.3, and 0.5. The maximum allowable substation loading factor remains at 85%.

The characteristics of the selected substations are listed in Tables 9, 10, and 11, while Table 12 provides a detailed breakdown of the costs. The selected distribution substations and their service areas for case 3, β= 0.2 are as case 2, β= 0.5. Substation allocation results for case3, and β of 0.3 and 0.5 are depicted in Figures 6, and 7.

Figure 8 illustrates the α values for different load points across the three cases studied in Case 3. As shown, the α values vary significantly, with some load points exhibiting a much higher allowable increase than others. The average α values for β= 0.2, β=0.3, and β=0.5 are 16.42%, 27.47%, and 42.88%, respectively. The results indicate that uncertainty values are anticipated to rise as β increases.

TABLE 9. Selected substation features in Case 3, β=0.2

Selected substations	Capacity (MVA)	Loaded capacity (MVA)	Loading percentage (%)
1	45	32.40	72.0
2	60	49.74	82.9
4	30	24.55	81.8
6	30	18.79	62.6
7	45	30.91	68.7
9	45	27.80	61.8
11	45	34.40	76.4
12	60	50.97	84.9

TABLE 10. Selected substation features in Case 3, β=0.3

Selected substations	Capacity (MVA)	Loaded capacity (MVA)	Loading percentage (%)
1	60	43.60	72.6
2	60	50.32	83.9
4	30	21.50	71.7
6	30	24.98	83.3
7	30	25.20	84.0
9	45	31.50	76.7
11	30	24.80	82.7
12	60	47.06	78.4
14	30	22.90	76.3

TABLE 11. Selected substation features in Case 3, β=0.5

Selected substations	Capacity (MVA)	Loaded capacity (MVA)	Loading percentage (%)
1	60	46.38	77.3
2	60	50.04	83.4
5	30	22.60	75.3
7	45	37.80	84.0
9	30	23.35	77.8
10	45	28.10	62.4
11	60	49.25	82.1
12	45	38.07	84.6
13	45	35.82	79.6

TABLE 12. Planning cost details in case 3

β	0.2	0.3	0.5
Investment cost (M$)	20.493	21.984	23.485
Maintenance cost (M$)	10.650	11.425	12.205
Loss cost (M$)	29.698	32.471	40.331
Total cost (M$)	60.841	65.881	76.022

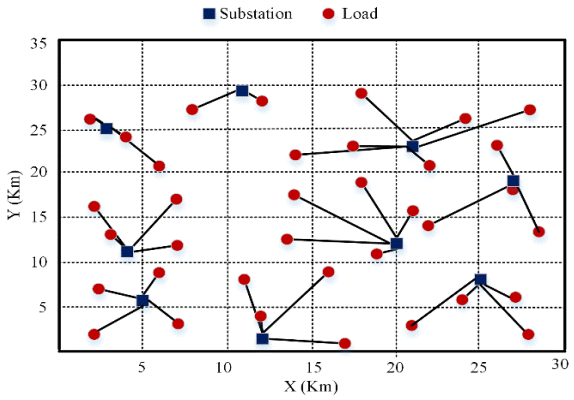


Figure 6. Substation allocation results for Case. 3, β=0.3

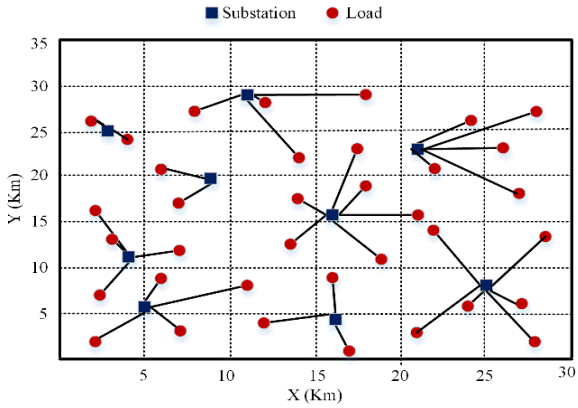


Figure 7. Substation allocation results for Case. 3, β=0.5

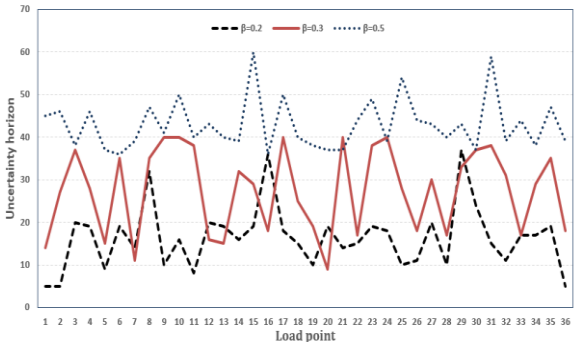


Figure 8. Uncertainty horizons (α) of load points in cases 3

To compare the proposed plans under deterministic and uncertain conditions, a study with β fixed at 0.5 and fixed α (Case 2- β=0.5) is compared with the base scenario (Case 1). When β is 0.5, 50% more financial resources are allocated to the proposed solution compared to the base scenario. In this case, the studies are conducted to cover the maximum uncertainty arising from load. However, if uncertainty does not materialize, in addition to the significant cost imposed on the system, many substations will operate at a low loading level.

Table 13 provides some suitable results in these cases. By considering perfect capacities for the selected substations of Case 2- β =0.5, 18 M$ would be saved. If the system encounters uncertainty, 76.65% of the substation capacity will be loaded in the design horizon year. But if uncertainty does not occur, the average transformer load will be 55% of their capacity, which indicates an unrealistic design. Therefore, before using this method to cover uncertainty, all influencing conditions such as social, economic, and political conditions must be considered.

TABLE 13. Comparing final solutions of case 2-β=0.5 with and without uncertainty

Selected substations	* Selected Capacity (MVA)	* Loading percentage (%)	** Loading percentage (%)	** Perfect Capacity (MVA)
1	60	84.6	60	45
2	60	84.6	60	45
4	45	65.8	46.7	30
6	45	72.1	51.1	30
7	45	78.3	55.55	30
9	45	78.3	55.55	30
11	60	82.3	58.33	45
12	60	75.2	53.31	45

*: Case 2- β=0.5
**: Not materializing of uncertainties
***: Perfect capacity proportional to substation loading in case of not materializing of uncertainties

8. CONCLUSION

Results indicate that an average load increment of approximately 42.88% can be accommodated with a corresponding 50% increase in permissible cost. In such scenarios, if potential load uncertainty does not materialize, the average substation load would be approximately 55% of their nominal capacities.

The findings reveal a clear correlation between increased permissible total cost and enhanced robustness against load uncertainty. Case 1 do not tolerate uncertainty by spending 50.7 M$ while case 3, β=0.5 can accommodate 42.88% uncertainty in average by spending M$ 76.022 for the provided solution.

The robustness is achieved through adjustments to substation's quantity, service area, and capacity, necessitating additional capital investment. Notably, in nearly all scenarios, the optimal selected substations capacities or service areas differ, demonstrating the increment substation deployment in vase of increasing budgetary allocation. The installed substation capacity reaches to 420 MVA in case 3, β=0.5 from 300 MVA in Case 1 without uncertainty. Also selected substation locations and service areas are considerably different.

When the uncertainty horizons fluctuated during the optimization process in Case 3, the final uncertainty values at load points exhibited significant variability. Specifically, the maximum α value was approximately seven times greater than the minimum α value, highlighting the non-linear relationship between β and α. Minimum and maximum load uncertainty horizons in case 3, β=0.5 are 36 and 60% while minimum and maximum load uncertainty horizons in case 3, β=0.5 are 5 and 37%.

The primary limitation of employing a fixed uncertainty horizon lies in its inflexibility. This constraint is effectively mitigated by adopting a variable uncertainty horizon, thereby enhancing the method's applicability to real-world scenarios. Conversely, the constant uncertainty horizon offers the benefit of a simplified design process, resulting in reduced computational time.

However, the risk-averse IGDT approach, while robust, suffers from a significant drawback: its conservative nature. By prioritizing the worst-case scenario, this methodology risks substantial financial losses if the actual uncertainty deviates from the assumed direction or if no uncertainty materializes. Therefore, a careful balance between robustness and cost-effectiveness must be considered when implementing this approach. In such scenarios, if potential load uncertainty does not materialize, the average substation load would be approximately 55% of their nominal capacities.

Possible directions for future research include: modeling uncertainties in additional parameters, such as the geographical coordinates of load points; performing simultaneous optimization of substation placement and medium-voltage feeder routing; and extending the problem formulation to incorporate considerations of a power market environment.

9. REFERENCES

1. Moradi SST. A probabilistic study on hole cleaning optimization. Записки Горного института. 2022;258:956-63. https://doi.org/10.1016/j.egypro.2018.11.109
2. Temraz H, Salama M. A planning model for siting, sizing and timing of distribution substations and defining the associated service area. Electric Power Systems Research. 2002;62(2):145-51. https://doi.org/10.1016/S0378-7796(02)00040-8
3. Mazhari SM, Monsef H, Falaghi H. A hybrid heuristic and learning automata-based algorithm for distribution substations siting, sizing and defining the associated service areas. International Transactions on Electrical Energy Systems. 2014;24(3):433-56. https://doi.org/10.1002/etep.1716
4. Alipour MA, Askarzadeh A. An efficient optimization framework for distribution network planning by simultaneous allocation of photovoltaic distributed generations and transformers. IET Renewable Power Generation. 2024;18(2):153-68. https://doi.org/10.1049/rpg2.12910

5. Behzadi S, Bagheri A. A convex micro-grid-based optimization model for planning of resilient and sustainable distribution systems considering feeders routing and siting/sizing of substations and DG units. Sustainable Cities and Society. 2023;97:104787. https://doi.org/10.1016/j.scs.2023.104787
6. Masud E. An interactive procedure for sizing and timing distribution substations using optimization techniques. IEEE Transactions on Power Apparatus and Systems. 2007(5):1281-6. https://doi.org/10.1109/TPAS.1974.293851
7. El-Fouly T, Zeineldin H, El-Saadany E, Salama M. A new optimization model for distribution substation siting, sizing, and timing. International Journal of Electrical Power & Energy Systems. 2008;30(5):308-15. https://doi.org/10.1016/j.ijepes.2007.10.002
8. Ravadanegh SN, Roshanagh RG. A heuristic algorithm for optimal multistage sizing, siting and timing of MV distribution substations. Electric power systems research. 2013;105:134-41. https://doi.org/10.1016/j.epsr.2013.07.019
9. Kiani Rad H, Moravej Z. Sub-transmission sub-station expansion planning based on bacterial foraging optimization algorithm. Journal of AI and Data Mining. 2017;5(1):11-20. https://doi.org/10.22044/jadm.2016.734
10. Mazhari SM, Monsef H, Romero R. A hybrid heuristic and evolutionary algorithm for distribution substation planning. IEEE Systems Journal. 2013;9(4):1396-408. https://doi.org/10.1109/JSYST.2013.2290973
11. Yu L, Shi D, Guo X, Jiang Z, Xu G, Jian G, et al. An efficient substation placement and sizing strategy based on GIS using semi-supervised learning. CSEE journal of power and energy systems. 2018;4(3):371-9. https://doi.org/10.17775/CSEEJPES.2017.00800
12. Wang S, Lu Z, Ge S, Wang C. An improved substation locating and sizing method based on the weighted voronoi diagram and the transportation model. Journal of Applied Mathematics. 2014;2014(1):810607. https://doi.org/10.1155/2014/810607
13. Salyani P, Salehi J, Gazijahani FS. Chance constrained simultaneous optimization of substations, feeders, renewable and non-renewable distributed generations in distribution network. Electric Power Systems Research. 2018;158:56-69. https://doi.org/10.1016/j.epsr.2017.12.032
14. Navidi M, Moghaddas Tafreshi SM, Anvari-Moghaddam A. Sub-transmission network expansion planning considering regional energy systems: a Bi-level approach. Electronics. 2019;8(12):1416. https://doi.org/10.3390/electronics8121416
15. Jalali M, Zare K, Tarafdarhagh M. Timing, Siting, and Sizing of Sub-transmission Substations and Distributed Generation Units Along Planning Horizon. Electric Power Components and Systems. 2016;44(10):1111-25. https://doi.org/10.1080/15325008.2016.1157107
16. Shayeghi H, Bagheri A. Dynamic sub-transmission system expansion planning incorporating distributed generation using hybrid DCGA and LP technique. International Journal of Electrical Power & Energy Systems. 2013;48:111-22. https://doi.org/10.1016/j.ijepes.2012.11.029
17. Ramezani M, Falaghi H, Moghaddam MP, Haghifam M-R, editors. Genetic based algorithm for optimal placement of distribution transformers. 2006 IEEE Power Engineering Society General Meeting; 2006: IEEE.
18. Abedi MH, Hosseini H, Jalilvand A. Sub-transmission substation expansion planning considering load center uncertainties of size and location. International Journal of Electrical Power & Energy Systems. 2019;109:413-22. https://doi.org/10.1016/j.ijepes.2019.02.032
19. Vahedi S, Banejad M, Assili M. GIS-based substation expansion planning. IEEE Systems Journal. 2020;15(1):959-70. https://doi.org/10.1109/JSYST.2020.3010565
20. Bosisio A, Berizzi A, Merlo M, Morotti A, Iannarelli G. A GIS-based approach for primary substations siting and timing based on Voronoi diagram and particle swarm optimization method. Applied Sciences. 2022;12(12):6008. https://doi.org/10.3390/app12126008
21. Esmaeeli M, Kazemi A, Shayanfar H, Haghifam M-R. Multistage distribution substations planning considering reliability and growth of energy demand. Energy. 2015;84:357-64. https://doi.org/10.1016/j.energy.2015.03.002
22. Sun L, You S, Hu J, Wen F. Optimal allocation of smart substations in a distribution system considering interruption costs of customers. IEEE Transactions on Smart Grid. 2016;9(4):3773-82. https://doi.org/10.1109/TSG.2016.2642163
23. Haghifam M, Shahabi M. Optimal location and sizing of HV/MV substations in uncertainty load environment using genetic algorithm. Electric Power Systems Research. 2002;63(1):37-50. https://doi.org/10.1016/S0378-7796(02)00087-1
24. Esmaeeli M, Kazemi A, Shayanfar H, Haghifam M-R, Siano P. Risk-based planning of distribution substation considering technical and economic uncertainties. Electric Power Systems Research. 2016;135:18-26. https://doi.org/10.1016/j.epsr.2016.03.013
25. Franco JF, Rider MJ, Romero R. Robust multi-stage substation expansion planning considering stochastic demand. IEEE Transactions on Power Systems. 2015;31(3):2125-34. https://doi.org/10.1109/TPWRS.2015.2443175
26. Sarhadi S, Amraee T. Robust dynamic network expansion planning considering load uncertainty. International Journal of Electrical Power & Energy Systems. 2015;71:140-50. https://doi.org/10.1016/j.ijepes.2015.02.043
27. Vahid-Ghavidel M, Shafie-khah M, Javadi MS, Santos SF, Gough M, Quijano DA, et al. Hybrid IGDT-stochastic self-scheduling of a distributed energy resources aggregator in a multi-energy system. Energy. 2023;265:126289. https://doi.org/10.1016/j.energy.2022.126289
28. Tostado-Véliz M, Kamel S, Hasanien HM, Turky RA, Jurado F. Uncertainty-aware day-ahead scheduling of microgrids considering response fatigue: An IGDT approach. Applied Energy. 2022;310:118611. https://doi.org/10.1016/j.apenergy.2022.118611
29. Liu J, Chen C, Liu Z, Jermsittiparsert K, Ghadimi N. An IGDT-based risk-involved optimal bidding strategy for hydrogen storage-based intelligent parking lot of electric vehicles. Journal of Energy Storage. 2020;27:101057. https://doi.org/10.1016/j.est.2019.101057
30. Moradi-Dalvand M, Mohammadi-Ivatloo B, Amjady N, Zareipour H, Mazhab-Jafari A. Self-scheduling of a wind producer based on Information Gap Decision Theory. Energy. 2015;81:588-600. https://doi.org/10.1016/j.energy.2015.01.002
31. Sabattin J, Fuertes G, Alfaro M, Quezada L, Vargas M. Optimization of large electric power distribution using a parallel genetic algorithm with dandelion strategy. Turkish Journal of Electrical Engineering and Computer Sciences. 2018;26(5):2648-60. https://doi.org/10.3906/elk-1801-26
32. Mahmoodabadi M, Nematia A, Daneshb N. A New Optimization Algorithm Based on Particle Swarm Optimization Genetic Algorithm and Sliding Surfaces. International Journal of Engineering Transactions C: Aspects. 2024;37(9):1716-35. https://doi.org/10.5829/ije.2024.37.09c.02
33. Esmaeilzadeh Z, Falaghi H, Ramezani M. Risk-Based AC/DC Hybrid Distribution System Planning. International Journal of Engineering, Transactions B: Applications. 2024;37(5):876-86. https://doi.org/10.5829/ije.2024.37.05b.06
34. Seifi H, Sepasian MS. Electric power system planning: issues, algorithms and solutions. Power Systems. 2011;49.

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

این مقاله یک روش نوین برای بهینه‌سازی مکان، اندازه و ناحیه سرویس‌دهی پست‌های توزیع، با در نظر گرفتن عدم قطعیت بار، ارائه می‌کند. حضور رو به افزایش تولیدات پراکنده تجدیدپذیر و خودروهای الکتریکی، عدم قطعیت‌های قابل توجهی را به بارهای شبکه توزیع وارد کرده است که از خطاهای پیش‌بینی بار مرسوم فراتر می‌رود. برای کاهش این چالش، یک چارچوب بهینه‌سازی مقاوم به کار گرفته می‌شود که مدل‌سازی ریسک را ممکن ساخته و عملکرد ایمن را در سناریوهای عدم قطعیت شدید تضمین می‌کند. نظریه تصمیم‌گیری شکاف اطلاعات (IGDT) برای مدل‌سازی موثر عدم قطعیت بار به کار گرفته می‌شود، یک استراتژی ریسک‌گریز برای افزایش مقاومت در برابر عدم‌قطعیت اتخاذ می‌شود. هدف، حداقل کردن هزینه کل برنامه‌ریزی سیستم توزیع، شامل هزینه‌های سرمایه‌گذاری و نگهداری پست‌های توزیع و هزینه‌های مرتبط با تلفات، ضمن رعایت محدودیت‌های فنی است. در ابتدا، مسئله جایابی پست‌های توزیع با استفاده از مقادیر میانگین بار حل می‌شود. سپس، یک رویکرد IGDT ریسک‌گریز برای شناسایی راه حل‌های مقاوم در برابر عدم قطعیت بار اعمال می‌شود. اثربخشی روش پیشنهادی با استفاده از یک سیستم آزمون نشان داده می‌شود. نتایج نشان می‌دهد که با افزایش ۵۰ درصدی هزینه مجاز، می‌توان به طور متوسط افزایش بار حدود ٤٢/٨٨ درصد را مدیریت کرد. در چنین سناریوهایی، اگر عدم قطعیت بار احتمالی محقق نشود، میانگین بار پست حدود ۵۵ درصد از ظرفیت اسمی آنها خواهد بود.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1101-1114

International Journal of Engineering

Journal Homepage: www.ije.ir

Sustainable High-performance Self-compacting Concrete with Multi Blended of Materials Powders: Fresh and Hardened Properties

R. K. Khoman*, H. M. Owaid

Civil Engineering Department, College of Engineering, University of Babylon, Babylon, Iraq

PAPER INFO

Paper history:
Received 09 May 2025
Received in revised form 23 May 2025
Accepted 07 June 2025

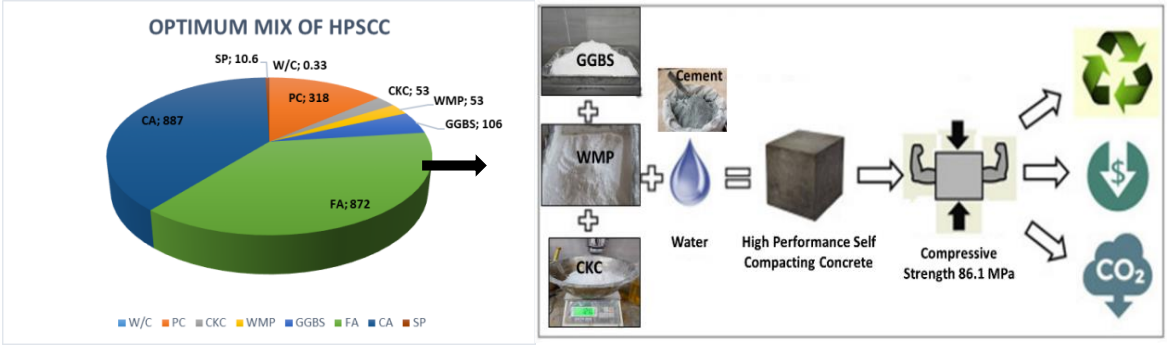
Keywords:
High-performance Self-consolidating Concrete
Calcined Kaolin Clay
Waste Marble Powder
Fresh-state Properties
Hardened-State Properties
Ground Granulated Blast Furnace Slag

ABSTRACT

High-performance self-consolidating concrete represents a significant advancement in construction technology. This study examined and compared the fresh-state properties (slump flow (D (mm), V-Funnel flow time, and L-box tests), hardened-state properties (unit weight and compressive strength), and microstructure characteristics (Scanning Electron Microscopy (SEM)) of high-performance self-consolidating concrete (HPSCC) with a reference mix (M0) as the control mix containing only Ordinary Portland Cement (OPC) without pozzolanic substitutions. Two blend systems were developed in this study: a binary blend system and a quaternary blend system, which included (OPC+CKC, OPC+WMP, OPC+GGBS, and OPC+CKC+WMP+GGBS). The results for the fresh properties of all HPSCC mixtures fell within the acceptable ranges set by EFNARC, with no segregation or bleeding observed, and all HPSCC mixtures demonstrated superior performance than the control mixture. From this study, it can be concluded that using 60% OPC + 10% CKC + 10% WMP + 20% GGBS yielded improved fresh, compressive strength and denser microstructure compared to control mix (M0), making it suitable for use as HPSCC. By cutting clinker content 40 %, the optimal quaternary blend lowers the binder's embodied CO_2 by roughly 30 %, offering a tangible environmental benefit while still achieving a 90-day compressive strength of 86.1 MPa.

***doi**: 10.5829/ije.2026.39.05b.05*

Graphical Abstract



NOMENCLATURE

OPC	Ordinary Portland cement	W/B	Water to binder ratio
CKC	Calcined kaolin clay	SP	Superplasticizer
GGBS	Ground granulated blast furnace slag	FA	Fine aggregate
WMP	Waste marble powder	CA	Coarse aggregate

*Corresponding Author Email: ruwaida.jabbar.engh354@student.uobabylon.edu.iq (R. K. Khoman)

Please cite this article as: Khoman RK, Owaid HM. Sustainable High-performance Self-compacting Concrete with Multi Blended of Materials Powders: Fresh and Hardened Properties. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1101-14.

1. INTRODUCTION

Self-compacting concrete (SCC) is a modern type of concrete known for its capability to flow and self-consolidate under its own weight, eliminating the necessity for mechanical vibration. It can fill formwork with heavy reinforcement while maintaining consistent homogeneity, preventing any migration or separation of its large components (1). The use of SCC addresses the problem of unskilled labor in construction, particularly in underwater concreting and heavily reinforced structures. The necessity for vibration is obviated, construction timelines are expedited, labor expenditures are diminished, and noise pollution is curtailed (2, 3). SCC is distinguished by a substantial amount of cement, a minimal quantity of coarse aggregates, and the incorporation of superplasticizers to reduce the water-to-binder ratio. High-performance concrete (HPC) is delineated as concrete that is meticulously formulated, mixed, placed, consolidated, and cured to attain elevated strength and diminished transport properties, thereby ensuring enhanced durability. It demonstrates remarkable resistance to segregation (4). The demand for HPC has experienced a consistent upward trajectory owing to its remarkable mechanical and durability attributes (5, 6). HPC can be engineered to manifest elevated workability, mechanical properties, and improved durability (7, 8). Innovations in concrete technology have culminated in the advent of a novel generation of concrete predicated on the principles of SCC and HPC, which is designated as High Performance Self-Compacting Concrete (HPSCC).

HPSCC is described as a type of SCC that exhibits great compressive strength and enhanced durability. Unlike conventional concrete, the mix design of HPSCC is characterized by higher cement content, the use of superplasticizers, and the incorporation of reactive materials like silica fume. As a result, HPSCC can effectively fill forms containing densely packed steel reinforcement and achieves self-leveling without the need for mechanical compaction, leading to exceptional strength and durability (9, 10). Due to the significant demand for cement in HPSCC production, it is essential to utilize alternative materials to partially replace cement, contributing to environmentally sustainable construction practices by reducing the carbon footprint. Currently, supplementary cementitious materials e.g. ground granulated blast furnace slag (GGBS), fly ash, silica fume, metakaolin (MK), marble powder, and limestone powder are increasingly employed in the development of next-generation concrete that is both eco-efficient and environmentally friendly (2, 3, 11-14). Metakaolin, a fine, white, and environmentally benign material, is derived from kaolin clay via a CO_2-free calcination process at 650–900 °C. This abundant natural mineral, primarily composed of silicate (SiO_2 50–55%) and aluminate (Al_2O_3 40–45%), along with minor oxides (Fe_2O_3, TiO_2, CaO, MgO), offers a sustainable alternative to cement. Research studies conducted by Vejmelková et al. (15) and Kavitha et al. (16) have evaluated the effects of metakaolin and ordinary Portland cement (OPC) on SCC by measuring CO_2 emissions, and have shown that metakaolin provides environmental benefits during its production. Additionally, metakaolin can enhances early-age concrete strength and more effectively fill the pores of the cement paste compared to other fillers. Its incorporation in concrete enhances the mixture's viscosity (17) and accelerates early strength gain owing to its higher pozzolanic reactivity at early-ages (18). In contrast, GGBS is derived from the steel and iron production. It is a non-metallic product mainly composed of aluminosilicates and calcium silicates along with other bases, and is formed in a molten state alongside iron during the blast furnace process (1). GGBS exhibits a chemical structure similar to that of cement and relatively consistent properties. Annually, approximately 250 million tons of GGBS are produced worldwide, of which only 90 million tons are utilized in concrete production (19). GGBS can enhance the hardened and fresh properties of SCC and supports sustainable construction practices. It delays the setting time of concrete, causing strength development over a longer period (20).

Studies have shown that GGBS increases workability and reduces water demand owing to the greater water absorption capacity of its smoother, glossier particles compared to cement particles. Concrete incorporating GGBS exhibits superior ultimate strength in comparison to that produced with Portland cement (PC), owing to an increased formation of strength-enhancing calcium silicate hydrates (CSHs) and a reduction in free lime content, that is not strength-enhancing. Marble waste, on the other hand, is regarded as a severely contaminating type of industrial byproduct, presenting substantial environmental and health hazards (21). The disposal of marble industry waste, which mainly involves of fine-powder, presents numerous environmental challenges (22). Nevertheless, these waste materials offer potential for beneficial and commercial application in improving SCC properties (23, 24). Marble waste is generated during both the early stages of mining and the various machining stages (cutting, polishing, and finishing) (25). According to Tennich et al. (26), marble waste improves the fresh properties of concrete via decreasing plastic contraction. Moreover, the higher calcium content of marble dust (MD) enhances the SCC hardened properties (27).

Few studies in the literature have examined HPSCC and its optimization. However, there is a lack of comprehensive investigations addressing the comparative effects of combining natural and waste-based pozzolanic materials in multi-blend systems for

HPSCC, particularly regarding their role in enhancing mechanical and durability properties under sustainable construction frameworks. HPSCC shows significant promise for a wide range of applications and structural uses. However, realizing this potential necessitates a thorough evaluation of its performance prior to widespread implementation in the construction industry. Existing research highlights developments in both the hardened and fresh properties of HPSCC mixtures achieved through the incorporation of CKC, GGBS, and WMP in binary and quaternary blend systems. Based on this context, the present study offers a thorough analysis of the overall properties of HPSCC and to compare these properties with those of conventional SCC.

2. EXPERIMENTAL PROGRAM

Constituent materials, mix proportion, specimens preparation, and tests are described in detail as follows:

2. 1. Materials All HPSCC mixes used locally sourced Ordinary Portland Cement (OPC)-CEM I from Lafarge, conforming to Iraqi Standard (28), with a specific gravity of 3.12 and a specific surface area (SSA) of 358 m²/kg. Tables 1 and 2 present chemical and physical properties of CKC,WMP and GGBS. Calcined kaolin clay (CKC) was produced by calcining kaolin clay from the Dewekhla region (Al Ramadi desert) at 800 ± 20 °C for 2 hours (heating rate: 5 °C/min), followed by slow cooling at room temperature over 24 hours (29). Following this, the CKC was ground via the air blast method to obtain a reactive material with greater fineness. The CKC's major oxide content (SiO_2 + Al_2O_3 + Fe_2O_3) exceeded 70%, meeting the minimum requirement for Class "N" pozzolans (30). Waste marble powder (WMP), a by-product of marble masonry, served as a supplementary cementitious material. Sludge collected from companies in Al-Hilla city (Iraq) was utilized as WMP industrial waste in this work. This material is produced as a waste by-product during the cutting and shaping of marble. The material contains a high percentage of CaO (85.62%), which is comparable to that of cement, potentially imparting cementing properties to WMP. GGBS is a highly sustainable construction material, offering significant ecological benefits by utilizing a by-product. GGBS production is significantly more energy-efficient and generates substantially fewer CO_2 emissions than PC production (31). GGBS exhibits hydration characteristics similar to those of PC and is ground to a fine cement-like consistency. The GGBS used in this study aligns with ASTM standards (32). It is in the form of a white powder and is incorporated into the mixtures as a partial replacement for cement. Locally sourced sand from the Al-Ukhaidher region served as the fine aggregate, meeting Iraqi standards and falling within the third gradient zone (28), and has a fineness modulus of (2.52). Coarse aggregate consisted of crushed and washed gravel (10 mm maximum size) sourced from Al-Nabai'i and meeting Iraqi standard specifications (28). A high-performance superplasticizer concrete admixture (SikaViscocrete-180GS) was employed; it is based on polycarboxylic ether, serves as an economical high-range water reducer, enhances early and final strength, and maintains workability, specifically developed for the ready-mix concrete industry and suitable for hot weather conditions. This admixture was utilized to achieve the preferred flowability of HPSCC. SikaViscocrete-180GS complies with the requirements of Type F&G upon dosage (33).

2. 2. Mixture Proportions A total of 13 mixtures (Table 3) were meticulously formulated and cast with a water to binder ratio (w/b) of 0.33, accompanied by a comprehensive binder content of 530 kg/m^3. The dosages of the superplasticizer were set at 2% by mass of the cementitious material to meet the performance requirements of SCC as per EFNARC (34). This work developed two blending systems: binary and quaternary, executed in a two-stage process. In the initial stage, the control mix was created using only OPC as the binder, while the other mixtures incorporated binary blend

TABLE 1. Chemical Compositions of CKC,GGBS and WMP

Chemical composition (%)	CKC	GGBS	WMP
SiO2	54.7	34.40	1.51
Al2O3	37.4	13.26	0.84
Fe2O3	1.72	0.43	0.44
CaO	0.84	39.94	85.62
MgO	0.42	8.82	0.54
Na2O	0.37	0.38	1.17
K2O	0.54	0.59	0.08
SO3	0.13	0.3	0.61
P2O5	0.29	-	0.28
TiO2	0.68	0.94	-
MnO	-	0.29	-
Loss on ignition (LOI)	2.91	0.2	8.31

TABLE 2. Physical Properties and Strength Activity Index (SAI) of CKCof CKC,GGBS and WMP

Property	CKC	GGBS	WMP
Specific gravity	2.6	2.9	2.68
Fineness, m^2/kg	1640	418	569
(SAI%), at 7 days	85	98	82
(SAI%), at 28 days	94	117	92

systems (OPC + CKC, OPC + WMP, and OPC + GGBS) where OPC was replaced by 10%, 15%, and 20% of CKC, 10%, 15%, and 20% of WMP, and 20%, 30%, and 40% of GGBS (by mass). The selection of replacement levels was based on previous studies indicating this range as optimal for balancing workability, pozzolanic reactivity, and compressive strength development in SCC. In the subsequent phase, pertaining to the quaternary systems (OPC+CKC+WMP+GGBS), the substitution ratios for CKC were established at (10%, 15%, and 20%) with optimum replacement ratio of WMP and GGBS (10% WMP + 20% GGBS) serving as a partial replacement of the binder (by mass).

2. 3. Preparation of Specimen All materials were used in their processed powder form, not raw. The consistency of the blended powders was ensured by first thoroughly mixing the pozzolanic materials (CKC, WMP, and GGBS) with the OPC to achieve a uniform blend before adding them to the aggregates. The powders were stored in sealed, dry conditions to prevent moisture absorption and maintain their physical properties. All materials used were in processed, finely ground form to enhance reactivity and ensure homogeneity in the concrete mix. The concrete mixtures were prepared for approximately 5 minutes with a rotating planetary mixer yielding a uniform slurry without any sign of agglomeration or segregation. Following the mixing process, tests were performed to assess the characteristics of fresh concrete, after which the mixtures were cast into specimens using standard 100 mm³ cube molds, as illustrated in Figure 1. Following the casting process, the specimens were covered with nylon sheets to maintain a humid environment. These specimens were then stored in a controlled environment within the casting room, maintained at a constant temperature of 25 ± 2°C, for a period of 24 hours. After this initial curing period, the molds were removed from the specimens. Subsequently, to ensure proper curing and preparation for subsequent testing, the specimens were immersed in a water basin at room temperature until the required testing ages were reached.

2. 4. Tests of Fresh Concrete In accordance with EFNARC (34) guidelines, the combination of slump flow, V-funnel, and L-box tests effectively evaluates the key fresh properties of self-compacting concrete—deformability, flowability, and passing ability, respectively. While other tests such as the J-ring provide additional insights, they often overlap in purpose with the L-box and slump flow tests. Thus, selecting these three standard EFNARC tests ensures comprehensive assessment of SCC fresh behavior with efficiency and reliability, avoiding redundant testing.

2. 5. Tests on Hardened Concrete Hardened concrete was evaluated through tests including unit weight (density), compressive strength, and Scanning Electron Microscopy (SEM) analysis. Concrete density (kg/m³) was determined as mass per unit volume (35). To determine the density, the dimensions and weights of concrete cubes, specifically those measuring 100 × 100 × 100 mm, were meticulously recorded prior to the execution of compressive strength assessments at curing intervals of 7 and 28 days. Compressive strength testing, performed according to relevant standards, is a widely accepted measure of concrete's mechanical performance

TABLE 3. Mixes Proportions of HPSCCs

Mix. no	Mix Notation	*W/B*	Quantities of ingredients (kg/m3)						
			OPC	CKC	WMP	GGBS	FA	CA	SP
M0	Control-OPC	0.33	530	-	-	-	872	887	10.6
M1	10%CKC	0.33	477	53	-	-	872	887	10.6
M2	15%CKC	0.33	450.5	79.5	-	-	872	887	10.6
M3	20%CKC	0.33	424	106	-	-	872	887	10.6
M4	10%WMP	0.33	477	-	53	-	872	887	10.6
M5	15%WMP	0.33	450.5	-	79.5	-	872	887	10.6
M6	20%WMP	0.33	424	-	106	-	872	887	10.6
M7	20%GGBS	0.33	424	-	-	106	872	887	10.6
M8	30%GGBS	0.33	371	-	-	159	872	887	10.6
M9	40%GGBS	0.33	318	-	-	212	872	887	10.6
M10	10%CKC+10%WMP+20%GGBS	0.33	318	53	53	106	872	887	10.6
M11	15%CKC+10%WMP+20%GGBS	0.33	291.5	79.5	53	106	872	887	10.6
M12	20%CKC+10%WMP+20%GGBS	0.33	265	106	53	106	872	887	10.6

Figure 1. HPSCC Specimen Preparation: Mixing, Casting, and Curing

Figure 2. Flowability Testing of HPSCC (Slump Flow, V-Funnel, and L-Box Tests)

and overall quality (36). For each mixture and curing age (7, 28, and 90 days), the mean compressive strength has been calculated from three cube specimens with dimensions of 100 × 100 × 100 mm. Throughout the testing procedure, each specimen was subjected to a progressively increasing load in a compression apparatus with a maximum capacity of 1900 kN (refer to Figure 3) until the point of failure was reached. The Strength was calculated as the maximum load divided by the cross-sectional area.

The microstructure of HPSCC specimens that incorporated blended materials powders was analyzed using Scanning Electron Microscopy (SEM). SEM was conducted on mixes (M1, M4, M7, and M10) and compared with the reference mix (M0). Samples for SEM examination were prepared following the methodology outlined according to ASTM (37). The samples were reduced in size using a saw, with 10 × 10 × 10 mm pieces selected from concrete cube cores that exhibited optimal compressive strength at 28 days. The SEM apparatus was set to HV: 20 KV with various magnification factors and high resolution.

Figure 3. Compressive strength test

3. RESULTS AND DISCUSSION

The test consequences for all mixes were reviewed and discussed in both the fresh and hardened states as follows:

3. 1. Fresh Results Table 4 summarizes the fresh-state test results (slump flow, V-funnel,and L-box,). The slump flow test assesses the flowability of freshly mixed concrete without confinement. A larger slump-flow diameter corresponds to greater flowability (or deformability) of the mixture. The outcomes of the testing demonstrate that the fresh characteristics of the manufactured concrete meet acceptance criteria, as illustrated in Figure 4. All concrete mixes met the SF3 class in terms of slump flow, except for the binary mixtures containing 20% CKC, which fell within the SF2 class. Therefore, all the mixes were assumed to have acceptable rheological consistency / uniformity of flow from the filling perspective. It was noted that the presence of 10%, 15%, and 20% of CKC reduced the flow diameter by 2.5%, 4.37%, and 7.5%, respectively. Even though the dosage of SP is 2%, the flow diameter decreased because of the inclusion of CKC. This is likely due to the particle size of CKC, which could be categorized as ultra-fine, and its pozzolanic reactivity, which increased the inter-granular friction. As a result, the elevated friction leads to diminished fluidity (deformability), whereas higher replacement levels of WMP contribute to enhanced flowability (deformability). This improvement is demonstrated by the M4, M5, and M6 mixes, which showed an increase in flowability by 0.63%, 1.25%, and 3.13% compared to M0, respectively. The fine particle size of the marble powder is the main

TABLE 4. Fresh-State Test Results of HPSCCs

Mix.no	Slump flow D(mm)	V-funnel Time(s)	PA
M0	800	6.2	0.948
M1	780	7.5	0.936
M2	765	8.1	0.912
M3	740	9.7	0.874
M4	805	6.5	0.962
M5	810	5.9	0.976
M6	825	5.6	0.98
M7	820	6	0.984
M8	835	5.7	0.991
M9	845	5.3	0.996
M10	800	6.8	0.963
M11	790	7	0.954
M12	780	7.5	0.942

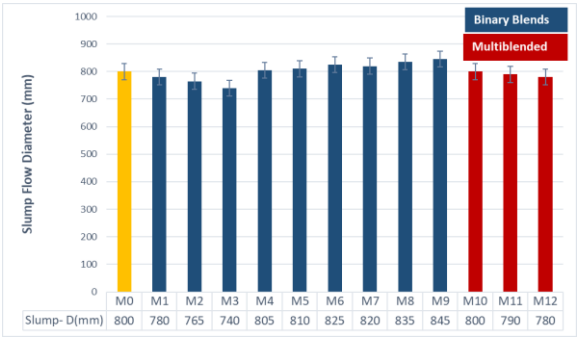


Figure 4. Slump flow diameter results of HPSCCs

reason for this modest improvement, as it leads to denser packing and reduced inter-particle friction (38-40). Additionally, escalating the replacement ratios of GGBS further enhances the mix flowability. This escalation is evidenced by M7, M8, and M9 mixes which increased flowability by 2.5%, 4.37%, and 5.62% in comparison with M0, respectively. The escalation in slump value is because of GGBS's fine, smooth particles. The fine GGBS particles occupy the spaces between the cement and aggregate sand, facilitating improved flow of cement concrete as less water is required for lubrication, resulting in an increased slump value (41). For the multiblended mixes (M11 and M12), it was observed that the slump diameter value decreased by 1.25% and 2.5%, respectively, compared to M0. In contrast, for M10, the slump flow diameter reached 800 mm, which has the same flowability as M0, thereby maintaining the workability of the concrete. When compared to binary blends with CKC, the slump diameter value of the multiblended mixes increased. Utilizing a quaternary blend of CKC, WMP, and GGBS as a partial cement replacement resulted in acceptable fresh-properties for HPSCC. Addition of GGBS to the HPSCC mixtures bettered the fresh characteristics and effectively offset the negative impact of CKC on these properties. These findings provide further evidence of the significant influence that supplementary cement-based materials exert on the rheology of cement-based systems (42). The V-funnel test assessed the HPSCC filling ability (inversely related to viscosity). Longer flow times indicate lower filling ability and higher viscosity. Observed flow times ranged from 5.3 to 9.7 seconds (see Figure 5). Consequently, all the mixes were classified into the VF1 class, except for the 20% CKC, which was classified into the VF2 class according to EFNARC. The V-funnel time for the control mixture M0 was 6.2 s, and it escalated for M3, which reached 9.7 s (the highest funnel time). The reduction in filling ability may be attributed to two factors: the particle size and the increased yield strength of CKC (43). Ultrafine CKC particles, due to their high reactivity, augmented mixture viscosity and accelerated hydration. The V-funnel times

for mixtures M4, M5, M6, M7, M8, and M9 were 6.5 s, 5.9 s, 5.6 s, 6.0 s, 5.7 s, and 5.3 s, respectively. The V-funnel time decreased with escalating waste marble powder quantity, indicating that viscosity diminishes with a higher marble powder ratio. This occurred because of the augmentation of the fines ratio and surface area. The lowest funnel time was recorded for M9, which was 5.3 s. It was detected that escalating the replacement amount of GGBS to a maximum of 40% can enhance the filling capacity by reducing the time required for filling. GGBS possesses larger particle sizes than CKC, which helps to decreases surface area and reactivity, thereby enhancing the concrete's flowability. Additionally, GGBS exhibits reduced yield strength because of reduced internal forces within the mixture, which contributes to increased fluidity of the concrete (1, 19). However, for multiblended mixes, the funnel time ranged from 6.8 to 7.5 seconds, indicating an increase compared to M0. The prospective application of an amalgamation of CKC, WMP, and GGBS as partial cement replacements in a quaternary blend can yield satisfactory fresh properties of HPSCC. These consequences further underscore the significant impact of supplementary cement-based substances on the rheology of cement-based systems (42). An L-box test was performed to evaluate the HPSCC's passing capacity.

The results of the passing ability (PA) produced ranged between 0.874 and 0.996 (see Figure 6), categorizing it as class two (PA2) within the satisfactory limits specified according to ASTM (34). All of the mixtures demonstrated homogeneous characteristics, with no observable segregation or blocking. The PA values measured for M0, M1, M2, and M3 were 0.948, 0.936, 0.912, and 0.874, respectively. The observed decline is because of the inclusion of CKC, which can diminish the passing ability because of reduced fluidity. This may be due to the long, hexagonal shape of the calcined kaolin clay particles being shaped and sized in such a way as to create obstructions within the fresh mixture and escalate inter-particle friction (43). The passing ability produced by mixtures M4, M5, and M6 was 0.962, 0.976, and 0.980, respectively, indicating that the L-box flow escalated with increasing WMP percentages. The higher WMP content improves HPSCC workability by increasing free water, thus enhancing the concrete's filling ability (44, 45). The passing ability for mixtures M7, M8, and M9 was 0.984, 0.991, and 0.996 (the highest passing ability), respectively, demonstrating that the L-box flow escalated with higher percentages of GGBS. The enhanced passing ability may be ascribed to the particle size of GGBS, which positively affects the reduction of internal friction, thereby contributing to enhanced fluidity and passability (1, 19). However, for multiblended mixes (M10 and M11), the passing ability (PA) was increased by 1.58% and 0.63% compared to M0. It is noted that the passing ability of (M12) decreased

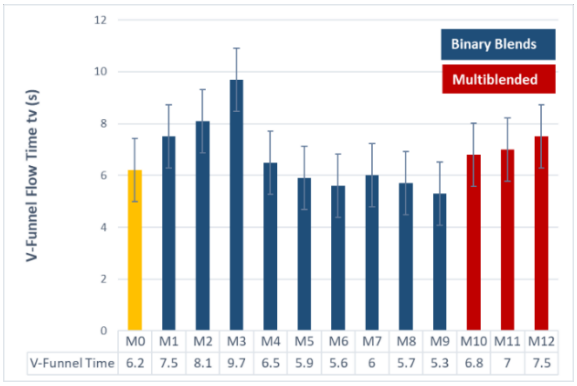


Figure 5. V-Funnel flow time results of HPSCCs

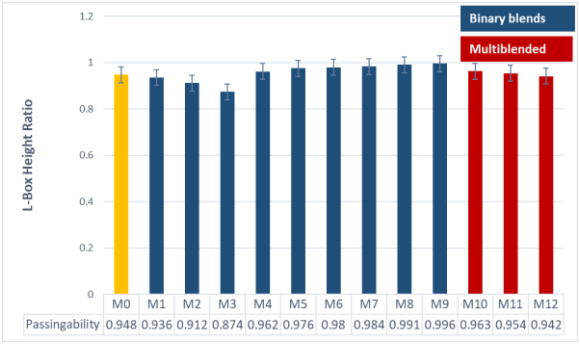


Figure 6. L-box height ratio results of HPSCCs

by 0.63% in comparison with M0. The incorporation of CKC, WMP, and GGBS as partial replacements for cement in a quaternary blend establishes promising fresh properties for HPSCC. In summary, all mixtures exhibited satisfactory workability, stability, and consistency.

3. 2. Hardened Results The self-weight of any structure is entirely dependent on the unit weight of the constituent materials. Therefore, it is a significant parameter for mortar or concrete. The HPSCC mixtures density reduced with increasing pozzolan (CKC, WMP, and GGBS) replacement (see Figure 7). The density of all HPSCC mixes increased with curing age. The percentage decrease in density of HPSCCs (M1-M12) was (0.77, 1.06, 1.42, 0.45, 0.77, 1.38, 0.28, 0.61, 1.22, 1.42, 1.71, and 2.04) % at 28 days, respectively, in comprasion with the control mixture (M0). The density diminished further in multi-blended mixes, with the most significant reduction observed in mixture M12 (20% CKC, 10% WMP, 20% GGBS). This outcome can be ascribed to the reduced specific gravity of the pozzolans relative to cement, resulting in a decreased mass per unit volume (19). Thus, CKC, WMP, and GGBS can be used in mortar or concrete as lighter alternatives to cement. The results are promising in decreasing the self-weight of the structure on soils with poor bearing capacity to some

extent. "Table 5 summarizes the development of compressive strength at 7, 28, and 90 days for binary and multiblended mixes using CKC, WMP, and GGBS respectively." The compressive strength results show that all the concretes samples produced within this investigated study exhibit high performance, with the 28-day compressive strength ranges from 63.7 MPa to approximately 81.2 MPa, while the 90-day compressive strength fell within the range of 69.7 to 86.1 MPa. "CKC, owing to its fine particle size and high contents of reactive SiO_2 and Al_2O_3, offers an extensive specific surface area that accelerates the formation of calcium silicate hydrate (C–S–H) gel, resulting in a noticeable gain in early-age strength. By contrast, GGBS—although similarly rich in silica and alumina—possesses a coarser granulometry, which lowers its effective surface area and slows its initial dissolution. The reduced reaction rate extends the setting time and, consequently, suppresses 7-day compressive strength. However, once hydration progresses, the latent hydraulic activity of GGBS becomes dominant; additional C–S–H and C–A–S–H phases precipitate, progressively densifying the microstructure. This delayed but sustained reaction produces the marked strength increments measured at 28 and 90 days, ultimately allowing the GGBS-containing mixes to surpass the control concrete.The primary benefits of WMP—better particle packing and dilution of the cement matrix—are realised gradually, so any strength enhancement becomes more evident at later ages once the denser microstructure compensates for the initial lack of chemical reactivity." It is worth noting that our findings are in line with those of (46, 47). The observed increase in compressive strength beyond 28 days results from a combination of ongoing hydration of the cementitious materials and pozzolanic reactions of the supplementary powders (CKC, WMP, and GGBS). While OPC hydration continues at a slower rate after 28 days, the pozzolanic materials contribute significantly by reacting with calcium hydroxide (CH) released during cement hydration to form additional calcium silicate hydrate (C–S–H) gel. This secondary C–S–H formation densifies the microstructure, reducing porosity and enhancing mechanical strength over time. Specifically, CKC and GGBS exhibit notable pozzolanic and latent hydraulic activity, respectively, which drives the substantial strength improvements observed at 28 and 90 days. WMP acts primarily as a microfiller, contributing mainly to physical densification rather than chemical reaction.

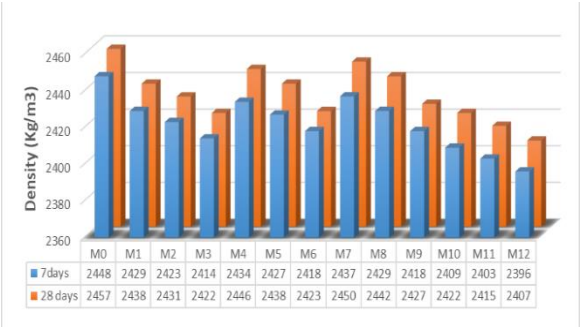


Figure 7. Density results of HPSCCs

TABLE 5. Compressive Strength Results of HPSCCs

Mix.no	7days	28days	90days
M0	53.6	68.7	75.1
M1	56.2	75.4	82.8
M2	55.4	73.6	80.6
M3	54.3	70.8	76.8
M4	52.3	72.6	79.4
M5	50.2	70.3	76.6
M6	47.6	63.7	69.7
M7	51.6	74.1	80.3
M8	49.7	71.2	75.7
M9	46.5	65.7	70.1
M10	55.3	81.2	86.1
M11	52.7	78.7	82.8
M12	50.6	65.2	69.2

Figures 8, 9, 10 and 11 show the development of compressive strength at 7, 28, and 90 days for binary mixes using CKC, WMP, and GGBS and multiblended mixes, respectively in comparison to control concrete (M0). At 90 days, the compressive strength of binary blends containing CKC (M1, M2, and M3) increased by 10.25%, 7.32%, and 2.26%, respectively, relative to the control mix (M0). It is evident that the compressive strength was highest for the 10% CKC (M1) mixture, achieving a strength of 82.8 at 90 days. This demonstrates that the optimal compressive strength was accomplished at a 10% CKC replacement level. This enhancement is ascribed to the increased fineness of CKC, promotes a more significant reaction with $Ca(OH)_2$, causing the development of further C–S–H and therefore increasing the compressive strength. Furthermore, the miniaturization of CKC particles also increases the compressive strength (48). The reduction in compressive strength for 15% CKC and 20% CKC, compared to 10% CKC, can be explained by the clinker dilution impact. The dilution influence, resulting from substituting part of the cement with CKC, is offset in CKC concrete by the filler influence, the pozzolanic reaction of CKC with calcium hydroxide, and the synergistic influence of the mineral admixture (48). The compressive strength of binary blends with WMP (M4 and M5) increased by (5.73% and 1.99%), while the compressive strength of (M6) diminished by (7.19% and 6.66%) at 90 days of age

versus the control mix (M0). The data indicate that the compressive strength was highest for the 10% WMP (M4) mixture, achieving strengths of 79.4 at 90 days. When the replacement level of WMP reaches 15% and 20%, the compressive strength starts to decrease, suggesting that the optimal limit for WMP replacement is 10%, beyond which the gains in strength are minimal or negative. The fine particles of WMP occupy the micro-voids of the concrete matrix, leading to a denser micro-structure in the HPSCC mix. This leads to enhanced mechanical properties in HPSCC mixes. In addition, marble powder decreases the porosity of the concrete mix and improves binding properties as a result of the chemical hydration of calcite and tricalcium aluminates (49). Furthermore, the compressive strength of binary blends with GGBS (M7 and M8) increased by (6.92% and 0.8%), while the compressive strength of (M9) decreased by (6.66%) at 90 days relative to the control mix (M0). Notably, the 20% GGBS (M7) blend achieved the highest compressive strength of 80.3 MPa at 90 days. The escalation in compressive strength of mix 7 can be ascribed to the incorporation of an optimal 20% GGBS content. The existence of calcium oxide and silicon dioxide in GGBS enables their reaction with calcium hydroxide, likely causing the establishment of supplementary crystalline calcium-silicate-hydrate gel.

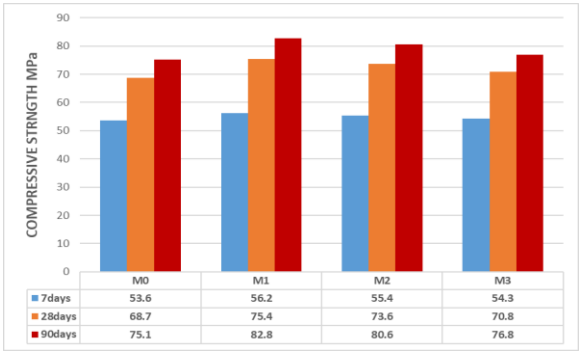


Figure 8. Compressive strength development results of HPSCCs with CKC blends

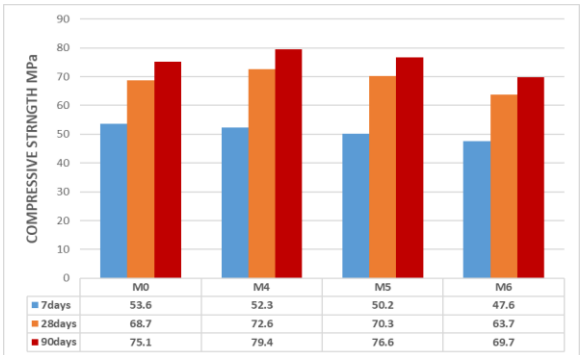


Figure 9. Compressive strength development results of HPSCCs with WMP blends

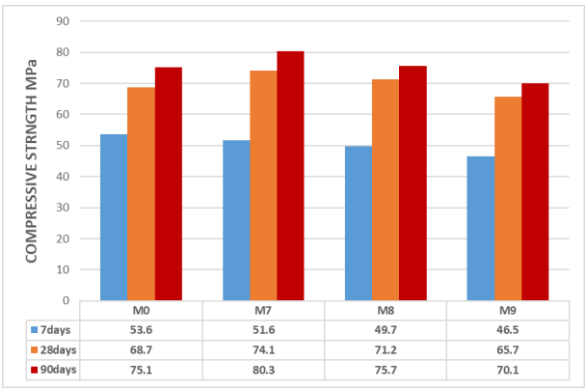


Figure 10. Compressive strength development results of HPSCCs with GGBS blends

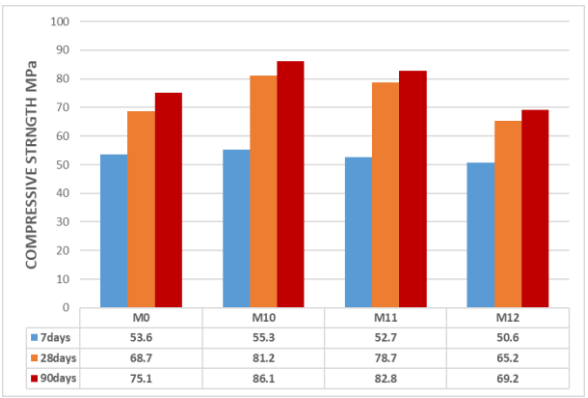


Figure 11. Compressive strength development results of HPSCCs with Multiblended mixes

These crystalline compounds are responsible for the improvement in strength properties (50). The compressive strength of multiblended mixes M10 and M11 increased by 14.65% and 10.25%, respectively, while the compressive strength of M12 decreased by 7.86% at 90 days of age in comparison with the control mix (M0). The maximum compressive strength was accomplished for the mix containing 10% CKC, 10% WMP, and 20% GGBS (M10), reaching a strength of 86.1 at 90 days. The observed enhancement can be ascribed to the synergy action of the pozzolans, wherein the optimal 20% GGBS content provided a notable filler or packing effect. The existence of calcium oxide and silicon dioxide in GGBS allows them to react with calcium hydroxide, which may result in the establishment of supplementary crystalline C–S–H gel. The inclusion of 20% CKC, 10% WMP, and 20% GGBS in the M12 mix gave rise to a marked lowering of compressive strength. This is because of the reduced supply of calcium hydroxide from OPC and the low reactivity of GGBS. “SEM analysis was conducted using a Hitachi TM4000Plus tabletop scanning electron microscope (Hitachi High-Tech Corp., Japan). The instrument

operated at an accelerating voltage of 20 kV under high vacuum conditions. Concrete specimens (10 mm × 10 mm × 10 mm) were dried and gold-coated prior to imaging." Figure 12 displays SEM images of different mixes, namely, M1, M4, M7, and M10, compared with the reference mix (M0) at ambient temperature. The mixes were selected based on their mechanical performance, specifically those with optimum strength at 28 days of age. The SEM analysis revealed that the presence of needles of ettringite (on a micron scale), thin hexagonal plates of portlandite (calcium hydroxide), and C-S-H plates as the main hydrated products in the control specimens. Due to the increased compressive strength from 10% CKC replacement, the SEM presented in Figure 9 shows pozzolanic reactions and gel development during cement hydration, with significant increases in C-S-H and decreases in $Ca(OH)_2$. Consequently, the microstructures of the cement have been enhanced because of the reaction between the active calcined kaolin clay and $Ca(OH)_2$ in the cement paste. Cement-based paste specimens with replaced calcined kaolin clay exhibited a more compact and denser pore structure (48). These findings are fully consistent with the compressive strength results. Clearly, the cement-marble powder paste specimens showed higher density and lower porosity than cement paste specimens. In this phase, the microstructure of the pastes consists of amorphous C-S- H particles, while Ettringite (E) needles are found within the pores, indicating that the paste is completely hydrated and all voids have been filled. Furthermore, with an increased amount of MP, the resulting SCC becomes denser and less porous (49). In contrast to the control concrete, pores are scarce in the 20% GGBS concrete. This effect is due to GGBS, which significantly decreases both the amount and average size and amount of $Ca(OH)_2$ crystals in the aggregate-mortar Interfacial Transition Zone (ITZ), thereby making the microstructure of the ITZ denser. These influences contribute to the great strength of concrete when an optimal amount of GGBS is used alongside PC (50, 51). Yuksel (52) noted that the micro-structure of concrete containing GGBS differs from that of PC concretes owing to further reactions involving GGBS. Generally, the cracks absence denotes effective mixing and strong bonding among the materials. In concrete, an ITZ forms where two distinct components meet. Thus, ITZ can be minimized through enhanced mixing and uniform distribution. It is concluded that a multi-blend of 10% CKC, 10% WMP, and 20% GGBS offers good bonding with cement. The lack of cracks indicates that the mixture composition and amounts are appropriate. The ITZ, resulting from the combination of various chemical elements, is obviously visible and represents a positive advancement in concrete strength. "Although EDS analysis was not conducted in this study, a qualitative porosity assessment was carried out using grayscale contrast in SEM images. The images of quaternary blends (e.g., Mix M12) showed significantly denser matrices and fewer pores compared to the control mix (M0), indicating enhanced microstructure due to synergistic pozzolanic action."

4. COST AND ENERGY PERSPECTIVE

Replacing 40 % of the binder with CKC, WMP, and GGBS lowers clinker content and therefore both energy demand and cost:

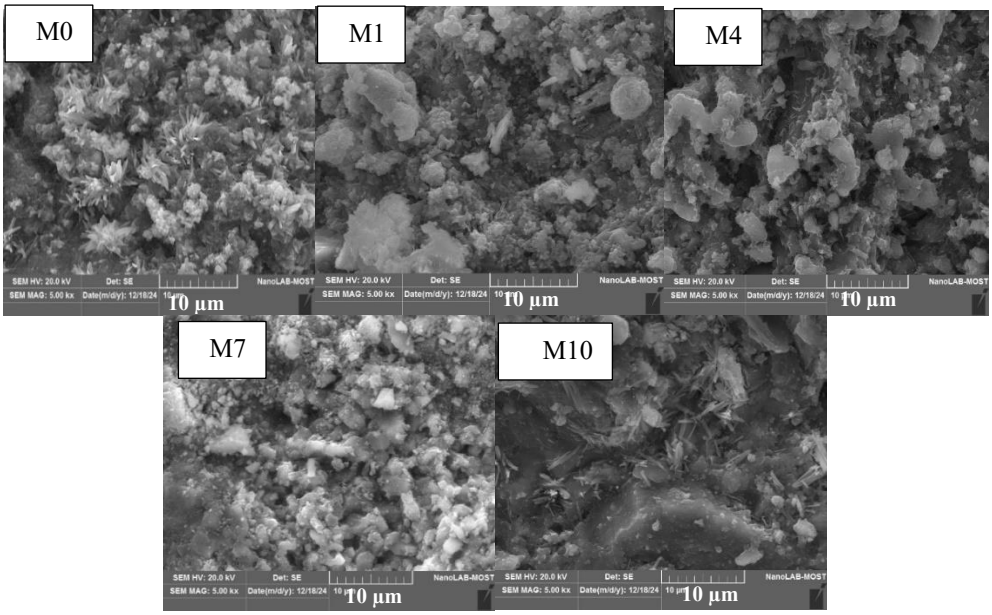


Figure 12. Scanning electron microscopy(SEM) results of HPSCCs at scale bars (10 µm) to assist with pore structure comparison

4. 1. ENERGY Producing 1 t of OPC clinker consumes ≈ 3.3 GJ and emits ≈ 850 kg CO_2. CKC is calcined at ~800 °C (≈ 1.8 GJ t^{-1}), GGBS is a by-product requiring only grinding (≈ 0.4 GJ t^{-1}), and WMP is simply milled waste (≈ 0.2 GJ t^{-1}). Weighted for a 60 % OPC + 40 % blend, the embodied energy is cut by about 28–30 % relative to 100 % OPC.

4. 2. COST In most markets GGBS and WMP are 30–50 % cheaper than OPC; CKC is close to parity. Using the same 60/40 proportion typically reduces binder cost by **10–15 %**. Any minor increase in grinding or transport is outweighed by the saving in clinker.

4. 3. SCALABILITY All three powders are already produced at industrial scale (GGBS) or can be scaled through existing calcination/grinding lines (CKC, WMP). No specialized equipment beyond additional silos and dosing screws is needed, so commercial SCC batching plants can adopt the blend with minimal capital expenditure. Thus, at scale the quaternary blend offers a tangible economic advantage while cutting embodied energy and CO_2—making it more sustainable than standard OPC-based SCC.

5. CONCLUSIONS

1. Increased cement usage results in higher carbon dioxide emissions, which subsequently cause detrimental environmental impacts. Utilizing blended pozzolanic materials (CKC, GGBS, and WMP together) up to 50% as partial cementitious replacements in concrete can mitigate these effects.
2. Slump flow diameters measured between 740 and 845 mm. Consequently, all concrete mixes met the SF3 class in terms of slump flow, except for the binary mixtures containing 20% CKC, which fell within the SF2 class according to EFNARC. Therefore, all mixes were considered to have acceptable rheological consistency / uniformity of flow from a filling perspective.
3. The V-funnel flow times (in seconds) ranged from 5.3 to 9.7. Accordingly, all mixes could be classified into the VF1 class, except for the 20% CKC mix, which was classified into the VF2 class based on EFNARC.
4. The results of the produced PA range from 0.874 to 0.996, which is categorized as class two (PA2) (within the EFNARC-recommended range). No segregation or blocking was noticed in any of the mixes.
5. Despite a reduction in certain fresh properties observed with CKC as a cement replacement in HPSCC, the resulting mixtures still met the requirements for SCC.
6. The addition of WMP and GGBS to the HPSCC mixes enhanced the fresh-state properties and mitigated the negative effects of CKC on these properties.
7. The use of CKC, WMP, and GGBS in a quaternary blend as partial cement replacements demonstrated the potential to achieve satisfactory fresh properties in HPSCC at replacement levels up to 50%.
8. The density of the HPSCC mixtures diminished with increasing pozzolan content (CKC, WMP, and GGBS) replacement increased, with the highest reduction in density observed for mixtures containing 20% CKC, 10% WMP, and 20% GGBS.
9. The HPSCC compressive strength outcomes confirm that all the concretes manufactured in this work are of high performance, with 28-day strength varying between 63.7 and approximately 81.2 MPa, and 90-day strength varying between 69.7 and 86.1 MPa.
10. For binary blends, the optimal replacement level was 10% CKC, 10% WMP, and 20% GGBS to achieve higher compressive strength.
11. The maximum compressive strength was attained with 10% CKC + 10% WMP + 20% GGBS (M10), reaching a strength of 86.1 at 90 days. The improvement in strength is attributed to the synergistic interaction between the filler effect, enhanced pozzolanic reactivity, and enhanced particle packing of the supplementary cementitious materials.
12. The quaternary blend exhibited a visibly denser microstructure; this is clearly illustrated in SEM image M10, where the pore network is markedly finer and less connected than in the reference mix, aligning with the observed strength gains.
١٣. The blended powders caused only subtle aesthetic changes. GGBS and CKC imparted a slightly darker grey tone than plain OPC, while the light-coloured WMP helped keep the hue balanced; overall, the hardened concrete retained a uniform medium-grey shade. Surface texture and finish quality (slump-flow-derived smoothness, absence of bleed lines) were indistinguishable from the reference SCC, so the mix remains suitable for exposed or architectural applications that demand consistent colour and a smooth, defect-free finish.

6. FUTURE WORK

Future research should focus on the long-term durability of the developed concrete mixtures under various environmental conditions, including freeze-thaw cycles and chemical exposure. Investigating the thermal cycling behavior and its impact on microstructure and mechanical properties would provide valuable insights for broader applications. Future work will incorporate EDS or image-based porosity quantification to provide more robust microstructural validation. Additionally, exploring the use of these sustainable concrete formulations in prefabricated structural elements could

help optimize production processes and enhance performance in industrial construction settings.

7. REFERENCES

1. Dadsetan S, Bai J. Mechanical and microstructural properties of self-compacting concrete blended with metakaolin, ground granulated blast-furnace slag and fly ash. Construction and Building Materials. 2017;146:658-67. https://doi.org/10.1016/j.conbuildmat.2017.04.158
2. Shi C, Wu Z, Lv K, Wu L. A review on mixture design methods for self-compacting concrete. Construction and Building Materials. 2015;84:387-98. https://doi.org/10.1016/j.conbuildmat.2015.03.079
3. Vivek S, Dhinakaran G. Durability characteristics of binary blend high strength SCC. Construction and Building Materials. 2017;146:1-8. https://doi.org/10.1016/j.conbuildmat.2017.04.063
4. Newman J, Choo BS. Advanced concrete technology 3: processes: Elsevier; 2003.
5. Caldarone MA, Taylor PC, Detwiler RJ, Bhidé SB. Guide specification for high performance concrete for bridges. 2005. http://worldcat.org/isbn/0893122459
6. Zhutovsky S, Kovler K. Influence of water to cement ratio on the efficiency of internal curing of high-performance concrete. Construction and Building Materials. 2017;144:311-6. https://doi.org/10.1016/j.conbuildmat.2017.03.203
7. Nadeem A, Memon SA, Lo TY. The performance of fly ash and metakaolin concrete at elevated temperatures. Construction and Building Materials. 2014;62:67-76. https://doi.org/10.1016/j.conbuildmat.2014.02.073
8. Afroughsabet V, Biolzi L, Ozbakkaloglu T. Influence of double hooked-end steel fibers and slag on mechanical and durability properties of high performance recycled aggregate concrete. Composite Structures. 2017;181:273-84. https://doi.org/10.1016/j.compstruct.2017.08.086
9. Dybeł P, Wałach D, Ostrowski K. The top-bar effect in specimens with a single casting point at one edge in high-performance self-compacting concrete. Journal of Advanced Concrete Technology. 2018;16(7):282-92. https://doi.org/10.3151/jact.16.282
10. Megid WA, Khayat KH. Effect of concrete rheological properties on quality of formed surfaces cast with self-consolidating concrete and superworkable concrete. Cement and Concrete Composites. 2018;93:75-84. https://doi.org/10.1016/j.cemconcomp.2018.06.016
11. Islam A, Alengaram UJ, Jumaat MZ, Bashar II, Kabir SA. Engineering properties and carbon footprint of ground granulated blast-furnace slag-palm oil fuel ash-based structural geopolymer concrete. Construction and building materials. 2015;101:503-21. https://doi.org/10.1016/j.conbuildmat.2015.10.026
12. Özbay E, Erdemir M, Durmuş Hİ. Utilization and efficiency of ground granulated blast furnace slag on concrete properties–A review. Construction and Building Materials. 2016;105:423-34. https://doi.org/10.1016/j.conbuildmat.2015.12.153
13. Mohan A, Mini K. Strength and durability studies of SCC incorporating silica fume and ultra fine GGBS. Construction and Building Materials. 2018;171:919-28. https://doi.org/10.1016/j.conbuildmat.2018.03.186
14. White G. Quantifying the impact of reclaimed asphalt pavement on airport asphalt surfaces. Construction and Building Materials. 2019;197:757-65. .https://doi.org/10.1016/j.conbuildmat.2018.11.189
15. Vejmelková E, Keppert M, Grzeszczyk S, Skaliński B, Černý R. Properties of self-compacting concrete mixtures containing metakaolin and blast furnace slag. Construction and Building Materials. 2011;25(3):1325-31. https://doi.org/10.1016/j.conbuildmat.2010.09.012
16. Kavitha O, Shanthi V, Arulraj GP, Sivakumar V. Microstructural studies on eco-friendly and durable Self-compacting concrete blended with metakaolin. Applied Clay Science. 2016;124:143-9. https://doi.org/10.1016/j.clay.2016.02.011
17. Melo KA, Carneiro AM. Effect of Metakaolin's finesses and content in self-consolidating concrete. Construction and Building Materials. 2010;24(8):1529-35. https://doi.org/10.1016/j.conbuildmat.2010.02.002
18. Ghoddousi P, Saadabadi LA. Study on hydration products by electrical resistivity for self-compacting concrete with silica fume and metakaolin. Construction and Building Materials. 2017;154:219-28. https://doi.org/10.1016/j.conbuildmat.2017.07.178
19. Boukendakdji O, Kadri E-H, Kenai S. Effects of granulated blast furnace slag and superplasticizer type on the fresh properties and compressive strength of self-compacting concrete. Cement and concrete composites. 2012;34(4):583-90. https://doi.org/10.1016/j.cemconcomp.2011.08.013
20. Ho L. Effect of curing regimes on mortar incorparating meatakaolin and slag at low water/binder ratio (Doctoral dissertation, Utar). 2013.
21. Ulubeyli GC, Bilir T, Artir R. Durability properties of concrete produced by marble waste as aggregate or mineral additives. Procedia engineering. 2016;161:543-8. https://doi.org/10.1016/j.proeng.2016.08.689
22. Kore SD, Vyas A. Impact of marble waste as coarse aggregate on properties of lean cement concrete. Case studies in construction materials. 2016;4:85-92. https://doi.org/10.1016/j.cscm.2016.01.002
23. Bhuva P, Patel A, George E, Bhatt D, editors. Development of self compacting concrete using different range of cement content. National Conference on Recent Trends in Engineering & Technology; 2011.
24. Elyamany HE, Abd Elmoaty M, Mohamed B. Effect of filler types on physical, mechanical and microstructure of self compacting concrete and Flow-able concrete. Alexandria Engineering Journal. 2014;53(2):295-307. https://doi.org/10.1016/j.aej.2014.03.010
25. Toubal Seghir N, Mellas M, Sadowski Ł, Krolicka A, Żak A, Ostrowski K. The utilization of waste marble dust as a cement replacement in air-cured mortar. Sustainability. 2019;11(8):2215. https://doi.org/10.3390/su11082215
26. Tennich M, Ouezdou MB, Kallel A. Thermal effect of marble and tile fillers on self-compacting concrete behavior in the fresh state and at early age. Journal of Building Engineering. 2018;20:1-7. https://doi.org/10.1016/j.jobe.2018.06.015
27. Suprakash AS, Karthiyaini S, Shanmugasundaram M. Future and scope for development of calcium and silica rich supplementary blends on properties of self-compacting concrete-A comparative review. Journal of Materials Research and Technology. 2021;15:5662-81. https://doi.org/10.1016/j.jmrt.2021.11.026
28. No I, Specification I. Central Agency for Standardization and Quality Control. 1984. https://iraqi-standards.org/Home/Ssection
29. Khoman R, Owaid H. Influence of Nanoparticles Additions on Fresh Properties and Compressive Strength of Sustainable Self-Compacting High-Performance Concrete Containing Calcined Pozzolanic Materials. International Journal of Mechanical Engineering. 2022;7(1):870-4. https://kalaharijournals.com/resources/101-120/IJME_Vol7.1_110.pdf

30. ASTM C. Standard specification for coal fly ash and raw or calcined natural pozzolan for use in concrete. ASTM Int. 2012;4. https://www.astm.org/c0618-19.html
31. Rajini B, Rao AN. Mechanical properties of geopolymer concrete with fly ash and GGBS as source materials. International Journal of Innovative Research in Science, Engineering and Technology. 2014;3(9):15944-53. http://dx.doi.org/10.15680/IJIRSET.2014.0309023
32. C989 A. Standard specification for ground granulated blast-furnace slag for use in concrete and mortars. American Society for Testing and Materials, Philadelphia. 2006. https://www.astm.org/c0989-22.html
33. ASTM A. ASTM C494: Standard specification for chemical admixtures for concrete. West Conshohocken, PA, USA: ASTM. 2005. https://www.astm.org/c0494-17.html
34. Bibm C, Ermco E. EFNARC: The European guidelines for self compacting concrete. Specification, Production and Use. 2005;63(3). https://efnarc.org/wpcontent/uploads/2019/12/efnarc_scc_guidelines_2005.pdf
35. Standard B. Methods for determination of density of hardened concrete. Part. 1881;114:1983-. https://www.bsigroup.com/en-GB/standards/
36. British S. 116: Testing concrete-Part 116: method for determination of compressive strength of concrete cubes. British Standards Institution London, UK. 1983. https://www.bsigroup.com/en-GB/standards/
37. Standard A. C1723–16 Standard Guide for Examination of Hardened Concret e Using Scanning Electron Microscopy. ASTM International, West Conshohocken PA. 2022. https://www.astm.org/c1723-22.html
38. Hilal N, Hamah Sor N, Hadzima-Nyarko M, Radu D, Tawfik TA. The influence of nanosunflower ash and nanowalnut shell ash on sustainable lightweight self-compacting concrete characteristics. Scientific reports. 2024;14(1):9450. https://doi.org/10.1038/s41598-024-60096-5
39. Amin M, Hadzima-Nyarko M, Agwa IS, Zeyad AM, Tayeh BA, Adesina A, et al. A Review on the Effect of Marble Powder on Properties of Self-Compacting Concrete. Advances in Science and Technology. 2024;152:61-72. https://doi.org/10.4028/p-gw4vSr
40. Shaker HR, AlSaraj W, Al-Jaberi L. Behavior of reinforced geopolymer concrete beams under repeated load. International Journal of Engineering, Transactions B: Applications. 2024;37(5):974-83. https://doi.org/10.5829/ije.2024.37.05b.14
41. Makki O, Al-Mutairee H. Response of rubcrete continuous deep beams under sinusoidal loads. International Journal of Engineering Transactions A Basics. 2022;35:1307-16. https://doi.org/10.5829/ije.2022.35.07a.10.
42. Bheel N, Benjeddou O, Almujibah HR, Abbasi SA, Sohu S, Ahmad M, et al. Effect of calcined clay and marble dust powder as cementitious material on the mechanical properties and embodied carbon of high strength concrete by using RSM-based modelling. Heliyon. 2023;9(4). https://doi.org/10.1016/j.heliyon.2023.e15029
43. Kavitha O, Shanthi V, Arulraj GP, Sivakumar P. Fresh, micro-and macrolevel studies of metakaolin blended self-compacting concrete. Applied Clay Science. 2015;114:370-4. http://dx.doi.org/10.1016/j.clay.2015.06.024
44. Makki O, Al-Mutairee H. Mechanical and dynamical properties of structural rubcrete mixes. International Journal of Engineering. 2022;35(09):1744-51. https://doi.org/10.5829/ije.2022.35.09C.10
45. Kadhim S, Al-Salima N, Kadhima M, Haleemb H. Structural Behavior of Concrete Members Strengthened Using Fiber Reinforced Polymer: A Review. 2025. http://dx.doi.org/10.5829/ije.2025.38.01a.15
46. Elbialy S, El-Latief A, Al-Jabali H, Elsayed H, Shawky S. Enhancing the properties of steel fiber self-compacting NaOH-based geopolymer concrete with the addition of metakaolin. Civ Eng J. 2024;10(7):2244-60. https://doi.org/10.28991/CEJ-2024-010-07-011
47. Hano MM, Hano SM, Al-Rawe HS. Examining the Compressive Behavior of SFRC and SCC Using Finite Element and Experimental Methods. Civil Engineering Journal. 2025;11(03). http://dx.doi.org/10.28991/CEJ-2025-011-03-017
48. Al-Hashem MN, Amin MN, Ajwad A, Afzal M, Khan K, Faraz MI, et al. Mechanical and durability evaluation of Metakaolin as cement replacement material in concrete. Materials. 2022;15(22):7868. https://doi.org/10.3390/ma15227868
49. Moreno AN, Ponce Palafox C, Múzquiz-Ramos EM, Avalos Belmontes F. Marble residues in construction materials: a review of the use of marble powder in mortars, concrete, and bricks. Revista ALCONPAT. 2022;12(2):162-83. https://doi.org/10.21041/ra.v12i2.522
50. Saranya P, Nagarajan P, Shashikala A, editors. Eco-friendly GGBS concrete: a state-of-the-art review. IOP conference series: materials science and engineering; 2018: IOP Publishing.
51. Bashir MT, Khan AB, Khan MMH, Rasheed K, Saad S, Farid F. Evaluating the implementation of green building materials in the construction sector of developing nations. J Hum Earth Future. 2024;5(3):528-42. https://doi.org/10.28991/HEF-2024-05-03-015
52. Yuksel I. Blast-furnace slag. Waste and supplementary cementitious materials in concrete: Elsevier; 2018. p. 361-415.

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

بتن خودتراکم با عملکرد بالا، پیشرفت قابل توجهی در فناوری ساخت و ساز محسوب می‌شود. این مطالعه به بررسی و مقایسه خواص حالت تازه جریان اسلامپ (D (mm، زمان جریان قیف V و آزمایش‌های جعبه L، خواص حالت سخت شده (وزن واحد و مقاومت فشاری) و ویژگی‌های ریزساختار (میکروسکوپ الکترونی روبشی (SEM) بتن خودتراکم با عملکرد بالا (HPSCC) با یک مخلوط مرجع (M0) به عنوان مخلوط کنترلی حاوی فقط OPC بدون جایگزینی پوزولانی پرداخت. دو سیستم ترکیبی در این مطالعه توسعه داده شد: یک سیستم ترکیبی دوتایی و یک سیستم ترکیبی چهارتایی که شامل (OPC+CKC، OPC+WMP، OPC+GGBS) و (OPC+CKC+WMP+GGBS) بود. نتایج مربوط به خواص تازه همه مخلوط‌های HPSCC در محدوده قابل قبول تعیین شده توسط EFNARC قرار گرفت، بدون اینکه جداشدگی یا آب انداختن مشاهده شود، و همه مخلوط‌های HPSCC عملکرد بهتری نسبت به مخلوط کنترلی نشان دادند. از این مطالعه می‌توان نتیجه گرفت که استفاده از ٦٠٪ OPC + 10% CKC + 10% WMP + 20% GGBS در مقایسه با مخلوط کنترل (M0) مقاومت فشاری تازه بهبود یافته و ریزساختار متراکم‌تری را ارائه داد که آن را برای استفاده به عنوان HPSCC مناسب می‌کند. حداکثر مقاومت فشاری در ۹۰ روز به ۸۵.۲ مگاپاسکال رسید. تجزیه و تحلیل هزینه نشان می‌دهد که خاک رس کائولن کلسینه شده و ضایعات صنعتی می‌توانند به عنوان مواد سبز امیدوارکننده عمل کنند. با کاهش محتوای OPC از طریق جایگزینی جزئی با CKC، WMP و GGBS، این مطالعه به کاهش قابل توجه انتشار CO_2 مرتبط با تولید سیمان کمک می‌کند و شیوه‌های ساخت و ساز پایدارتری را ترویج می‌دهد.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1115-1121

International Journal of Engineering

Journal Homepage: www.ije.ir

Investigation of Cement Compositions of Backfill Systems for Lining Wells with Inclined and Horizontal Sections

V. V. Nikishin*, A. V. Mysin, D. S. Kuznetsova

Department of Well Drilling, Oil and Gas Faculty, Empress Catherine II Saint Petersburg Mining University, Saint Petersburg, Russia

P A P E R I N F O

Paper history:
Received 23 May 2025
Received in revised form 02 June 2025
Accepted 28 June 2026

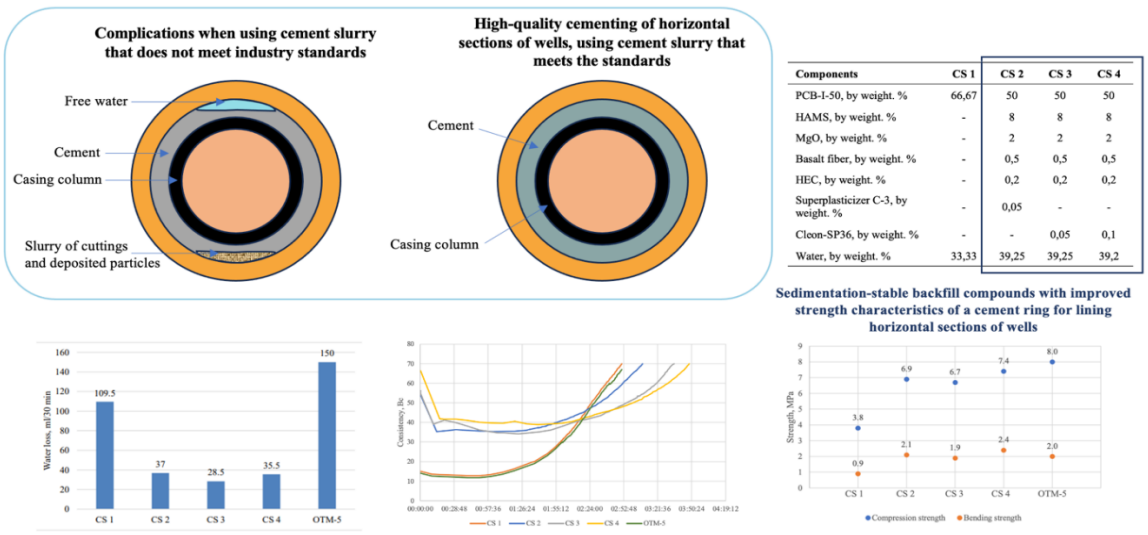
Keywords:
Cement Slurry
Cement Stone
Lining Horizontal Sections of Wells
Sedimentation Stability
Water Loss
Strength of Cement Stone

A B S T R A C T

In order to increase the flow rate of oil wells in oil production and gas extraction industry, there is an increase in the construction of inclined-directed and horizontal wells. Despite the experience of drilling horizontal wells, their cementing still does not have the required quality. For these areas, incomplete replacement of drilling mud with cement slurry is noted due to irregularities in the walls of the well, insufficient centering of the casing in the wells, sedimentation of process fluids. As a result, there may be no contact of the cement stone with the casing and rock in this zone, which leads to a violation of the tightness of the cement stone and the appearance of cross fluid flows. Thereby, the development of new compositions of cement slurries and well completion methods is a crucial task in the field of construction of horizontal wells for hydrocarbons. The research work presents the results of a study of the effect of various additives - hollow aluminosilicate microspheres, hydroxyethyl cellulose, magnesium oxide, polymer material "Cleon-SP36", superplasticizer C-3, basalt fiber on the structural, rheological and strength properties of cement slurry and cement stone. Sedimentation-stable compositions with improved strength characteristics of a cement ring for lining horizontal sections of wells, allowing to increase the efficiency of oil field development are proposed. The developed backfill compound based on PCB I-50 is designed for cementing oil and gas well casings at low and normal temperatures (15-50 °C), as well as taking into account the physico-mechanical characteristics of the hollow aluminosilicate microspheres (HAMS) – the recommended application depth is up to 3000 m vertically.

***doi**: 10.5829/ije.2026.39.05b.06*

Graphical Abstract



*Corresponding Author Email: nikishin_vv@pers.spmi.ru (V. V. Nikishin)

Please cite this article as: Nikishin VV, Mysin AV, Kuznetsova DS. Investigation of Cement Compositions of Backfill Systems for Lining Wells with Inclined and Horizontal Sections. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1115-21.

1. INTRODUCTION

The process of oil and gas production is not getting simpler, which requires to develop technologies and introduce new methods, including horizontal drilling (1, 2). For directional and horizontal wells, the total volume is steadily increasing. The construction of horizontal wells is effective in the reservoir of small thickness, while directional wells are highly effective in formations of large thickness (3-5).

The quality of well cementing depends on effectiveness of oil and gas fields and the use of modern methods to increase reservoir recovery (6, 7). In cementing deviated and horizontal wells compared to vertical ones, the pumping time for cement slurry increases. To reduce near-wellbore formation contamination by cement filtrate, it is necessary to decrease wellbore hydrostatic pressure using lightweight additives. During cementing of deviated and horizontal wells, standard cement slurry is prone to sedimentation and stratification under gravity. Additionally, incomplete displacement of drilling fluid by cement slurry is observed. This issue is controlled by adding a viscosifier, which enhances stability and creates an optimal rheological profile for the displacement process. However, the viscosifier increases slurry viscosity, requiring the introduction of dispersant additives. Another critical factor is cement shrinkage during curing, compensated by expansive additives. The quality of set cement is determined by high strength characteristics, ensuring reliable zonal isolation, especially across productive formations. This is achieved by incorporating reinforcing additives. Presently, regulating the properties of cement slurries, the most common way is to improve the quality of well cementing. Addition of additives to cement various geological factors of wells makes it possible to regulate their basic properties and select the necessary compositions (8). The change in the strength characteristics of cement stone is based on the changes in the physico-chemical processes occurring during the hardening of cement slurry.

Despite the experience of building horizontal wells, their cementing still does not have the required quality. On the inclined upper and horizontal surfaces, the casing tends to foot on the low wall of the well, which prevents its proper centering and has a negative effect on the formation of cement stone. A channel of deposited solid particles (Boycott effect) is observed on the underside of the annular space, and the filtration crust and thickened drilling mud are located around the circumference. The removal of solid particles, filtration crust and thickened drilling mud from these two problem areas is crucial to ensure proper cementing results. However, in conditions of horizontal wells, the completeness of the replacement of drilling mud with cement is difficult (9, 10). When cementing horizontal walls of wells, the process of settling of dispersed phase particles and delamination of cement slurry is accelerated under the influence of gravitational convection and a channel with liquid is formed at the upper wall of the well. In this layer, the continuity of the cement stone is disrupted and there is no contact between the cement stone and the rock, which can lead to cross flow behind casing. "On an inclined section of the well, the separated water can be displaced by a heavier and more mobile cement slurry, however, on an extended horizontal section, this water remains at the upper wall of the well" (11).

Despite the advantages of Portland cement as the main component of the backfill compound over other materials, filling the space between the well walls with a free cement slurry reduces the quality of wall reinforcement – the insulating characteristics of cement stone deteriorate during deformation processes. Therefore, to increase the lifetime of the well, it is necessary to use additional additives in the backfill compound (12).

Improving the quality of cementing horizontal sections of wells is achieved by using backfill materials with the following properties:

- high sedimentation stability and low water loss;
- high degree of pumpability;
- tight contact of the cement stone with the casing wall and rock formations;
- high strength characteristics of cement stone (11).

2. MATHERIALS AND METHOD

The properties of the developed backfill compounds were determined according to API 10B-2, ISO 10426-2 and API 10B-6 standards and recommendations on well cementing at Gazprom Neft fields:

- the density of cement slurry;
- mobility of the mortar on the AzNII cone;
- thickening time (consistency) per ZM1002 consistency meter;
- water loss on the FANN filter press;
- viscosity of the slurry, shear stress values on the FANN 35 rotary viscometer;
- setting time of the solution on the «Vika» device;
- compressive and bending strength of the stone on the test press Controls Digimax3.

The studies were carried out with the following backfill compounds:

CS 1 - Base cement slurry consisting of PCB-I-50 Portland cement and water at water-cement ratio = 0,5;

CS 2 - Cement slurry with the addition of hollow aluminosilicate microspheres (HAMS), hydroxyethyl cellulose (HEC), magnesium oxide, superplasticizer C-3, basalt fiber;

CS 3 (CS 4) - Cement slurry with the addition of hollow aluminosilicate microspheres (HAMS), hydroxyethyl

cellulose (HEC), magnesium oxide, polymer material "Cleon-SP36", basalt fiber;
OTM-5 is a lightweight backfill material containing easily dispersed and finely dispersed lightening additives with a density from 300 to 2600 kg/m^3.

The ratio of the components of cement slurries is summarized in Table 1.

Functions of the reagents used:

- hollow aluminosilicate microspheres - a lightweight supplement;
- hydroxyethyl cellulose - stabilizer;
- superplasticizer C-3 – plasticizer;
- Cleon-SP36 – plasticizer;
- magnesium oxide - expansion supplement;
- basalt fiber - providing high strength characteristics of cement stone.

3. RESULT AND DISCUSSION

3. 1. Studies of Rheological Characteristics, Sedimentation Stability and Water Loss of Cement Slurry

Hollow aluminosilicate microspheres are used to develop a lightweight cement slurry. Microspheres as a by-product are formed in the process of high-temperature coal combustion in Power plants as a result of granulation of the melt of the mineral part of the coals and the blowing of crushed small droplets by internal gases. Their distinctive feature is their low price, low density, thermal stability, significant strength and crack resistance (13). Adding of HAMS into the backfill compound increases the setting time of the cement slurry, its water retention capacity and accelerates the process of structure formation due to the formation of crystallization centers, which improves the pumpability of the mortar and the tightness of the structure. HAMS is characterized by high compressive strength: the ultimate strength is 15-30 MPa, which allows microspheres to maintain their structure even in aggressive highly viscous media. The physical properties of HAMS are presented in Table 2.

TABLE 1. Compositions of cement slurries

Components	CS 1	CS 2	CS 3	CS 4
PCB-I-50, by weight. %	66,67	50	50	50
HAMS, by weight. %	-	8	8	8
MgO, by weight. %	-	2	2	2
Basalt fiber, by weight. %	-	0,5	0,5	0,5
HEC, by weight. %	-	0,2	0,2	0,2
Superplasticizer C-3, by weight. %	-	0,05	-	-
Cleon-SP36, by weight. %	-	-	0,05	0,1
Water, by weight. %	33,33	39,25	39,25	39,2

TABLE 2. Physical properties of HAMS (17)

Fractional composition	Percentage content
< 100 microns	10 - 16%
100 - 160 microns	35 - 41%
160 - 250 microns	38 - 44%
250 - 400 microns	7 - 11%
400 - 500 microns	< 2%

The sedimentation stability of cement slurry is determined by the high strength of its thixotropic structure (14) and ensures the uniformity of the cement ring obtained during hardening (15). The water loss of the cement slurry leads to the formation of channels with the tempering fluid in the well, for example, with a water loss of up to 4%, channels up to 7 mm thick are formed. In this regard, a polymer-based structure-forming component, hydroxyethylcellulose (HEC), is added to the cement slurry. The introduction of a thickener (polymer) into the system affects an increase in the viscosity of the tempering fluid and, according to Stokes' law, leads to a decrease in the rate of precipitation of solid particles, which increases the sedimentation stability of the mortar.

HEC is also used to reduce the filtration of the dispersion medium from the cement slurry in the intervals of permeable formations. When pumping cement slurry into the well, filtering free water from the mortar contributes to its thickening, premature setting, and contamination with the filtrate of the mortar of the bottom-hole zone, which affects the productivity of the well. Addition of HEC to mortar increases its viscosity and binds some of the free water due to its structure, in which the molecules adsorbed on cement particles and create protective shells.

The mechanism of action of HEC against sedimentation:

a) Polymer mesh formation

- Adsorption on cement particles: HEC molecules are adsorbed on the surface of cement grains through hydrogen bonds and electrostatic interactions. They create "bridges" between the particles, preventing them from settling.
- Forming a 3D structure:

Long chains of HEC intertwine to form a three-dimensional grid that:

- increases the structural viscosity (the solution behaves like a gel at rest);
- captures and holds solid particles and air bubbles in suspension.

b) Viscosity increase

Deceleration of gravity deposition: The HEC dramatically increases the shear stress limit. It takes more energy than gravity to start moving the particles.

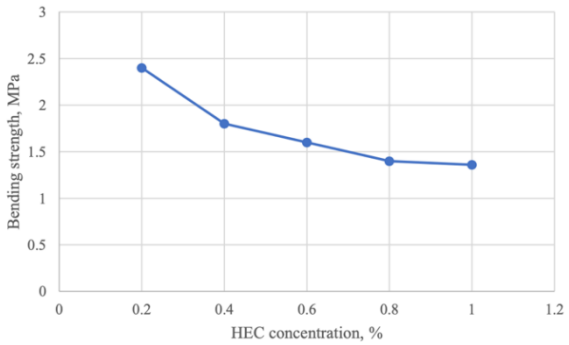


Figure 1. Effect of the HEC on the bending strength of cement stone

- Thixotropic behavior: at rest, the solution thickens, and when stirred, it liquefies. This prevents stratification during transportation.

c) Water retention

- Hydrogel formation: HEC binds free water in its structure, reducing its migration to the surface. There is no "watery" stratification.

With an increase in the concentration of HEC in the backfill system, the water loss of the cement slurry decreases, however, at a concentration of HEC of more than 0.2%, a decrease in the strength of cement stone is observed due to a slowdown in hydration (HEC forms films on cement particles) and an increase in porosity due to air entrainment. Figure 1 shows the effect of the HEC concentration in the backfill compound on the bending strength of cement stone.

When preparing the cement slurry in the laboratory, the HEC was dissolved in advance in water at a temperature of 45 °C with stirring for 3 hours to form a solution exhibiting pseudoplastic properties. Due to its non-ionic nature, HEC is characterized by high compatibility with various reagents. Another significant advantage of HEC is that it is environmentally friendly, as it is capable of biodegradation without forming harmful substances.

According to the results of the experiments, it was noted that cementing slurries of CS 2-4 with a concentration of 0.2% hydroxyethylcellulose polymer have zero water loss (Table 3).

The results of a study of the filtration index of cement slurries at an excess pressure of 690 kPa on a FANN filter press are shown in Figure 2.

For cementing shanks, the water loss value according to the Requirements of PJSC Gazprom Neft should not exceed 50 ml/30 min. Based on Figure 2, the water loss rates of CS 2-4 mortars meet the requirements, which will allow pumping the maximum amount of cement slurry into the well, obtaining the necessary compressive strength and good adhesion between cement stone, rock and casing, and the water loss rate of cement slurry without special additives (CS-1) is unsatisfactory and

TABLE 3. The effect of reagents on the density and sedimentation stability of cement slurry

Composition number	Density, kg/m^3	Presence of water loss in the vertical cylinder	Presence of water loss in the horizontal cylinder
CS 1	1890	yes	yes
CS 2	1460	no	no
CS 3	1460	no	no
CS 4	1500	no	no
OTM-5	١٥٠٠	yes	yes

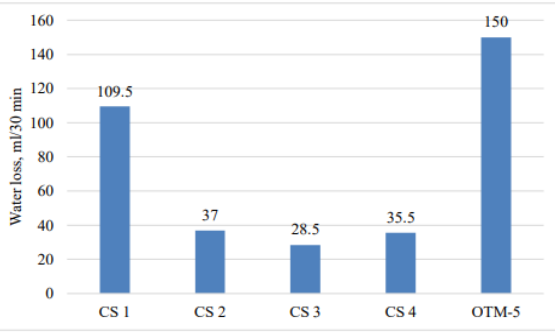


Figure 2. Water loss of backfill compounds

exceeds the required value. Cement slurries containing the polymer additive "Cleon-SP36" have the best rheological properties.

The disadvantage of the stabilizing additive in the form of HEC in the backfill compounds is a decrease in the spreadability of the slurry. Superplasticizer C-3 and polymer material "Cleon-SP36" were used to increase the mobility of the cement slurry. Superplasticizer C-3 is a product of condensation of naphthalene sulfonic acid with formaldehyde, the degree of polymerization is 2-10. It has no inhibitory effect on hydration processes. The main component of the C-3 superplasticizer is sulfonated naphthalene-formaldehyde polycondensates (82-84%). These organic compounds belong to surfactants that affect the structure of the mixture, reducing the need for water and increasing the strength of the cement stone. They are supplemented with 8-10% sodium sulfate and up to 10% water. Polymer material "Cleon SP36" manufactured by «Orgpolymersynthesis SPb» LLC is an aqueous polymer system based on acrylic and vinylsulfonic acids and their salts. The presence of several types of functional groups in the polymer provides a high affinity for cement, and its branched structure leads to the fact that cement particles repel each other not only due to electrostatics, but also due to steric difficulties. The use "Cleon SP36" improved the rheological characteristics of backfill compounds, an increase in strength, a decrease in porosity, and a reduction in the water-cement ratio. It has thermal stability and belongs to a non-toxic substance.

Figure 3 shows the results of measurements of the spreadability of cement slurries. The change in the consistency of the compounds was studied at a temperature of 50 °C and a rotation speed of the consistency mixer at 150 rpm, the results are illustrated in Figure 4.

The thickening time is considered to be the time that has elapsed from the moment of initial application of pressure and temperature to the moment when the cement slurry reaches a consistency of 70 Bc – the generally accepted consistency value at which the cement slurry is considered not to be pumped. CS-4 has the longest thickening time at a temperature of 50 °C.

The measurement results of the setting time of cement slurries are shown in Figure 5.

When choosing cementing modes, the rheological parameters of the mortar are important for predicting the magnitude of the hydraulic resistances of the cement slurry that occur during the initiation of its flow, and the completeness of the replacement of drilling mud with cement slurry (16). Figure 6 shows the results of research on the FANN 35 rotary viscometer device.

3. 2. Investigation of the Physico-mechanical Characteristics of Cement Stone After the cement slurry hardens, the cement stone must have significant strength characteristics, low permeability, and good adhesion to the casing and rock formations (17, 18).

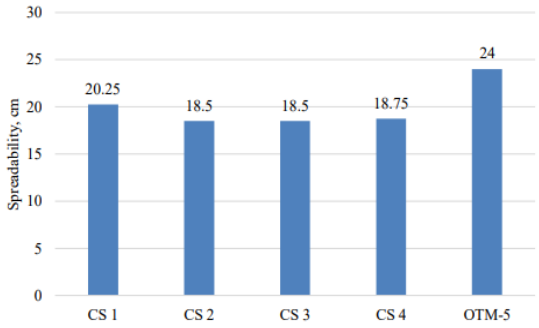


Figure 3. Spread ability of cement slurries

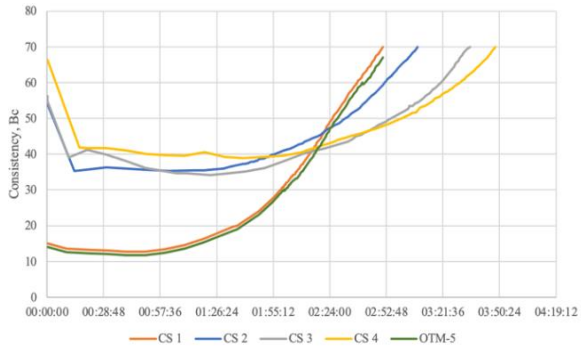


Figure 4. Changes in the consistency of cement slurries

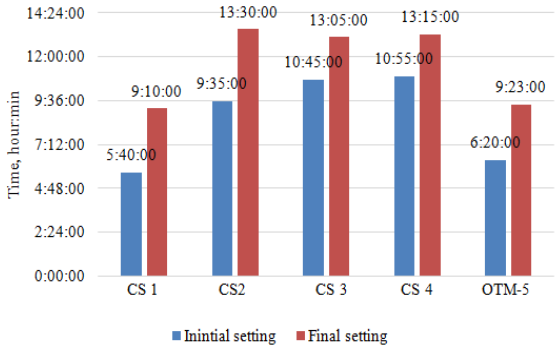


Figure 5. Setting time of cement slurries

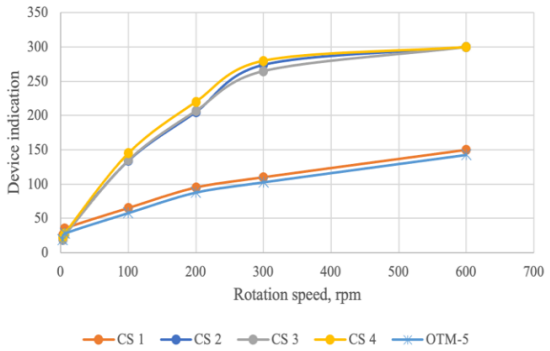


Figure 6. Research results on the FANN 35 rotary viscometer

Inappropriate contact between rock formation – cement stone – casing prevents their strong adhesion and the formation of reliable tightness of casing string annulus, and therefore an expansion additive is introduced into the backfill compound. Expansion of the oxide type is considered the most effective, providing maximum linear expansion with a minimum amount of a component – magnesium oxide or calcium oxide. This expansion is caused by the formation of corresponding metal hydroxides the oxides of which occupy a smaller volume than the hydration products (19-21).

Volumetric reinforcement is used to increase the strength and increase the ability of cement stone to withstand dynamic loads (6, 22). Reinforcing fibers as a short lengths of basalt fiber are used. The fiber hardening may occur due to the transfer of the applied load caused by the cement matrix to the fibers by means of tangential forces acting along the interface. If the modulus of elasticity of the fibers exceeds the modulus of elasticity of the cement matrix, then the fibers perceive the main part of the acting stresses. Then, total strength of the cement stone is proportional to their volume content (23). The cement matrix provides resistance to compressive stresses, the basalt fiber provides resistance to tensile and bending stresses.

Besides, adding basalt fiber into the backfill compound increases the expansion of the cement slurry-

stone by transferring the crystallization pressure of the expanding additive to the frame formed by the fiber (23).

The study conducted by Rozhkova et al. investigated the effect of cement reinforcement to the coefficient of linear stone expansion which has revealed that the fiber concentration of more than 0.5% reduces the expansion values of cement, and also leads to fiber supersaturation in the cementing slurry and the formation of lump. The results of studies of backfill compound for strength characteristics are shown in Figure 7.

According to the requirements of PJSC Gazprom, the recommended minimum value of compressive strength of cement stone after 2 days is 3.4 MPa. Compressive strength testing of each of the cement stones showed sufficient values, the cement stone formed from CS-4 exceeds the required compression strength by 118%.

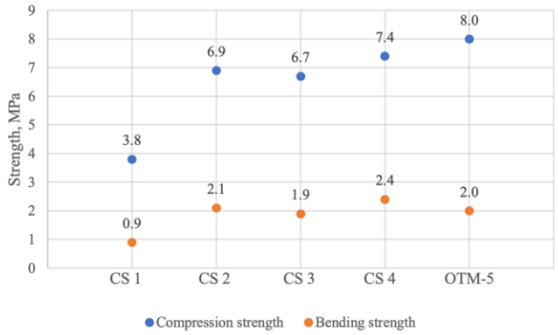


Figure 7. Strength characteristics of cement stone (Compiled by the authors)

4. CONCLUSION

The research results show that a non-additive cement slurry consisting of PCB 1-50 and process water, as well as lightweight backfill material OTM-5, do not meet the requirements of high-quality cementing of horizontal sections of wells and, therefore, it is necessary to use various reagents that improve the properties of the cement slurry. The use of hydroxyethyl cellulose has a positive effect on the filtration characteristics and sedimentation stability of the cement slurry. The water loss decreased by 74% with a concentration of 0.2% HEC, in comparison with CS-1. To increase the mobility of cement slurries, sulfonated naphthalene-formaldehyde polymers such as Superplasticizer C-3 are mainly used. However, in conditions of increased requirements for the quality of cementing horizontal sections of wells and environmental safety, the polymer material Cleon-SP36 is more effective in influencing the rheological and physico–mechanical characteristics of cement slurry - stone. The addition of basalt fiber to the backfill systems has a positive effect on the strength of cement stone. With a fiber concentration of 0.5%, the increase in compression and bending strength was 49% and 63%, respectively. The developed backfill compound based on PCB I-50 is designed for cementing oil and gas well casings at low and normal temperatures (15-50 °C), as well as taking into account the physico-mechanical characteristics of the HAMS – the recommended application depth is up to 3000 m vertically. Further detailed studies are planned for high-depth conditions and elevated bottomhole temperatures.

5. REFERENCES

1. Litvinenko VS, Tsvetkov PS, Dvoynikov MV, Buslaev GV. Barriers to implementation of hydrogen initiatives in the context of global energy sustainable development. Записки Горного Института. 2020;244:428-38. 10.31897/PMI.2020.4.5
2. Litvinenko V. The role of hydrocarbons in the global energy agenda: The focus on liquefied natural gas. Resources. 2020;9(5):59. 10.3390/resources9050059
3. Belikov SA ST, Kardopoltsev AS, Leskin FYu. . Efficiency analysis of horizontal wells with multistage hydraulic fracturing using the example of the Kondinsky field. Petroleum Engineering. 2023;21(1):39-50. https://doi.org/10.17122/ngdelo-2023-1-39-50
4. Pryakhin E, Azarov V. Increasing the adhesion of fluoroplastic coatings to steel surfaces of pipes with a view to their use in gas transmission systems. Chernye Met. 2024;3:69-75. https://doi.org/10.17580/chm.2024.03.11
5. Kadochnikov VG, Dvoynikov MV. Development of technology for hydromechanical breakdown of mud plugs and improvement of well cleaning by controlled buckling of the drill string. Applied Sciences. 2022;12(13):6460. https://doi.org/10.3390/app12136460
6. Shao M, Aleksander P, Xia Y, Xu H, Tian Y, Fetisov V, et al. Understanding the phase behavior during CO2 flooding by dissipative particle dynamics. Journal of Molecular Liquids. 2024;409:125514. https://doi.org/10.1016/j.molliq.2024.125514
7. Chaliapin DV BD, Fattakhov MM, Chaliapina AD, Kuznetsov VG. . Investigation of factual information on lining of wells using machine learning and neural networks. Oil and gas. 2021;3:108-19. 10.31660/0445-0108-2021-3-108-119
8. Xu F, Xu Z, Tang S, Ren Q, Guo Y, Wang L, et al. Evolution of physical and mechanical properties of cementing materials during underground energy exploitation and storage. Journal of Energy Storage. 2022;45:103775. 10.1016/j.est. 2021.103775
9. Sabins FL. Problems in cementing horizontal wells. Journal of petroleum technology. 1990;42(04):398-400. 10.2118/20005-PA
10. Blinov PA, Sadykov MI, Gorelikov VG, Nikishin VV. Development and research of backfill compounds with improved elastic and strength properties for oil and gas well lining. Записки Горного института. 2024(268 (eng)):588-98.
11. G. Romeu Delgado CDLBn. The main problems in the cementing long extended reach wells in Cuba. Petroleum Engineering. 2023;21(2):86-91. https://doi.org/10.17122/ngdelo-2023-2-86-91
12. Agwu OE, Akpabio JU, Akpabio MG. Potentials of waste seashells as additives in drilling muds and in oil well cements. Cleaner Engineering and Technology. 2020;1:100008. https://doi.org/10.1016/j.clet.2020.100008
13. Teryaeva T, Kostenko O, Ismagilov Z, Shikina N, Rudina N, Antipova V. Physico-chemical properties of aluminosilicate hollow microspheres. Vestnik KuzGTU. 2013;5:86-90.
14. Blinov PA SA, Lobachev IM, Volkov AR. . Evaluation of Strength Properties of Cement Containing NaCl in Mixing Water for Cementing Wells in Halite Strata. International Journal of

Engineering, Transactions A: Basics. 2024;37(10):2109-15. 10.5829/IJE.2024.37.10A.21

15. Tabatabaee Moradi SSh NN, Nikolaeva TN. . Development of spacer fluids and cement slurries compositions for lining of wells at high temperatures. Journal of Mining Institute. 2020;242:174-8. 10.31897/PMI.2020.2.174
16. Tao C, Rosenbaum E, Kutchko BG, Massoudi M. A brief review of gas migration in oilwell cement slurries. Energies. 2021;14(9):2369. https://doi.org/10.3390/en14092369
17. Fetisov V. Analysis of numerical modeling of steady-state modes of methane–hydrogen mixture transportation through a compressor station to reduce CO2 emissions. Scientific Reports. 2024;14(1):10605. https://doi.org/10.1038/s41598-024-61361-3
18. AA. K. Research and development of expansion additives for backfill compounds. Bulletin of the Perm National Research Polytechnic University, Geology Oil and gas and mining. 2015;16:46-53. 10.15593/2224-9923/2015.16.5
19. Bugrayev A, Nafikova S, Taoutaou S, Timonin A, Gurbanov G, Burkenya A, et al., editors. Case Studies of Expanding Cement to Improve Wellbore Sealing. SPE Gas & Oil Technology Showcase and Conference; 2019: SPE.
20. Ghanbari M, Kohnehpooshi O, Tohidi M. Experimental study of the combined use of fiber and nano silica particles on the properties of lightweight self compacting concrete. International Journal of Engineering, Transactions B: Applications. 2020;33(8):1499-511. 10.5829/ije.2020.33.08b.08
21. Pryakhin E, Azarov V. Comparative analysis of the use of epoxy and fluoroplastic polymer compositions as internal smooth coatings of the inner cavity of steel main gas pipelines. CIS Iron and Steel Review. 2024;28:93-8. https://doi.org/10.17580/cisisr.2024.02.16
22. Ishbaev G, Dil'miev M, Ishbaev R, Latypov T. Development of grouting materials with high impact strength. Burenie i neft. 2015;9:38-41.
23. Hassan Q, Jabbar A, Mohammed D. Experimental and Numerical Investigation of the Impact of Basalt Fibers and Tie Spacing on Short Concrete Column Behavior. International Journal of Engineering, Transactions A: Basics. 2023;36(7):1287-99. 10.5829/ije.2023.36.07a.10

COPYRIGHTS

©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

به منظور افزایش سرعت جریان چاههای نفتی در صنعت تولید نفت و استخراج گاز، افزایش ساخت چاههای شیبدار و افقی وجود دارد .با وجود تجربه حفاری چاه های افقی، سیمان آنها هنوز کیفیت مورد نیاز را ندارد .برای این مناطق ، جایگزینی ناقص گل حفاری با دوغاب سیمان به دلیل بی نظمی در دیواره های چاه، مرکز ناکافی پوشش در چاه ها ، رسوب مایعات فرآیند، که در نتیجه ممکن است هیچ ارتباطی از سنگ سیمان با پوشش و سنگ در این منطقه وجود نداشته باشد، که منجر به نقض سفتی سنگ سیمان و ظاهر جریان مایع متقاطع می شود. در نتیجه، توسعه ترکیبات جدید دوغاب سیمان و روش های تکمیل چاه یک کار مهم در زمینه ساخت چاه های افقی برای هیدروکربن ها است. این کار تحقیقاتی نتایج یک مطالعه در مورد اثر مواد افزودنی مختلف - میکروسفرهای آلومینوسیلیکات توخالی، هیدروکسی اتیل سلولز، اکسید منیزیم، مواد پلیمری "Cleon-SP36"، سوپرپلاستیزر C-3، فیبر بازالت بر خواص ساختاری، رئولوژیکی و استحکام دوغاب سیمان و سنگ سیمان را ارائه می دهد. رسوب-ترکیبات پایدار با ویژگی های قدرت بهبود یافته یک حلقه سیمان برای پوشش بخش های افقی چاه ها، اجازه می دهد تا بهره وری توسعه میدان نفتی را افزایش دهد. ترکیب پشتی توسعه یافته بر اساس pcb I-50 برای سیمان پوشش های چاه نفت و گاز در دمای پایین و عادی (50-15) درجه سانتیگراد و همچنین با در نظر گرفتن ویژگی های فیزیکی-مکانیکی ژامبون ها طراحی شده است – عمق کاربرد توصیه شده تا 3000 متر عمودی است.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1122-1138

International Journal of Engineering

Journal Homepage: www.ije.ir

Mapping Flood Zone in Khorram Abad River by Integrating Hydrological and Hydraulic Models

E. Fatapour[a], A. Afrous*[b], B. Aminnejad[c], A. Saremi[a], A. Khosrojerdi[a]

[a] Department of Water Science and Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
[b] Department of Water Science and Engineering, Dezful Branch, Islamic Azad University, Dezful, Iran
[c] Department of Water Science and Engineering, Rudehen Branch, Islamic Azad University, Rudehen, Tehran, Iran

PAPER INFO

Paper history:
Received 27 January 2025
Received in revised form 08 April 2025
Accepetd 07 May 2025

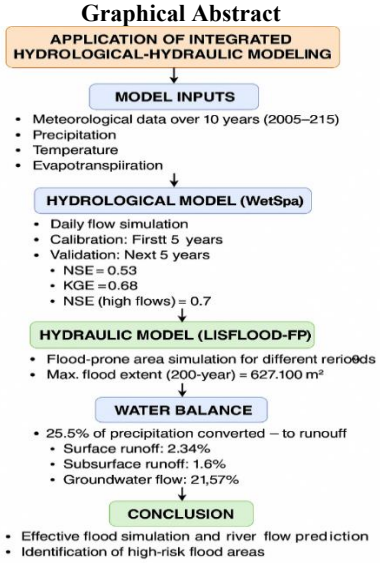
Keywords:
Floods
WetSpa Model
LISFLOOD-FP Model
River Flow Simulation
Khorram Abad

ABSTRACT

The aim of this study is to investigate the application of an integrated hydrological-hydraulic modeling approach to simulate river flows, flood hydrographs, and flood-prone areas in the Khorram Abad (Iran) drainage basin. In this study, the WetSpa hydrological model and the LISFLOOD-FP hydraulic model were used to simulate river flows and flood zones. Ten years of meteorological data (water years 2005–2006 to 2014–2015) were employed for the simulations. The models were calibrated using data from the first five years, while the subsequent five years of data were used for validation. The WetSpa model was applied to simulate daily flows, and the LISFLOOD-FP model was used to simulate flood-prone areas for different return periods. The WetSpa model successfully simulated daily flows with acceptable accuracy, achieving performance metrics such as the Nash-Sutcliffe Efficiency (NSE) of 0.53 and the Kling-Gupta Efficiency (KGE) of 0.68. For high-flow events, the NSE coefficient reached 0.7, indicating the model's effectiveness in flood prediction. The results from the LISFLOOD-FP model showed that the maximum flood extent for the 200-year return period was 627,100 square meters. The water balance analysis revealed that 25.5% of the precipitation was converted into runoff, comprising 2.34% surface runoff, 1.6% subsurface runoff, and 21.57% groundwater flow. This study demonstrates that the integration of hydrological and hydraulic models can effectively predict flood behavior and delineate high-risk areas.

doi: *10.5829/ije.2026.39.05b.07*

Graphical Abstract



*Corresponding Author Email: Ali.afrous@gmail.com (A. Afrous)

Please cite this article as: Fatapour E, Afrous A, Aminnejad B, Saremi A, Khosrojerdi A. Mapping Flood Zone in Khorram Abad River by Integrating Hydrological and Hydraulic Models. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1122-38.

1. INTRODUCTION

In recent decades, climate change has become one of the most pressing global issues and environmental crises. The far-reaching effects of these changes, observable at both local and global scales, manifest in temperature variations, altered precipitation patterns, and increased climatic instabilities (1). These transformations have led to serious and multifaceted consequences for water resources, food security, ecosystems, infrastructure, and public health (2). Among the most concerning impacts of climate change is the heightened frequency and intensity of extreme weather events, particularly floods and droughts (3). These phenomena pose especially severe threats in countries like Iran, which possesses distinct climatic and geographical characteristics (4). Iran's location within the global arid belt, coupled with its unique topographical features, makes it one of the most vulnerable countries to climate change (5). Decreasing precipitation in many regions, rising temperatures, and the occurrence of sudden heavy rainfall leading to severe flooding exemplify the complex and destructive impacts of these changes. The increasing frequency of large-scale and abrupt floods in recent years underscores the urgent need for comprehensive studies and effective risk management strategies (6).

Flooding, defined as the overflow of water beyond its natural boundaries, often results in widespread damage to agricultural lands, urban infrastructure, financial assets, and human life (7). Floods have been recognized as the most frequent and devastating natural disasters worldwide over the past decade, causing profound environmental, social, and economic consequences (8). Rapid population growth and accelerating industrialization have driven human settlements closer to riverbanks and concentrated economic activities in floodplains. This trend has markedly increased the frequency and severity of flood events (5). In Iran, reports indicate a 40% rise in flood incidents over recent decades (9). Additionally, official statistics reveal that over the past 53 years, more than 12,000 Iranian citizens have lost their lives due to floods. These alarming figures emphasize the critical importance of addressing flood-related issues to mitigate their disastrous impacts (10).

Lorestan Province, due to its distinct hydrological and geomorphological features, is one of the most flood-prone regions in Iran. The combination of narrow riverbeds and the presence of villages and agricultural lands along riverbanks has amplified the financial and human losses caused by floods (5). Over the past 25 years, this province has experienced 30 significant flood events, resulted in the loss of approximately 60 lives during this period (4). These statistics highlight the province's heightened vulnerability and the urgent need for advanced flood risk assessment and management.

Khorramabad River, a crucial watershed in western Iran, plays a vital role in supplying water resources and maintaining the region's ecological balance. In recent decades, this river has experienced severe and unexpected flooding, inflicting extensive damage on agricultural productivity, infrastructure, and water resources (3). Beyond physical destruction, these floods have triggered numerous social and economic challenges, such as forced migrations and escalating poverty in local communities. With ongoing population growth and urban expansion in areas surrounding the river, the region's susceptibility to climate change and flood events has significantly intensified (5). Accordingly, one of the most fundamental approaches to mitigating the destructive impacts of floods is the precise identification of flood-prone areas within watersheds. Floodplains and river-adjacent regions, which serve as hubs for economic and social activities due to their advantageous characteristics, are persistently at risk of flood-related hazards (6). Therefore, in such vulnerable zones, determining the extent of flood inundation, floodwater elevation relative to ground level, and flood characteristics across various return periods commonly known as flood zoning becomes critically important (8).

Despite extensive research on flood mapping and prediction, a significant and well-defined study gap persists regarding the integration of advanced hydrological and hydraulic models with GIS for high-accuracy flood simulation. This gap becomes even more pronounced when considering the complex and multifaceted impacts of climate change. This study explicitly addresses this gap by proposing a comprehensive and innovative approach that combines these techniques for the Khorramabad River watershed. By doing so, this research not only enhances the precision of flood prediction but also establishes a robust framework for future studies in regions with similar climatic and hydrological characteristics. Moreover, this approach offers practical applications for disaster risk reduction and informed urban and environmental planning. The study aims to fill this critical gap by providing a detailed, high-accuracy model that incorporates advanced hydrological and hydraulic modeling techniques integrated with GIS tools.

To date, several methods have been proposed for identifying flood-prone areas, including synoptic flood forecasting (11), flood prediction through river hydrograph analysis (12), and statistical routing and probability-based forecasting of flood occurrence and return periods (13). Hydrological and hydraulic models have proven to be advanced and effective tools for simulating and forecasting various hydrological conditions. These models analyze complex water flow processes, climatic variations, geographical changes, and meteorological data such as rainfall and temperature,

enabling precise predictions of flood timing, magnitude, and location (13). Among these techniques, the integration of Geographic Information Systems (GIS) with hydrological and hydraulic models has opened new frontiers in flood forecasting and simulation (14). GIS-based analyses enable researchers to scrutinize an area's geographical and topographical features in detail and incorporate climate change impacts more accurately into simulation models. This integration not only improves prediction accuracy but also plays a vital role in identifying vulnerable areas susceptible to future flooding (13).

In this study, the selection of hydrological and hydraulic models is meticulously justified based on their demonstrated efficiency, reliability, and adaptability to regions with similar climatic and geographical conditions. The models employed, including advanced rainfall-runoff and water surface elevation simulation techniques, have been extensively validated in previous research and have consistently shown high performance in capturing the unique hydrological dynamics of the Khorramabad River basin. Furthermore, the combination of these models with GIS tools creates a sophisticated and highly accurate framework for detailed flood mapping and comprehensive risk assessment. This integrative approach offers a scientifically robust methodology for addressing the pressing challenges of flood prediction and management, making it the most suitable choice for this research.

The present study, focusing on the Khorramabad River as a representative of sensitive watersheds in Iran, aims to simulate and predict flood occurrences while analyzing the far-reaching impacts of climate change on this basin. For the first time, this research employs an integrated approach that combines advanced hydrological and hydraulic models with GIS-based data analysis. The objective is to develop a comprehensive and highly accurate model for flood prediction and the identification of high-risk areas. Given climate change projections indicating a potential 2°C rise in Iran's temperature by 2030, the significance of this study becomes even more apparent. This anticipated temperature increase, alongside diminishing runoff and escalating drought conditions, is expected to exacerbate the frequency and severity of sudden and intense floods in regions like the Khorramabad basin. Additionally, factors such as rapid population growth, urban sprawl, and vegetation loss further amplify the area's vulnerability, underscoring the urgent need for proactive research and effective flood management strategies.

2. RESEARCH LITERATURE

Rad et al. (10) used the HEC-RAS model in Khorramabad watershed, Iran, to simulate flood profiles and identify flood-prone areas. Their findings showed that significant portions of agricultural and residential land are at risk during floods, emphasizing the need for accurate flood mapping for mitigation planning. Hassas et al. (11) developed a web-based tool for flood spreading modeling using Google Maps DEM data. The tool, applied to the Ayutthaya River flood in Thailand and the Shiraz Flood in 2019, effectively assessed flood risks and aided disaster preparedness planning. Tsegaye et al. (12) integrated GIS with hydrologic-hydraulic modeling to assess flood risks in Bonita Bay, Florida. They have simulated flood inundation due to sea-level rise, land-use changes, and groundwater rise, offering a comprehensive framework for flood risk management.

Chen (13) combined GIS with multi-criteria analysis techniques like AHP and TOPSIS for flood hazard mapping in the Dadu River basin, China. Their approach was successfully identified flood-prone areas, validated by historical flood data. Flores et al. (14) examined flood hazard disparities in Houston, highlighting that minority neighborhoods are disproportionately affected by floods not well-represented in FEMA's maps. The study stressed the need for revised flood risk mapping. Madi et al. (6) integrated the HEC-RAS model into a decision support system for flood prediction in Algeria's Tamanrasset Valley. The study used GIS to assess flood risks under various return periods, demonstrating effective flood zone identification.

Kastridis and Stathis (15) applied the HEC-RAS and SCS models to analyze flash floods in Greece. They found that human activities significantly influenced flood generation, highlighting the importance of land management and GIS-based flood risk analysis. Mudashiru et al. (16) reviewed flood hazard mapping methods, comparing physical, empirical, and modeling techniques. The study emphasized the need for accurate hazard assessment for better flood management strategies. Ahmed et al. (17) used remote sensing, GIS, and AHP techniques to identify flood hazard zones in Tripura, India, showing that AHP is an efficient and cost-effective method for flood risk assessment in data-scarce areas.

El-Bagoury and Gad (18) utilized HEC-HMS and HEC-RAS software for assessing flash flood hazards in two basins in southeastern Cairo, Egypt. By employing meteorological and morphological data and the SCS curve number method, the study calculated discharge volume, peak flow, and flood depth. Runoff volumes for 25-, 50-, and 100-year storms were estimated for An-Nawayah and Al-Rashrash basins. The study identified high, extreme, and medium flood risk zones and proposed constructing dams as a mitigation strategy. Javidan and Bahremand (19) applied the WetSpa model in the Ziyarat basin, Iran, to separate flood flow components. They found that 97% of flood water was base flow during events, providing insights into flood

flow dynamics. El-Saoud and Othman (20) applied GIS and hydrological modeling tools to assess flash flood impacts in the Wadi Mehassar basin near Makkah. Using ArcGIS, ERDAS Imagine, and WMS software, the study processed geological and meteorological data. HEC-1 and HEC-RAS models were utilized to simulate hydrograph curves and inundation extents, identifying 44% of the area as prone to flash flooding. The study highlighted the vulnerability of urban and sacred sites to short-duration, high-intensity rainfall. Cea et al. (21) examined an extreme flood event in Maputo, Mozambique, using high-resolution hydrologic–hydraulic models combined with global data sources. The study validated model outputs with observed data, highlighting the utility of integrated models for flood hazard assessment in data-scarce regions. Results emphasized the role of reservoirs in flood management and the need for enhanced flood risk plans. Hasan et al. (22) investigated urban flash flooding in Aur River catchment, Malaysia, employing XPSWMM software for hydrological and hydraulic simulation. The study used rainfall intensity-duration-frequency data to forecast peak flows and water levels. Results revealed the hydrological changes due to urbanization and emphasized the importance of integrating hydrological analysis in urban flood mitigation planning.

Venkata Rao et al. (23) employed the SWAT and HEC-RAS models to simulate floods in India's Nagavali and Vamsadhara basins. Using 48-hour lead-time rainfall forecasts, the study achieved high accuracy in predicting flood inundation extents and depths. The research highlighted the utility of integrated modeling systems in improving flood preparedness and management during tropical cyclones. Anselmo et al. (24) presented a multi-step integrated modeling approach for flood risk assessment at Montalto di Castro power plant, Italy. The study involved rainfall-runoff modeling, statistical rainfall analysis, and two-dimensional hydraulic modeling. The approach effectively simulated flood events and proved transferable to other flood-prone areas. Ruidas et al. (25) reviewed advancements in flood hazard modeling, focusing on machine learning and metaheuristic techniques. The study evaluated global flood mitigation strategies and highlighted gaps in knowledge regarding regional susceptibility patterns. The research emphasized developing advanced flood prediction models to support early warning systems and flood risk management. Fernández-Nóvoa et al. (26) reviewed global Flood Early Warning Systems (FEWS), focusing on numerical modeling integration. The study discussed the evolution of FEWS from basin-scale applications to global systems, emphasizing their role in mitigating climate-induced flood risks. Key advancements in FEWS development were highlighted as tools for resilience and disaster preparedness. Borga et al. (27) described the Hydrate project's efforts to enhance flash flood forecasting across Europe. The project developed advanced observation strategies and tools for ungauged basins, improving the understanding of flash flood mechanisms. The study underscored the importance of harmonized data and innovative methodologies for effective flash flood risk management. Kumar et al. (28) provided a comprehensive review of flood modeling techniques, including hydrologic, hydraulic, AI-based, and multi-criteria approaches. The study evaluated their effectiveness, challenges, and future prospects, advocating for integrated and advanced models to address data limitations and improve flood risk mitigation strategies.

Al-Omari et al. (29) developed a comprehensive flood hazard map for the King Talal Dam area in Jordan using remote sensing (RS) and GIS methodologies. By integrating key environmental and geographical factors with the Analytic Hierarchy Process (AHP), the study effectively categorized flood risk levels and provided a robust framework for flood risk assessment and disaster management. Yanfatriani et al. (30) analyzed the correlation between extreme weather indices and hydrometeorological disasters in Jambi Province using IMERG rainfall data and disaster records. Their findings revealed significant trends in extreme rainfall and their links to floods, landslides, and droughts, offering valuable insights for regional risk reduction and climate adaptation strategies. Supratman et al. (31) investigated the impact of land use changes on flood hazards in Jakarta's Pondok Karya area by combining Landsat imagery with hydrological and hydraulic modeling tools (HEC-HMS and HEC-RAS). The study highlighted increased flood risk due to urban expansion and vegetation loss, emphasizing the need for effective water management and flood mitigation interventions.

In recent years, numerous studies have focused on flood simulation and prediction, particularly using hydrological and hydraulic models, as well as statistical methods and GIS-based analyses. These studies have explored hydrological processes and flood prediction across various geographic and climatic scales. While many efforts have aimed at achieving more accurate flood simulations and predictions, the use of hydrological and hydraulic models, often in isolation or limited combination with other tools, remains prevalent. This limited integration has led to less accurate predictions and risk assessments in some regions. Given the complexities inherent in hydrological and hydraulic processes, employing advanced models alongside GIS analysis can significantly enhance flood prediction accuracy, improve flood-prone area identification, and optimize water resource management. Additionally, such an approach can lead to more effective crisis planning in vulnerable regions.

Despite significant advances in flood simulation and prediction, the integrated use of hydrological, hydraulic,

and GIS models has not been sufficiently explored in certain regions, particularly in areas with limited data availability. This gap underscores the need for new research that employs more comprehensive approaches to substantially improve prediction accuracy and flood risk management. While numerous studies have focused on hydrological or hydraulic modeling individually, the combination of these models with GIS and other advanced analytical techniques remains underexplored. The present study aims to address these gaps by proposing a holistic and innovative approach to flood analysis. By integrating hydrological and hydraulic models with GIS analysis, this research seeks to enhance the precision of flood simulations, provide a more accurate identification of high-risk areas, and contribute to the development of more effective flood management strategies, particularly in data-scarce and sensitive watersheds. Ultimately, this research will not only improve flood prediction accuracy but also support the creation of more robust flood mitigation strategies, water resource management, and disaster planning.

TABLE 1. Geographical and Equivalent Dimensions of the Studied Watershed

Parameter	Value
Area (km²)	250.2
Perimeter (km)	277
Watershed Length (km)	75
Minimum Elevation (m)	1166
Maximum Elevation (m)	2850
Average Elevation (m)	1598
Average Slope (%)	19
Main Stream Pure Slope (%)	0.7
Main Stream Average Slope (%)	1.4
Main Stream Length (km)	7
Equivalent Rectangle Length (km)	114
Equivalent Rectangle Width (km)	21.9
Diameter of Equal-Area Circle (km)	56.5

3. RESEARCH METHODS AND TOOLS

3. 1. Characteristics of Watershed and the Studied River Section The studied section of this river spans 30 kilometers and is located in Khorramabad County, west of the city of Khorramabad. The river flows predominantly in a northeast direction until it converges with the Kashkan River. The starting point of the studied section is located 8 kilometers southwest of the city of Khorramabad. Initially, the river passes through the vicinity of Qarq Chaqa village, and at the confluence of the Khorramabad River and the Kashkan River, known as Doab Veysian, it has the following geographic coordinates:

$X = 217898$
$Y = 3710118$

The watershed area of the studied region is 250 square kilometers, with a perimeter of 277 kilometers. The Khorramabad River serves as the access route to the Khorramabad section. The river flows parallel to this route and, after merging with the Kashkan River, which approaches from the northwest (Doab Veysian), continues along the studied section under the name of the Kashkan River, ultimately flowing toward the city of Poldokhtar.

3. 2. Research Method The WetSpa model, developed by VUB University of Brussels in Belgium, is a distributed and physically-based model designed to predict the transfer of water and energy among soil, vegetation, and the atmosphere at the watershed or regional scale with daily time steps (31). This model represents the hydrological system of a watershed, encompassing components such as the atmosphere, canopy layer, stem zone, transmission zone, and saturated zone. The watershed is divided into cells distributed heterogeneously. Each cell is divided into two main sections: bare soil and vegetated soil, both of which play significant roles in maintaining water and energy balance. The calibration of the WetSpa model was carried out using observed data from local meteorological stations, streamflow measurements, and soil moisture records. The model parameters, such as infiltration rates, evapotranspiration, and soil parameters, were calibrated through an iterative process that involved comparing model outputs with observed hydrological data and adjusting the parameters to minimize the model error. The optimization process used the Nash-Sutcliffe Efficiency (NSE) index and root mean square error (RMSE) to assess model performance, with the goal of achieving the best fit between simulated and observed data. Figure 1 illustrates a schematic representation of the hydrological processes modeled by WetSpa.

The WETSPASS model is employed to estimate long-term spatial patterns of groundwater recharge, which can serve as input for regional groundwater flow models and the analysis of regional groundwater flows. This model, developed by Batelaan and De Smedt (2001) based on the WetSpa model, is shown in Figure 2. The WetSpass model examines the transfer of water and energy among soil, vegetation, and the atmosphere under quasi-steady-state conditions. As a GIS-based distributed hydrological model, WetSpass is used to estimate evapotranspiration, surface runoff, and groundwater recharge on seasonal and annual time scales.

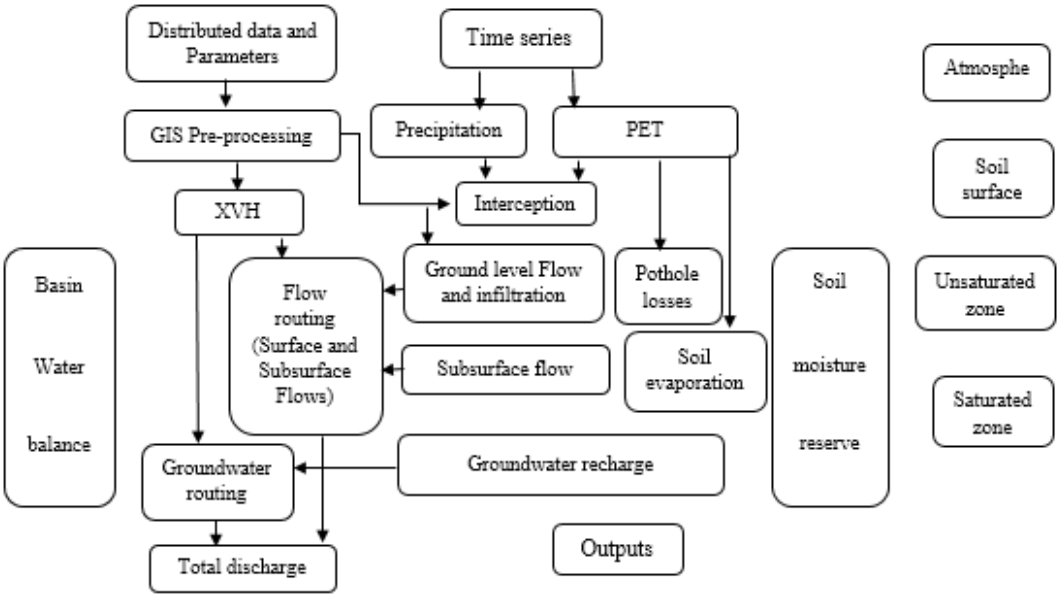


Figure 1. Main hydrological processes in the WetSpa model structure

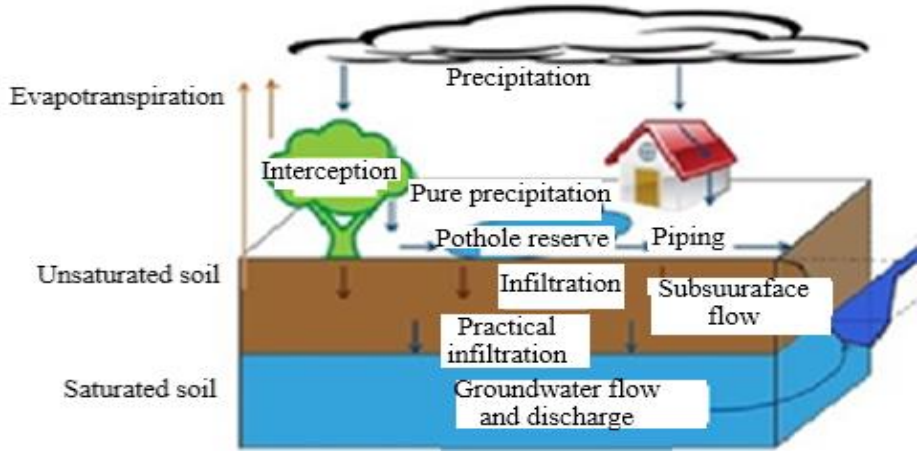


Figure 2. WetSpa model structure

The WetSpa Extension model is another GIS-based distributed hydrological model designed for flood prediction and water balance simulation at the watershed scale. This model has the capability to predict flow hydrographs at the outlet of a watershed or any other convergence point within the watershed and can simulate varying time intervals (16, 20). In addition to flood prediction, the model investigates the causes of flooding, particularly the spatial distribution of topography, land use, and soil types.

Finally, the LISFLOOD-FP hydrodynamic model is a hybrid model that simulates channel flows in one dimension and overland flows in floodplains in two dimensions. The inputs for this model are files in Ascii format that are executed in the DOS environment. The outputs of this model, also in Ascii format, include water depth at each pixel. The LISFLOOD-FP model was calibrated by adjusting parameters related to the channel flow resistance, overland flow roughness, and floodplain infiltration. The model calibration process involved using observed flood extent and water depth data, adjusting parameters through a trial-and-error method, and employing statistical measures like RMSE and the coefficient of determination (R^2) to evaluate model performance.

Regarding the performance of the WetSpa model in low-flow conditions, it has been noted that the model's accuracy decreases when simulating low-flow events. To address this limitation, model performance can be improved by refining the parameterization of infiltration

processes, particularly in areas with low runoff potential. Additionally, implementing more sophisticated land surface parameterization, such as considering the effect of soil compaction and vegetation changes on infiltration, can improve model performance during low-flow conditions. It is also recommended to integrate more local data on soil moisture and evapotranspiration to better capture the dynamics of water movement in low-flow scenarios.

4. RESEARCH FINDINGS

4. 1. Hydrological Model Calibration The calibration of the hydrological model was conducted using data from the first five years (water years 2005-2006 to 2009-2010), incorporating daily records of precipitation, temperature, and potential evapotranspiration. These datasets were extracted from a comprehensive 10-year record spanning water years 2005-2006 to 2014-2015. Additionally, daily discharge measurements from the Cham Anjir hydrometric station, located at the drainage basin outlet, were employed to compare simulated hydrographs with observed flow data during the calibration period. Calibration was performed manually using a trial-and-error approach, adjusting 11 global parameters to optimize the agreement between simulated and observed hydrographs. Both graphical and statistical comparisons were utilized to evaluate model performance. Initial parameter values were determined based on model structure, study area characteristics, and key hydrological variables. These values were subsequently refined to enhance the model's representation of observed conditions. The final calibrated parameter values are presented in Table 2, with a summary provided in Table 3. These values represent optimal estimates for simulating hydrological processes in the study area. Figure 3 illustrates a comparative analysis of the simulated and observed daily flow hydrographs, demonstrating a strong alignment between the datasets. This successful calibration validates the model's ability to replicate real-world hydrological behavior, confirming its reliability and applicability for simulating hydrological conditions in the study area.

Table 2 outlines the initial values and variation ranges of key parameters in the WetSpa model, which are essential for simulating hydrological processes. The ki parameter (sub-surface flow factor) has an initial value of 0.007, ranging from 0.1 to 10, allowing flexibility in soil permeability representation. The kg parameter (groundwater loss coefficient) starts at 0.0001, with a broad variation range (1×10^{-7} to 0.1), enabling adjustments based on local groundwater dynamics. Kss, the raw initial moisture parameter, is set at 1, ranging from 0.1 to 2, influencing soil moisture retention and runoff rates. The kep parameter (evapotranspiration correction factor) is initialized at 0.1, with a range from 0.3 to 2, allowing adaptation to vegetation and climate conditions. The Go parameter (initial groundwater reserve) is set at 10 mm, with a Gmax value reaching 5,550,000 mm, providing significant flexibility in representing groundwater storage. The To parameter (threshold degree coefficient for snowmelt) ranges from -3 to 2, accommodating diverse climatic conditions, while the ksnow coefficient (degree-day coefficient) varies from -3 to 3.2, affecting snowmelt dynamics.

The krain parameter (precipitation degree-day coefficient) ranges from 0.001 to 0.3, governing the impact of precipitation on snowmelt. The Krun parameter (surface runoff capacity) spans from 4 to 7, influencing runoff generation. Finally, the pmax parameter (maximum precipitation severity) varies from 100 mm to 5000 mm, enabling the simulation of extreme weather events. In summary, these parameters provide substantial

TABLE 2. Initial parameters and values and their ranges of variations

Sign	Parameters	Initial value	Min.	Max.
k_i	Sub-surface flow factor (-)	0.007	0.1	10
k_g	Groundwater loss coefficient (d-1)	0.0001	0.1*10-7	0.1
k_{ss}	Raw initial moisture (-)	1	0.1	2
k_{ep}	The correction factor of potential evapotranspiration (-)	0.1	0.3	2
G_o	Initial groundwater reserves (mm)	10	0	100
G_{max}	Max. initial groundwater reserve (mm)	5550000	100	10000
T_o	Threshold degree coefficient (c0)	0	-3	2
k_{snow}	Degree-day coefficient (mm c0 d-1)	0.045	-3	3.2
k_{rain}	Precipitation degree-day coefficient (mm mm c0 d-1)	0	10-3	0.3
K_{run}	Surface runoff capacity (-)	0.001	4	7
p_{max}	Max. precipitation severity (mm)	100	100	5000

flexibility, allowing the WetSpa model to be effectively calibrated for diverse environmental conditions, ranging from minimal groundwater recharge zones to regions experiencing significant precipitation and snowmelt events.

Table 3 presents the calibrated values of 11 global parameters for the WetSpa model, reflecting key hydrological processes such as infiltration, runoff, evapotranspiration, and groundwater dynamics. The ki value of 1 suggests a standard infiltration rate, allowing for effective water absorption into the soil, which is supported by the moderate kss value of 0.44, indicating substantial subsurface water storage. However, the low kg value (0.00003) implies limited groundwater recharge, suggesting that groundwater does not significantly contribute to the basin's water balance. The Go value of 140 and Gmax of 1020 highlight the capacity for groundwater storage, although recharge is slow due to the low kg coefficient. The kep value of 0.29 indicates moderate evapotranspiration, typical for areas with mixed vegetation. The low To value of 0.06 reflects initially dry soil, leading to higher runoff early in the rainfall events. With ksnow at 0.042, the snowmelt rate is low, which aligns with the similarly low krain value of 0.007, suggesting minimal runoff from rainfall. The Krun value of 3.8 suggests that runoff can become significant during intense rainfall events, and the high pmax value of 3000 enables the model to simulate extreme precipitation and flood conditions. Overall, the calibrated parameters indicate a model suited for regions where groundwater storage plays a role, but recharge is slow, and surface water dynamics, particularly runoff, are more significant during heavy rainfall events.

4. 2. WetSpa Model Validation Results The WetSpa model was validated using the global parameter values obtained during the calibration phase over a 5-year period (water years 2009–2010 to 2014–2015). The validation process involved comparing the observed daily flow hydrographs with those simulated by the model. The results of this comparison are presented in Figure 4, which offers a comprehensive overview of the model's performance during the validation phase.

The output files generated by the model provide a comprehensive view of the basin's water balance, including surface flows, subsurface flows, base flows, and total flow at each time step. These outputs are critical for assessing the model's performance and understanding the movement of water through the basin. According to the model's results, during the calibration period, 26.15% of the total precipitation is converted into runoff. This runoff is divided into surface runoff (2.37%), subsurface flow (1.61%), and groundwater discharge (base flow), which constitutes the largest portion at 22.17%.

In the validation period, the total runoff share from precipitation is slightly higher, reaching 26.42%. This slight increase suggests that the model's performance was consistent, with similar runoff dynamics observed in both

TABLE 3. Values of 11 global parameters of the WetSpa model determined in the calibration phase

Sign	Calibrated Numerical Value
ki	1
kg	0.00003
kss	0.44
kep	0.29
Go	140
Gmax	1020
To	0.06
ksnow	0.042
krain	0.007
Krun	3.8
pmax	3000

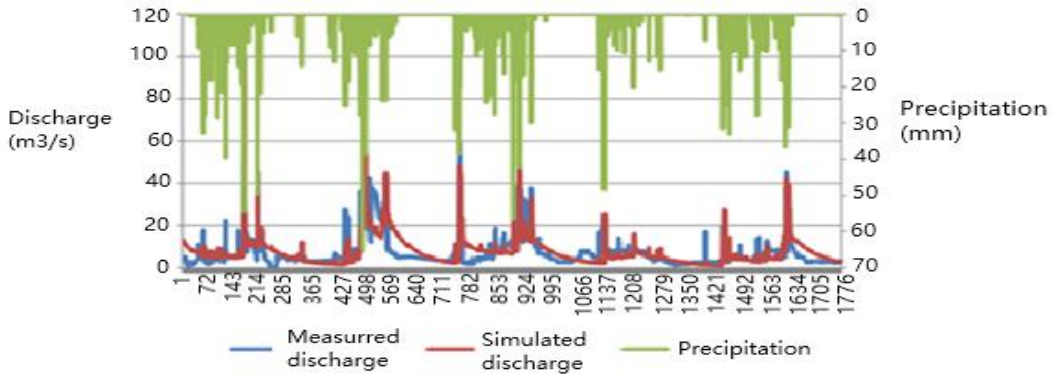


Figure 3. Comparison of the observational and simulated daily discharge hydrographs by the WetSpa model in the calibration period

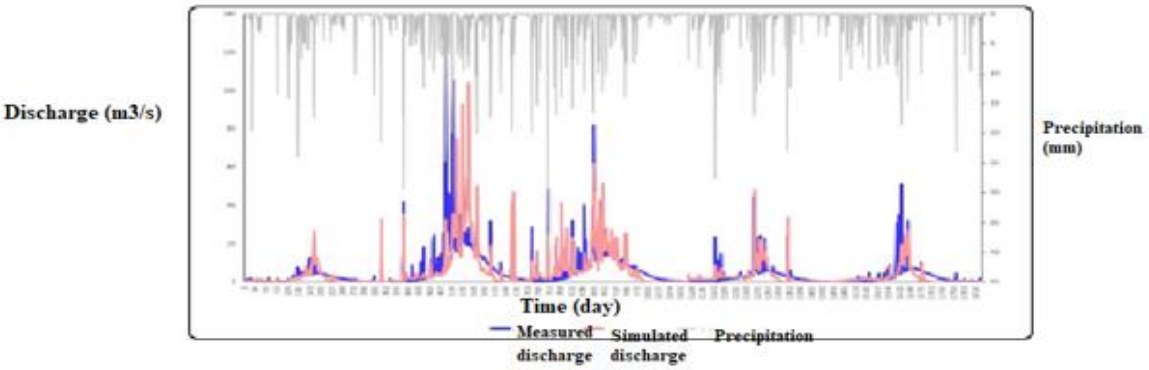


Figure 4. Comparison of the observational and simulated daily discharge hydrographs by the WetSpa model in the validation period

periods. Additionally, the model simulation results show that the ratio of evaporation to precipitation is 57.18% in the calibration period and increases to 69.20% in the validation period, indicating a higher rate of evapotranspiration during the validation phase.

Table 4 presents the water balance components calculated by the model for both the calibration and validation periods, offering a clear comparison of these key hydrological processes over time.

The variables are defined as follows: P represents the observed precipitation (mm), I denotes interception losses (mm), Ds is the change in soil moisture between the start and end of the time step (mm), F refers to infiltration losses (mm), E is evapotranspiration (mm), $PERS$ represents the main depth of infiltration outside the stem zone (mm), Ri denotes surface runoff (mm), Rg represents groundwater flow (mm), R is the total runoff (mm), and Dg refers to the change in groundwater reserves between the start and end of the time step (mm). All these components were thoroughly explained in the preceding section.

TABLE 4. Water balance components calculated by the model in the calibration period

	Sum(mm)	Precipitation (0.0)	Mean
P	2779.4	-	1.522
I	133.1	4.79	0.073
D_S	42.6	1.53	172.851
F	2067.1	74.37	1.132
E	1589.2	57.18	0.780
PERS	767.6	27.62	0.420
R_S	65.9	2.37	0.036
R_I	44.8	1.61	0.025
R_G	616.1	22.17	0.0337
R	726.9	26.15	0.398
D_G	-55.2	-1.98	108.793

4. 3. Evaluating the Efficiency of the WetSpa Model

The efficiency of the model in simulating river flows was assessed using two widely recognized performance metrics: the Nash-Sutcliffe efficiency criterion (Equation 3) and the Kling-Gupta efficiency (KGE) criterion (Equations 1 and 2). Additionally, a visual comparison of the simulated and observed hydrographs was performed to further evaluate the model's performance. The Nash-Sutcliffe efficiency (NSE) is a commonly used indicator of the model's predictive ability, with higher values reflecting better simulation accuracy. The Kling-Gupta efficiency (KGE), which incorporates key components such as correlation, bias, and variability, offers a more comprehensive assessment of model performance. Both metrics were employed to ensure a robust evaluation of the model's capability to replicate observed river flow dynamics (Tables 5 and 6).

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2}$$

$$KGE' = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\gamma-1)^2}$$

$$\beta = \frac{\mu_s}{\mu_o}$$

$$\alpha = \frac{\sigma_s}{\sigma_o}$$

$$\gamma = \frac{CV_s}{CV_o} = \frac{\sigma_s/\mu_s}{\sigma_o/\mu_o}$$

Where r is the correlation coefficient between simulated and observational discharges, α the standard deviation ratio of simulated values (σ_s) to the standard deviation of observational values (σ_o), and β the mean ratio of simulated values (μ_s) to the mean observational values (μ_o). The best value for the Kling-Gupta assessment criterion is 1, which indicates full correspondence with hydrographs.

According to the modified Kling-Gupta (KGE') criterion, γ is used instead of β, which is the ratio of the coefficient of simulated value changes (CV_s) to the coefficient of observational value changes (CV_o).

$$NSE = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}$$

TABLE 5. Classification of WetSpa model efficiency

Class	Excellent	Very good	Good	Weak	Very weak
Collective criterion	<85	0.7-0.8	0.7-0.55	0.55-0.4	0.4<

TABLE 6. Model assessment results based on statistical indices in calibration and validation periods

Efficiency criterion	Calibration	Validation
Nash-Sutcliffe criterion	0.53	0.48
Collective criterion	0.703	0.723
Kling-Gupta criterion	0.68	0.74

Where *NSE* is the Nash-Sutcliffe index, O_i observational discharge, P_i the predicted discharge, and $\overline{O}$ the mean observational discharges.

4. 4. LISFLOOD-FP Hydraulic Model

Inputs Required for Model Implementation

4. 4. 1. Digital Model of Area Under Study The model requires a high-resolution digital elevation model (DEM) of the study area, with an appropriate cell size and accuracy to effectively capture height variations and key features of both the main channel and the floodplain. The DEM should cover a sufficiently large portion of the floodplain, including areas that are likely to be affected by flooding, to ensure a comprehensive simulation of flood dynamics and accurately represent flood risk zones. The spatial resolution of the DEM must be high enough to capture fine-scale features that influence water flow, such as floodplain depressions and embankments.

4. 4. 2. Map of the Floodplain Roughness Coefficient In areas where the floodplain does not exhibit uniform roughness, a spatially variable map of roughness coefficients must be provided. This can be achieved by assigning Manning's roughness coefficient (n) values to individual cells within a raster map, aligned with the DEM in terms of cell size and spatial resolution. This map plays a critical role in modeling the varying resistance to flow across different floodplain areas, which can significantly affect flood behavior and water distribution during high-flow events.

4. 4. 3. River Characteristics A river characteristics file is essential for defining the geometry and hydraulic properties of the river. This file includes cross-sectional profiles of the river, with the number of sections determined by changes in the river's characteristics. Each section should include specific coordinates, Manning's roughness coefficient, and the dimensions (width and depth) of the riverbed or subsurface. Accurate flow simulations depend on properly defining these river characteristics, which help to predict the flow dynamics and flood extent within the river and its floodplain.

4. 4. 4. Defining Boundary Conditions for Model Implementation Proper boundary conditions are necessary for the model's accurate implementation. The upstream boundary should be specified using an inflow hydrograph or a constant discharge value, depending on the available data. For the downstream boundary, conditions can be defined based on downstream water levels or the outflow hydrograph if a measurement station is available. In the absence of such data, boundary conditions may be inferred by calculating the normal depth and measuring the global slope of the river section. These calculations can assist in defining the downstream behavior of the flow, ensuring accurate flood predictions.

4. 5. Model Outputs

The key outputs generated by the model are as follows:

4. 5. 1. Mass Balance File This output provides a comprehensive summary of the flood event, including the extent of the flooded area, water depths, and outflow discharges in the downstream areas of the model. The mass balance file is produced for a specified time interval, which can be determined by the user, and is crucial for understanding the flow dynamics and changes over time within the study area. This file helps identify peak flood events and the associated flow characteristics, which are vital for flood risk management.

4. 5. 2. Water Depths and Water Levels for Each Pixel The model generates water depths and water level heights for each pixel within a raster network at specific time intervals defined by the user. This granular output provides detailed spatial information about water levels and depths across the entire model domain. By analyzing these outputs, researchers can pinpoint flood-prone areas and assess the dynamics of water distribution during the simulation. This information is essential for identifying regions that experience deeper or more widespread flooding.

4. 5. 3. Maximum Water Levels and Maximum Water Depths The model also outputs the maximum water level height and maximum water depth predicted for each pixel over the entire simulation period. This output is valuable for identifying areas most susceptible to extreme flooding and understanding the peak flood conditions across the entire study area. The maximum values can help in delineating flood-prone areas that may require special attention in terms of mitigation and preparedness strategies.

4. 5. 4. Channel Water Level Profile A channel water level profile is generated at specific intervals set by the user. This profile provides an in-depth view of the water levels along the river channel, helping to assess the river's behavior under different flow conditions. By analyzing this output, stakeholders can gain insights into the hydrodynamics at various locations within the channel, such as the impact of bends, obstructions, or changes in riverbed elevation on water levels during floods.

The outputs generated by the LISFLOOD-FP model provide detailed information on flood dynamics, such as water levels, water depths, and flood extents. However, integrating this data with socio-economic information is essential for effective flood risk management. Understanding the spatial distribution of flood-prone areas is key to assessing the potential social and economic impacts of flooding in vulnerable communities.

The model's ability to identify areas with varying degrees of flood susceptibility, based on different return periods, can be enhanced by incorporating socio-economic factors such as population density, land use, infrastructure, and economic activity in the floodplains. Mapping these factors in relation to the model outputs can help in identifying high-risk zones where floods may lead to significant economic losses, displacement, and disruption to critical services. Furthermore, by assessing the vulnerability of specific sectors, such as agriculture, housing, and infrastructure, policymakers can prioritize areas for flood mitigation and adaptation strategies.

Incorporating socio-economic analysis into flood risk mapping can also facilitate the development of targeted flood protection measures, including the relocation of vulnerable populations, flood-proofing infrastructure, and improving emergency response plans. The integration of these insights can make flood risk management strategies more effective by addressing both the physical and socio-economic dimensions of flooding.

4. 6. LISFLOOD-FP Hydraulic Model Validation Results In this study, the LISFLOOD-FP hydraulic model was validated using two distinct flood events with discharges of 72.64 m^3/s and 207.86 m^3/s (Table 7). These events were selected to represent moderate and extreme flooding scenarios, providing a basis for assessing the model's accuracy in predicting flood dynamics. The first event, with a discharge of 72.64 m^3/s, was chosen to evaluate the model's performance under typical flood conditions. The second event, with a higher discharge of 207.86 m^3/s, was used to test the model's ability to simulate flood behavior during more severe events. By comparing the model's simulated water levels with actual observed data from gauge stations, the validation aimed to assess the accuracy of the model's flood predictions. The results helped to identify the model's strengths in simulating flood extents and water levels, as well as any discrepancies, which may have been due to factors such as gauge accuracy or sediment accumulation. The validation process is essential to understanding the reliability of the LISFLOOD-FP model under different discharge conditions and flood intensities, and to refine its predictive capabilities for future flood risk management.

4. 7. Evaluating the Hydraulic Model in the Validation Phase based on Error Percentage Index The discrepancy observed in the model's performance, particularly for the highest discharge of 207.86 m^3/s, can be attributed to several potential factors. One primary source of error could be the accuracy of the water gauge used for validation. The gauges may be subject to technical issues such as malfunctioning sensors, which can affect the accuracy of the recorded water levels. Additionally, the accumulation of sediment at the base of the gauge is another factor that can lead to incorrect readings of the water level. This sediment buildup could obstruct the gauge's proper functioning, causing discrepancies between the actual and simulated water heights. Moreover, the quality of the input data used by the model—such as the Digital Elevation Model (DEM), roughness coefficients, and boundary conditions—could also contribute to model discrepancies. Inaccuracies in these parameters can lead to small errors in the predicted water levels and flood extents. Despite these potential sources of error, the model performed relatively well overall, demonstrating good predictive capability under most conditions.

The observations highlight the need for further refinement in both the model and the quality of the input data, particularly under extreme discharge scenarios. Improving the accuracy of these parameters could enhance the model's reliability for future flood predictions and risk assessments. Figures 5 and 6 provide visual representations of the simulated flood zones during the validation phase. These figures correspond to

TABLE 7. Model error rate of various discharges in several implementations

Discharge (m^3/s)	Water gauge height (m)	Water height in the model at the station (m)	Error rate (m)	Error percentage
50.01	2.31	2.37	0.06	2.95
72.64	2.39	2.43	0.04	1.67
207.86	3.3	3.81	0.15	15.45

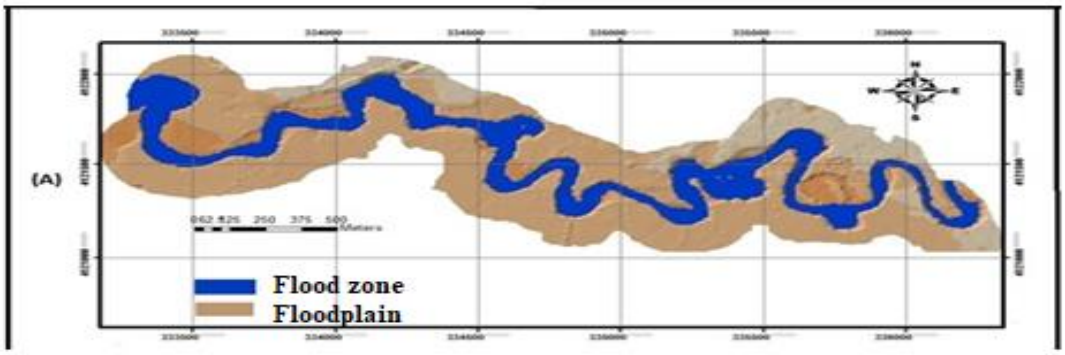


Figure 5. Flood zone simulated by the model with a discharge of 72.64 m^3/s

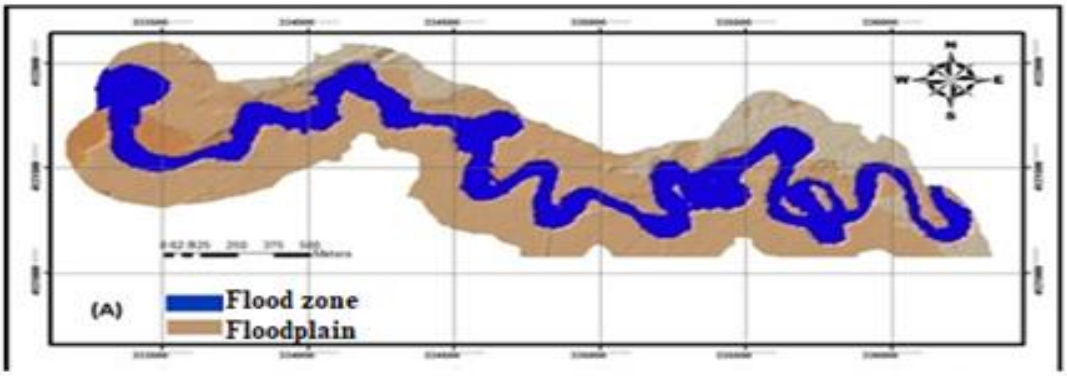


Figure 6. Flood zone simulated by the model with a discharge of 207.86 m^3/s

the two discharge events of 72.64 m^3/s and 207.86 m^3/s. The simulation results show that, in general, the flood zones align well with the areas that are expected to be prone to flooding under the given conditions. However, some minor discrepancies remain, likely due to the previously discussed factors such as sediment accumulation and input data uncertainties. By comparing the observed flood extents with the model's simulations, we can pinpoint areas where the model performed well and areas that may require further improvement. This analysis is essential for refining the model and ensuring more accurate flood risk management in future studies.

4. 8. Evaluating Integration of Two Hydrological and Hydraulic Models To assess the effectiveness of the integration between the hydrological and hydraulic models, three flood events were analyzed. These events had already been used for the calibration and validation of the hydraulic model. In the standalone application of the hydraulic model, observed hydrographs were directly used as input. However, in the integrated model approach, simulated hydrographs generated by the hydrological model were used as inputs to the hydraulic model, allowing for the simulation of the flood zones.

The integration code was applied separately for each year in which the analyzed flood events occurred. This method not only facilitated the simulation of flow hydrographs but also ensured a seamless transfer of the simulated hydrographs from the hydrological model to the hydraulic model through the integration code. This integration aimed to enhance the accuracy of flood simulations by combining the strengths of both models.

Figure 7 presents the daily flow hydrograph for a flood event with a discharge of 72.64 m^3/s, simulated using the integration code. The simulated hydrograph from the hydrological model was automatically transferred to the hydraulic model, where it was used as an upstream boundary condition for the flood simulation. This approach streamlined the process and ensured that the hydraulic model received accurate input for better flood zone predictions.

Figure 8 shows the final outlet conditions for the studied flood, illustrating the effectiveness of the integrated approach. The results demonstrate that the integration code successfully facilitated the transfer of data between the hydrological and hydraulic models, ensuring accurate simulations of the flood event.

Moreover, the findings of the LISFLOOD-FP model, particularly in identifying flood-prone areas for different return periods, offer valuable insights into areas most vulnerable to flooding. However, providing further insights into the spatial distribution of these flood-prone

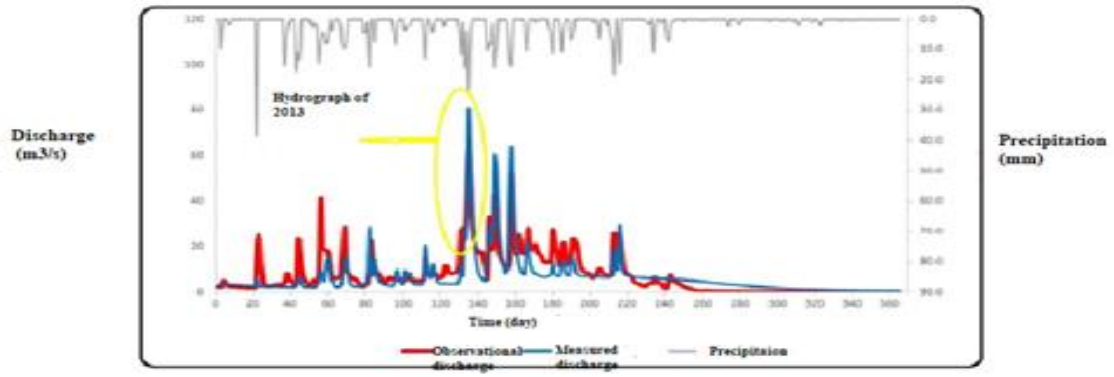


Figure 7. Observational and simulated daily flow hydrographs produced through the integration code of two models about the flood with a discharge rate of 72.64 m^3/s

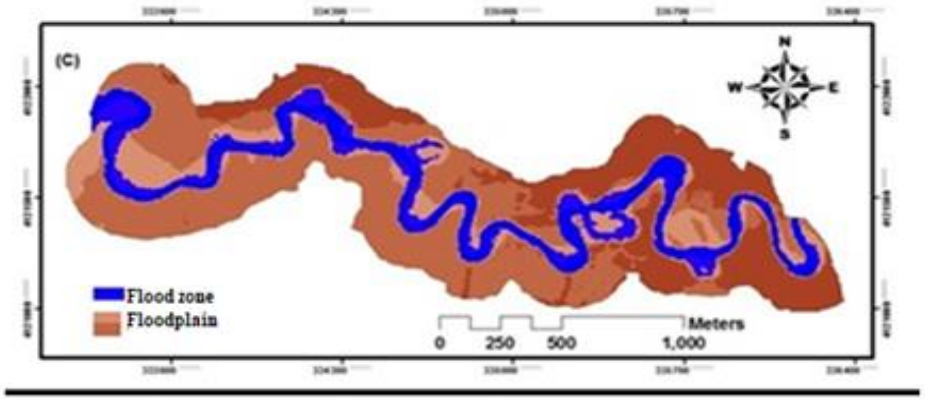


Figure 8. Flood zone with a discharge rate of 72.64 m3/s caused by a simulated inflow hydrograph

areas and their potential socio-economic impacts can significantly enhance risk management strategies. Understanding the locations and characteristics of high-risk zones can guide more targeted and effective flood mitigation efforts, helping to prioritize resources and interventions in the most vulnerable regions. The integration of both models, therefore, not only improves flood simulations but also contributes to the development of more comprehensive and informed flood risk management strategies, incorporating socio-economic factors for more holistic planning.

4. 9. Estimates of Discharges with Different Return Periods and Simulation of Flood Zones

In this study, the discharge values corresponding to return periods of 25, 50, 100, and 200 years were calculated based on a 50-year dataset that includes the maximum annual instantaneous discharges recorded at the Cham Anjir Hydrometric Station. These calculations were carried out using SMADA software, which facilitated the comparison of various statistical distributions to determine the most suitable model for the data. The chi-square goodness-of-fit tests were employed to assess the fit of the statistical distributions, ultimately leading to the selection of the Log-Pearson Type III distribution as the most appropriate model for the observed discharge data.

Following the determination of discharge estimates for the different return periods, the calibrated hydrological model was used to simulate the corresponding flood zones for each return period. The model generated spatial flood extents and water depths, providing a comprehensive view of flood behavior for the selected return periods. This simulation is essential for assessing the varying magnitudes of flood events and understanding the potential impacts of flooding on the surrounding area.

Table 6 presents the simulated results, including the maximum water depths and the total inundated areas for floods associated with different return periods. These outputs are critical for understanding the extent and severity of flood hazards under different scenarios and serve as a foundation for developing flood risk mitigation strategies. The results also provide valuable information for designing flood-control infrastructure, ensuring that the region is better prepared for future flood events.

In addition to providing detailed flood predictions, this study integrates advanced statistical methods with hydrological modeling to offer a robust framework for estimating flood risks and their spatial impacts. This integrated approach enhances the accuracy of flood predictions and contributes to more effective water resource management and disaster preparedness. By incorporating various return periods, the study also helps in understanding long-term flood risks and their implications for vulnerable areas, guiding informed decision-making and efficient resource allocation for flood risk management.

The analysis of Table 9 reveals a clear and direct relationship between discharge rates and the corresponding simulated flood zone area for return periods of 25, 50, 100, and 200 years. As the return period increases, the discharge rates also increase, from 414 m^3/s for a 25-year flood event to 620 m^3/s for a 200-year event. This increase in discharge is accompanied by a rise in the simulated flood zone area, from 539,800 m^2 for the 25-year event to 627,100 m^2 for the 200-year event. This pattern is consistent with the hydrological principles that suggest higher discharge events result in the inundation of larger areas, highlighting the escalating risk of flooding during more extreme events. Interestingly, the rate at which the flood zone area expands decreases with increasing return periods. For example, the area increases by 41,100 m^2 (7.6%) between the 25- and 50-year events, but only by 20,900 m^2 (3.4%) between the 100- and 200-year events. This diminishing rate of expansion implies that there are topographical and morphological limitations within the basin that restrict the extent of floodplain expansion despite significant increases in flow volumes. These constraints indicate that the landscape's capacity to accommodate larger flood zones becomes more limited as discharge levels rise.

The findings underscore the importance of considering both the magnitude of discharges and the spatial extent of inundation in flood risk management. The outputs generated by the calibrated model offer valuable insights into areas at higher risk of flooding, especially in extreme hydrological scenarios like the 200-year return period, which exhibits the largest zone area and discharge. This information is crucial for the development of targeted flood risk mitigation measures, land-use planning, and the design of infrastructure capable of withstanding extreme flood events. Ultimately, these results contribute to more informed decision-making in flood-prone regions, ensuring resilience and preparedness in the face of increasing flood risks.

TABLE 9. Maximum water depth and the zone area of floods with different return periods

Return period (year)	25	50	100	200
Discharge (m^3/s)	414	484	552	620
Model-simulated zone area (m^2)	539800	580900	606200	627100

5. DISCUSSION

The WetSpa model effectively simulated hydrological processes within the Khorram Abad basin, accounting for topographical, land use, and soil properties. These results align with the findings of Rad et al. (10), where the HEC-RAS model was used to assess flood-prone areas in the same region, underlining the importance of accurate flood mapping. The water balance analysis also revealed that 25.5% of the total precipitation in the basin was converted into runoff, with 2.34% as surface runoff, 1.6% as subsurface runoff, and 21.57% as groundwater discharge. However, the model performed less accurately for low-flow conditions compared to high-flow events. This discrepancy was largely due to the use of structural relationships for baseflow calculations and the limitations of the Thiessen polygon method for spatial precipitation distribution, a challenge similar to those reported in studies like Madi et al. (6) on spatial precipitation methods in Algeria.

The LISFLOOD-FP model was employed on a 7-kilometer stretch of the Khorram Abad River to simulate flood-prone areas under return periods of 25, 50, 100, and 200 years. The results indicated that as return periods increased, the flood extent also expanded, with the 200-year return period (discharge of 620 m^3/s) inundating 627,100 m^2. However, the rate of flood zone expansion diminished at higher return periods, likely due to topographical and morphological limitations within the basin. These findings corroborate the results reported by Ruidas et al. (25), who observed similar dynamics in the Gorganroud basin, emphasizing the significant role of basin-specific topography in flood modeling.

The integration of the WetSpa and LISFLOOD-FP models was implemented using Python-based code, facilitating the seamless transfer of hydrological outputs as inputs for the hydraulic model. This integration allowed for three different simulation scenarios:

1. Steady-flow simulations using constant peak discharge.
2. Unsteady-flow simulations using dynamic hydrographs over defined time steps.
3. Custom flow scenarios with user-defined time steps demonstrated excellent performance in simulating flood hydrographs and delineating flood-prone areas. This integration approach aligns with findings from Tsegaye et al. (12), where the combination of GIS and hydrological-hydraulic models proved effective in flood prediction in Florida. Moreover, the integration facilitated better management of spatial and temporal flood data, improving prediction accuracy and offering critical insights for flood management. The

study shows that integrating WetSpa and LISFLOOD-FP provides a powerful framework for simulating and managing flood events in sensitive watersheds. The models captured hydrological and hydraulic processes with accuracy, identified high-risk flood-prone areas, and contributed valuable insights for flood management strategies.

Based on the findings of this research, several recommendations can be made to enhance flood risk management and water resource planning:

Applying the Integrated Models in Other Watersheds: Extending the application of the integrated WetSpa and LISFLOOD-FP models to other basins across Iran would help identify flood-prone areas in diverse geographical contexts. However, careful consideration should be given to the hydrological and topographical characteristics of each region when transferring this modeling approach. Regions with significantly different climate, land use, or topographical features may require adjustments in model parameters, calibration, and validation methods. For instance, areas with high urbanization might require additional considerations of impervious surfaces, while regions with significant groundwater flow may need refinement in subsurface runoff simulations.

Leveraging Advanced Tools for Improved Simulation: Incorporating advanced technologies such as remote sensing and three-dimensional spatial data can further enhance simulation accuracy, particularly in regions where detailed topographical and land-use data is limited. The use of more precise elevation models and land cover datasets can help refine flood zone simulations and make predictions more reliable.

Designing Resilient Infrastructure: The results of this study can be used to inform the design of flood-resistant infrastructure, ensuring that critical infrastructure in flood-prone areas can withstand extreme events, particularly those with high return periods. Understanding the spatial extent of flooding can help in better floodplain management and design of flood barriers, drainage systems, and flood detention ponds.

Implementing Land-Use Planning: The insights gained from flood zone simulations can assist in better land-use planning, particularly in flood-prone regions, to reduce exposure and vulnerability to flood risks. By incorporating flood risk maps into urban and rural planning processes, regions can develop more sustainable practices that minimize flood risks while ensuring the continued growth of urban areas.

Impact of Climate Change: The potential impacts of climate change on flood modeling, particularly concerning changing precipitation patterns and flood intensities, should be further explored. Future studies could investigate how changing rainfall patterns and the intensity of extreme events may affect the flood-prone areas identified in this study. This would help adapt flood management strategies to future climatic conditions.

Challenges in Model Integration: A notable challenge faced during the integration of WetSpa and LISFLOOD-FP was related to computational time and the compatibility of data formats. The transfer of large datasets between the models required careful preprocessing to ensure compatibility, particularly when dealing with the spatial-temporal data outputs from the WetSpa model. Optimization of the integration process to reduce computation time is a potential area for improvement.

6. CONCLUSION

This study successfully integrated the WetSpa hydrological model and the LISFLOOD-FP hydraulic model to simulate daily flows, flood hydrographs, and flood-prone areas in the Khorram Abad drainage basin. Using a decade of meteorological data (2005–2006 to 2014–2015) as model inputs, the calibration was conducted for the first five-year period (2005–2006 to 2009–2010), while the validation relied on data from the subsequent five-year period (2009–2010 to 2014–2015). The WetSpa model demonstrated high accuracy in simulating river flows and flood hydrographs, with key performance metrics such as the Nash-Sutcliffe Efficiency (NSE) of 0.53 and the Kling-Gupta Efficiency (KGE) of 0.68. Notably, for high-flow events, the NSE reached 0.7, highlighting the model's robust capability for flood prediction.

Climate Impact on Hydrological Modeling: Future research should also consider the potential impact of climate change on hydrological modeling and flood predictions. Shifts in precipitation patterns, such as more intense rainfall or extended dry periods, could significantly alter flood dynamics and exacerbate the frequency and intensity of extreme flood events, which should be incorporated into the models for future projections.

By providing accurate flood zone simulations and offering practical recommendations, this research contributes significantly to reducing flood risks and improving water resource management in vulnerable regions like Khorram Abad. The integration of hydrological and hydraulic models has proven to be an effective tool in understanding and mitigating the impacts of flood events, fostering more informed decision-making and improved disaster preparedness.

7. REFERENCES

1. Karimi P, Safaval PA, Behzadi S, Azizi Z, Zarkash MMK, Kalashami HK. Flood risk zoning using geographical information system case study: Khorramabad Flood in April 2019. Acta hydrotechnica. 2022;35(63):89-100. https://doi.org/10.15292/acta.hydro.2022.07

2. Shayan S, Dehestani H, Hosein Zadeh MM. Geomorphologic classification of rivers in mountainous basins (case study: Khorramabad River). Quantitative Geomorphological Research. 2018;6(2):133-47.
3. Khaleghi M, Zeinivand H, Haghizadeh A. Sensitivity analysis of the spatially distributed hydrological WetSpa model parameters in estimation of daily flow in Khorramabad river. Iranian Water Researches Journal. 2016;10(4):143-8.
4. Robati M, Najafgholi P, Nikoomaram H, Vaziri BM. Zoning of Critical Hubs of Climate Change (Flood-Drought) Using the Hydrologic Engineering Center-Hydrologic Modeling System and Copula Functions Case study: Khorramabad Basin. 2024. https://doi.org/10.21203/rs.3.rs-5390435/v1
5. Pralle S. Drawing lines: FEMA and the politics of mapping flood zones. Climatic Change. 2019;152(2):227-37. https://doi.org/10.1007/s10584-018-2287-y
6. Madi H, Bedjaoui A, Elhoussaoui A, Elbakai LO, Bounaama A. Flood vulnerability mapping and risk assessment using hydraulic modeling and GIS in Tamanrasset valley watershed, Algeria. Journal of Ecological Engineering. 2023;24(7). https://doi.org/10.12911/22998993/163252
7. Di Baldassarre G, Schumann G, Bates PD, Freer JE, Beven KJ. Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches. Hydrological Sciences Journal–Journal des Sciences Hydrologiques. 2010;55(3):364-76. https://doi.org/10.1080/02626661003683389
8. Ghosh A, Chatterjee U, Pal SC, Towfiqul Islam ARM, Alam E, Islam MK. Flood hazard mapping using GIS-based statistical model in vulnerable riparian regions of sub-tropical environment. Geocarto International. 2023;38(1):2285355. https://doi.org/10.1080/10106049.2023.2285355
9. Jha V, Patel J, Sawant V, Tandel Y. Flood Hazard Assessment of Water-Front Geosynthetic Reinforced Soil Wall for Dam Regulation Rule Level. International Journal of Engineering Transactions A: Basics. 2023;36(4):640-8. https://doi.org/10.5829/ije.2023.36.04a.04
10. Rad M, Vafakhah M, Gholmalifard M. Flood mapping using HEC-RAS hydraulic model in part of Khorramabad watershed. Journal of Natural Environmental Hazards. 2018;7(16):211-26. https://doi.org/10.22111/jneh.2017.3343
11. Hassas H, Azizian A, Ghasemi M. A Simple Model for determining flood hazard areas. Journal of Natural Environmental Hazards. 2020;9(25):81-100. https://doi.org/10.22111/jneh.2020.31011.1549
12. Tsegaye S, Kebedew MG, Albrecht KK, Missimer TM, Thomas S, Elshall AS. Integrated GIS-hydrologic-hydraulic modeling to assess combined flood drivers in coastal regions: a case study of Bonita Bay, Florida. Frontiers in water. 2024;6:1468354. https://doi.org/10.3389/frwa.2024. 1468354
13. Chen Y. Flood hazard zone mapping incorporating geographic information system (GIS) and multi-criteria analysis (MCA) techniques. Journal of Hydrology. 2022;612:128268. https://doi.org/10.1016/j.jhydrol.2022.128268
14. Flores AB, Collins TW, Grineski SE, Amodeo M, Porter JR, Sampson CC, et al. Federally overlooked flood risk inequities in Houston, Texas: novel insights based on dasymetric mapping and state-of-the-art flood modeling. Annals of the American Association of Geographers. 2023;113(1):240-60. https://doi.org/10.1080/24694452. 2022.2085656
15. Kastridis A, Stathis D. Evaluation of hydrological and hydraulic models applied in typical Mediterranean Ungauged watersheds using post-flash-flood measurements. Hydrology. 2020;7(1):12. https://doi.org/10.3390/hydrology7010012
16. Mudashiru RB, Sabtu N, Abustan I, Balogun W. Flood hazard mapping methods: A review. Journal of hydrology. 2021;603:126846. https://doi.org/10.1016/j.jhydrol.2021.126846
17. Ahmed I, Pan ND, Debnath J, Bhowmik M, Bhattacharjee S. Flood hazard zonation using GIS-based multi-parametric Analytical Hierarchy Process. Geosystems and Geoenvironment. 2024;3(2):100250. https://doi.org/10.1016/j.geogeo.2023.100250
18. El-Bagoury H, Gad A. Integrated hydrological modeling for watershed analysis, flood prediction, and mitigation using meteorological and morphometric data, SCS-CN, HEC-HMS/RAS, and QGIS. Water. 2024;16(2):356. https://doi.org/10.3390/w16020356
19. Javidan N, Bahremand A. Separate components of the flood flow using the distributed hydrological WetSpa model in Ziarat-Gorgan watershed. Journal of Water and Soil Conservation. 2015;22(5):233-46.
20. El-Saoud WA, Othman A. An integrated hydrological and hydraulic modelling approach for flash flood hazard assessment in eastern Makkah city, Saudi Arabia. Journal of King Saud University-Science. 2022;34(4):102045. https://doi.org/10.1016/j.jksus.2022.102045
21. Cea L, Álvarez M, Puertas J. Using integrated hydrological–hydraulic modelling and global data sources to analyse the February 2023 floods in the Umbeluzi Catchment (Mozambique). Natural Hazards and Earth System Sciences. 2024;24(1):225-43. https://doi.org/10.5194/nhess-24-225-2024
22. Hasan HH, Mohd Razali SF, Ahmad Zaki AZI, Mohamad Hamzah F. Integrated hydrological-hydraulic model for flood simulation in tropical urban catchment. Sustainability. 2019;11(23):6700. https://doi.org/10.3390/su11236700
23. Venkata Rao G, Nagireddy NR, Keesara VR, Sridhar V, Srinivasan R, Umamahesh N, et al. Real-time flood forecasting using an integrated hydrologic and hydraulic model for the Vamsadhara and Nagavali basins, Eastern India. Natural Hazards. 2024;120(7):6011-39. https://doi.org/10.1007/s11069-023-06366-3
24. Anselmo V, Galeati G, Palmieri S, Rossi U, Todini E. Flood risk assessment using an integrated hydrological and hydraulic modelling approach: a case study. Journal of hydrology. 1996;175(1-4):533-54. https://doi.org/10.1016/S0022-1694(96)80023-0
25. Ruidas D, Pal SC, Saha A, Roy P, Pande CB, Islam ARMT, et al. Flood hazard forecasting and management systems: A review of state-of-the-art modelling, management strategies and policy-practice gap. International Journal of Disaster Risk Reduction. 2024;108:104539. https://doi.org/10.1016/j.ijdrr.2024.104539
26. Fernández-Nóvoa D, González-Cao J, García-Feal O. Enhancing flood risk management: a comprehensive review on flood early warning systems with emphasis on numerical modeling. Water. 2024;16(10):1408. https://doi.org/10.3390/w16101408
27. Borga M, Anagnostou E, Blöschl G, Creutin J-D. Flash flood forecasting, warning and risk management: the HYDRATE project. Environmental Science & Policy. 2011;14(7):834-44. https://doi.org/10.1016/j.envsci.2011.05.017
28. Kumar V, Sharma KV, Caloiero T, Mehta DJ, Singh K. Comprehensive overview of flood modeling approaches: A review of recent advances. Hydrology. 2023;10(7):141. https://doi.org/10.3390/hydrology10070141
29. Al-Omari AA, Shatnawi NN, Shbeeb NI, Istrati D, Lagaros ND, Abdalla KM. Utilizing remote sensing and GIS techniques for flood hazard mapping and risk assessment. Civil Engineering Journal. 2024;10(5):1423-36. https://doi.org/10.28991/CEJ-2024-010-05-05
30. Yanfatriani E, Marzuki M, Vonnisa M, Razi P, Hapsoro CA, Ramadhan R, et al. Extreme rainfall trends and hydrometeorological disasters in tropical regions: implications for climate resilience. Emerg Sci J. 2024;8(5):1860-74. https://doi.org/10.28991/ESJ-2024-08-05-012

31. Supratman M, Kusuma MS, Cahyono M, Kuntoro AA. Flood hazard assessment due to changes in land use and cover. Civil Engineering Journal. 2024;10(12):3874. https://doi.org/10.28991/CEJ-2024-010-12-04

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

هدف این مطالعه بررسی کاربرد رویکرد یکپارچه مدل‌سازی هیدرولوژیکی-هیدرولیکی برای شبیه‌سازی جریان‌های رودخانه‌ای، هیدروگراف‌های سیلاب و نواحی مستعد سیلاب در حوضه آبریز خرم‌آباد، ایران است. در این مطالعه، از مدل هیدرولوژیکی WetSpa و مدل هیدرولیکی LISFLOOD-FP برای شبیه‌سازی جریان‌های رودخانه‌ای و نواحی سیلابی استفاده شد. برای شبیه‌سازی‌ها از داده‌های هواشناسی ده‌ساله (سال‌های آبی ۲۰۰۵-۲۰۰۶ تا ۲۰۱۴-۲۰۱۵) استفاده گردید. مدل‌ها با استفاده از داده‌های پنج سال اول کالیبره شدند و داده‌های پنج سال بعدی برای اعتبارسنجی به کار رفت. مدل WetSpa برای شبیه‌سازی جریان‌های روزانه و مدل LISFLOOD-FP برای شبیه‌سازی نواحی مستعد سیلاب تحت دوره‌های برگشتی مختلف استفاده شد. مدل WetSpa توانست جریان‌های روزانه را با دقت قابل قبولی شبیه‌سازی کند و شاخص‌های عملکردی مانند کارآیی نش-ساتکلیف (NSE) معادل ۰.۵۳ و کارآیی کلینگ-گوپتا (KGE) معادل ۰.۶۸ را به‌دست آورد. برای رویدادهای جریان بالا، ضریب NSE به ۰.۷ رسید که نشان‌دهنده کارآیی مدل در پیش‌بینی سیلاب‌ها است. نتایج مدل LISFLOOD-FP نشان داد که حداکثر وسعت سیلاب برای دوره برگشتی ۲۰۰ ساله برابر با ۶۲۷,۱۰۰ متر مربع بوده است. تحلیل تعادل آب نشان داد که ۲۵.۵٪ از بارش‌ها به رواناب تبدیل شده است که شامل ۲.۳۴٪ رواناب سطحی، ۱.۶٪ رواناب زیرسطحی و ۲۱.۵۷٪ جریان آب‌های زیرزمینی می‌باشد. این مطالعه نشان می‌دهد که یکپارچگی مدل‌های هیدرولوژیکی و هیدرولیکی می‌تواند به‌طور مؤثری رفتار سیلاب‌ها را پیش‌بینی کرده و نواحی پرخطر را شبیه‌سازی کند.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1139-1151

International Journal of Engineering

Journal Homepage: www.ije.ir

Effect of Radio Frequency Sputtering and Laser Remelting on Ti6Al4V Behavior in Medical Applications

M. A. Ali Bash[a], A. N. Jasim[b], A. M. Resin*[a], S. A. Ajeel[a]

[a] *College of Production Engineering and Metallurgy, University of Technology, Baghdad 10066, Iraq*
[b] *Department of Materials Engineering, University of Diyala, Iraq*

PAPER INFO

Paper history:
Received 13 March 2025
Received in revised form 15 June 2025
Accepted 17 June 2025

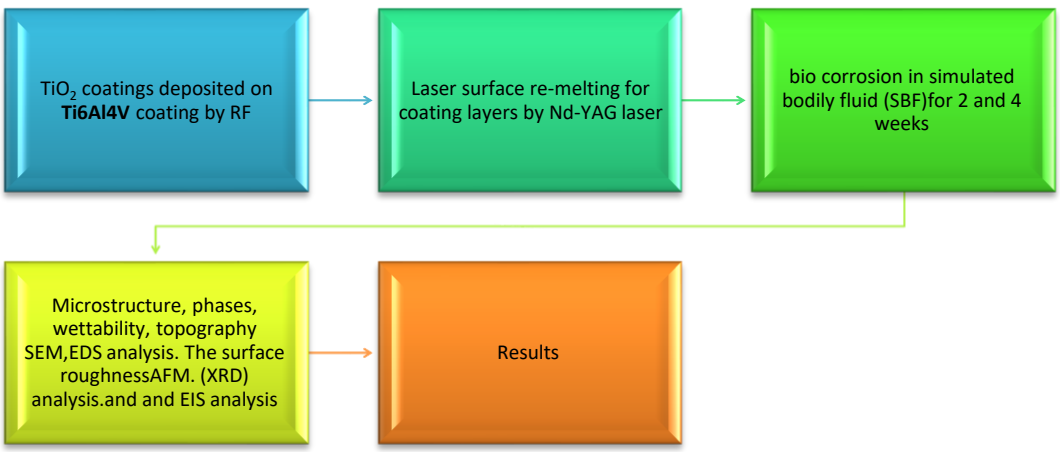
Keywords:
Biomaterial
Corrosion
Radio Frequency Sputtering
Ti6Al4V
Laser Treatment
Electrochemical Impedance Spectroscopy

ABSTRACT

In the field of orthopedics, the Ti6Al4V alloy is frequently used as a biomaterial, due to their characteristics such as high toughness, mechanical properties, high corrosion resistance, fracture toughness, biocompatibility and high strength. Yet, it has specific issues, such as corrosion and the release of vanadium element to the body fluids. Surface treatment methods promote osseointegration and biocompatibility by providing bioactivity and a barrier effect. In this study, the biological behavior of Sputtering titanium oxide (TiO_2) using radio frequency magnetron sputtering (RF) method as well as laser remelting of the pre-sputtered Ti6Al4V surface was investigated using Nd: YAG laser. Microstructure, phases, wettability, topography, composition, bioactivity of the modified surfaces were characterized using SEM, XRD, EDS, roughness, hardness, electrochemical corrosion, and Electrochemical Impedance Spectroscopy (EIS) analysis. The results demonstrate that the behavior of the substrate was positively impacted by the application of both surfaces engineering techniques, including the sputtering of ceramic coating and surface modification by laser remelting treatment. The improvement is owing to the formation of bioactive and more resistant surfaces. The modification of the surface performs two functions. The first serves as a barrier that prevents body fluids from reaching the implant, and it also provides a surface roughness and porosity similar to the roughness of the bone surface for the growth of surrounding tissue.

doi: *10.5829/ije.2026.39.05b.08*

Graphical Abstract



*Corresponding author email: ali.m.resen@uotechnology.edu.iq (A. M. Resin)

Please cite this article as: Ali Bash MA, Jasim AN, Resin AM, Ajeel SA. Effect of Acid Treatment and Clay Mineral Type on Acidic Clay Catalyst Performance in Esterification. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1139-51.

1. INTRODUCTION

Titanium and its alloys gained widespread use in the biomedical industry, where biological characteristics such as low elastic modulus, good fatigue strength, formability, and corrosion resistance were of crucial significance (1-3). This material is preferred because of its exceptional strength-to-weight ratio and remarkable biocompatibility, which means the body accepts it well and does not react unfavorably (4-6). Due to its ability to tolerate the mechanical stresses encountered in daily activities, Ti6Al4V is frequently utilized in orthopedic devices, dental implants, and joint replacements (7-9). The long-term clinical research confirms Ti6Al4V good biocompatibility. However, it raised concerns that the cytotoxic component vanadium might be released and cause both local and systemic issues (10, 11). Long amounts of time spending in close contact with semi-planted cells might be dangerous (12). Biomedical applications must cover innovative materials to increase the functionality and longevity of medical devices. Recently, surface modification techniques have attracted attention because they may speed up or improve "bone-anchorage" by creating a bio-interface between the implant and the surrounding bone tissue (13-15). So, the implant's surface characteristics are critical for osteoblast adhesion and growth during the early stages of osseointegration as well as long-term bone remodeling (16, 17). In this instance, the methods that primarily utilized involve adjusting the implant's roughness and texture. Chemical and structural changes to the composition of the surface that make touch with the bone to facilitate a direct chemical connection, strengthening other physical elements like bone tissue guidance, biomechanical stability, and bone anchoring (18, 19).

Thereby, a variety of techniques have been employed to locally modify the implants' surface energy, biochemistry, and topography (20). Simultaneously, efforts have been taken to enhance the mechanical characteristics of implants, with particular emphasis on improving the hardness as well as improving the wear resistance of metallic implants (21). Radio frequency magnetron sputtering is a more attractive plasma/vacuum-based deposition technique in this regard. One of the most prevalent plasma technologies in industry for coating and surface modification is magnetron sputtering, which produces highly conformal coatings which adhere well to a wide range of substrates (22-24). Titanium oxide (TiO_2) is a well-known coating material because of its exceptional biocompatibility, or ability to live with living tissue without having an adverse impact (25, 26). For materials that come into direct contact with biological systems, this feature is essential. Another novel technique for modifying the surface features of Ti6Al4V is laser surface treatment (27-30). The surface can be melted and then resolidified with lasers, producing microstructures that improve the material's characteristics (31). This procedure may prolong the implant's lifespan overall, increases surface hardness, and reduce wear. Furthermore, by producing textured surfaces that simulate real bone, laser treatment can improve the osseointegration, which is the process by which bone cells attach to the implant surface (32). Improved osseointegration is vital for the success of implants. By tailoring the surface characteristics of Ti6Al4V, engineers and scientists can develop implants that not only perform better but also lead to enhance patient outcomes (33).

In this investigation, studying the impact of using combination of advanced surface treatment techniques like RF sputtering and laser treatment on Ti6Al4V surface were focused. Through a series of experiments and tests, the microstructural, phases, roughness, and biocorrosion resistance of the treated specimens were evaluated compared to the untreated standard.

2. EXPERIMENTAL

2. 1. Coating of the Ti6Al4Vsubstrate The experimental material used as substrate in this investigation is a plate of Ti6Al4V alloy 3 mm thick. Its chemical composition is demonstrated at Table 1. Ceramic coating is TiO_2 nano-coating using RF sputtering. Before RF sputtering, the surface of substrate materials was ground with 400-grit waterproof abrasive paper to increase the bonding between the substrate and ceramic coating. The sputtering chamber was evacuated to a working pressure of 5×10^{-3} mbar. The deposition of TiO_2 film was conducted using the RF sputtering system in a pure argon atmosphere (99.9%) at a pressure of 5×10^{-3} mbar for a duration of 90 minutes.

2. 2. Laser Surface Re-melting of Ti6Al4V Pre-sputtered with TiO_2 A Nd: YAG laser was applied to process the surface of Ti6Al4V pre -sputtered with TiO_2. The processing parameters of the laser surface treatment were applied by frequency: 2 Hz, power: 650W, and traverse speed: 8 mm/s.

TABLE 1. The chemical composition of the standard Ti6Al4V alloy

Chemical element	Al	V	Fe	O	N	H	C	Ti
Standard (ASTM)	5.5/6.75	3.5/4.5	0.30	0.20	0.05	0.01	0.10	Balance

2. 3. Wettability Measurements The contact angle measurement method used to determine the wettability of the surface by the droplet suspension method with a contact angle tester type (Creating Nano Tech., Model: CAM110P, Serial No. 113031201804W). Less than a 90° angle of water contact corresponds to a hydrophilic surface, which is thought to promote the nutrients and cell attachments associated with bone formation.

2. 4. Surface Roughness Roughness can be described as the irregularity in surface profile that includes valleys and peaks, and it provides a significant impact on cellular responses (i.e., proliferation, differentiations, and adhesion) for implanting the surface. It has been evaluated using a Surface Roughness Tester type (JITAI8103Plus Surface Roughness Tester). The range was -80/80µm and the cutoff: 2.5mm. consists mostly of the contour's most significant height (R_z), its arithmetic mean deviation (R_a), and its root-mean-square deviation of the contour mean line (R_q).

2. 5. Analysis Methods FESEM was utilized to characterize the surface of Ti6Al4V, after RF sputtering the surface with TiO_2 and, after laser remelting of pre-sputtered Ti6Al4V. The compositions were analyzed with EDS analysis. As well as X-ray diffractometer (XRD) analysis used to identify the phases existing on the treated surfaces.

2. 6. Biocorrosion Behavior and EIS Measurements To investigate the behavior and corrosion rate of specimens in the simulated bodily fluid (SBF) for 2 and 4 weeks, electrochemical evaluations of the uncoated Ti6Al4V substrate, TiO_2-sputtered, and laser remelting of the pre-sputtered surface have been carried out. The corrosion rate was measured in millimeters per year (mm/y) using a CHI 604e electrochemical system. The AC impedance spectra for Ti6Al4V samples in simulated body fluids (SBF) were obtained. This analysis serves two purposes: first, it provides a diagnostic assessment of the surface's characteristics; second, it aids in the assessment of the enhancements that perform to overcome corrosive circumstances with frequencies up to 1 MHz.

3. RESULTS AND DISCUSSION

3. 1. Microstructural Features of the Coatings Figure 1 depicts typical FESEM micrograph images of Ti6Al4V after RF sputtering with TiO_2 and laser surface remelting. The ceramic TiO_2 films deposited on the surface of Ti6Al4V seemed thin since the grinding grooves of the substrate were still visible as shown in Figure 1(a, b). The grinding process of the substrate exhibited parallel grooves in the grinding direction to promotes the mechanical anchoring of coatings. Figure 1 (c, d) shows the laser surface remelting of pre-sputtered

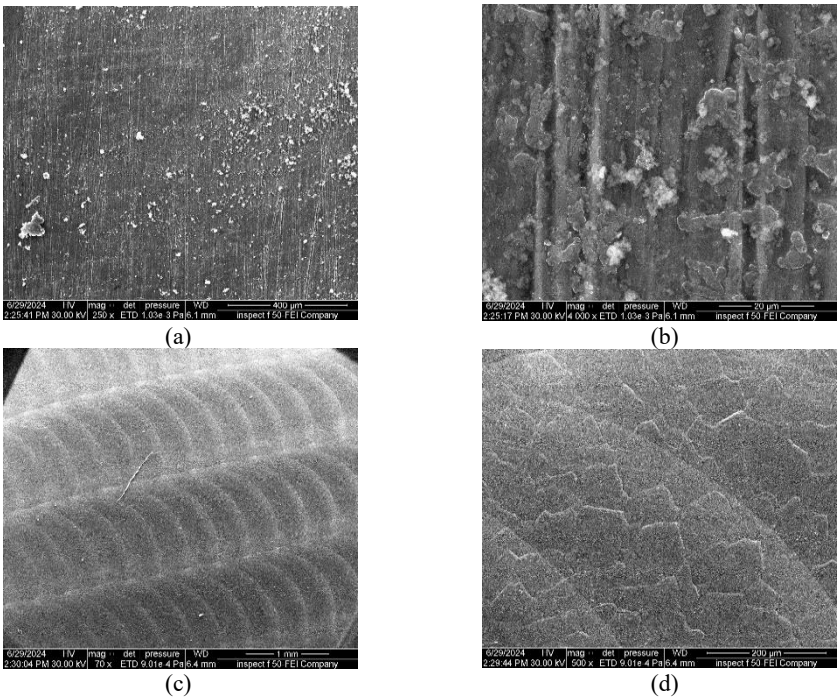


Figure 1. Plan view of Ti6Al4V after (a, and b) RF sputtering of TiO_2 on the ground of Ti6Al4V surface at different magnifications (c, and d) laser remelting of the pre-sputtered Ti6Al4V surface

Ti6Al4V. It can observe the uniform laser tracks overlap according to the investigated parameters used in this paper. Also, the martensite phase of the Ti alloy can be clearly observed at high magnification of Figure 1(d). The martensite phase was performed due to the high cooling rate that occurred after laser treatment.

The interesting think is the presence of ceramic coating on the surface of metallic substrate. It is playing a two crucial role. The first one is in protecting the metallic surface of the implants. Because the ceramic nano coating considered more resistant to the surrounding environment from the mechanical and biological point. The second role in increasing the surface response of Ti6Al4V to the laser beam interaction. As it is known the response of metallic surface to the laser beam lesser than the ceramic one. In other words, the laser beam absorption in the case of ceramic coating is high. Therefore, control over surface details is better.

Figure 2 displays the XRD spectra of substrate Ti-6Al-4V and modified specimens. Ti6Al4V alloy before coating showed three different 2θ peaks are related for Ti6Al4V [32.76θ (100), 35.24θ (002) and 40.13 θ (101)]. They identified according to JCPDS card No 33-0397. Also, it is indicated that this structure includes (α+β), while the α phase is representing the α stabilizing element as Al element has the tendency of stabilizing the α phase, whereas (V) element is acting as β phase stabilizer.

There are no characteristic peaks of TiO_2 after RF sputtering of Ti-6Al-4V with TiO_2. It disappears due to Titania which is a thin layer and it could not find by XRD analysis. But it can be characterized using GXRD analysis of sputtering TiO_2 on the surface of glass as shown in Figure 3. It's important to note, after laser treatment, it can obtain martensitic phase (α') as a result of high cooling rate. Such high cooling rates force formation of acicular martensite structure and enhances the materials wear and corrosion resistance (34). The lattice plane spacing in the constituent grains changed from a stress-free value to a new value that matched the amount of stress applied when Ti6A14V was subjected to laser treatment, which caused uniform strain to be generated over comparatively wide distances (Table 2).

The crystallinity of phases increases after laser treatment due to the high temperature of laser remelting of Ti6Al4V sputtered with TiO_2. There are no characteristic peaks of TiO_2 after laser remelting of pre-sputtering of Ti-6Al-4V with TiO_2. Table 2 lists the crystalline size for Ti6Al4V before and after modification.

3. 2. Surface Wettability of Titanium Alloy The biological system is significantly affected by the wettability of implant surfaces, including protein adhesion, surface interactions with hard and soft tissue cells, bacterial affinity, and biofilm formation. Figure 4 depicts the surface samples' wetting behavior. As previously noted, the contact angle of water droplets on the specimen surface was used to assess the wettability of the Ti6Al4V alloy under all conditions. Prior to coating, the ground substrate Ti6Al4V alloy had a contact angle value of 70.415. The value of the contact angle increased to 101.637 upon the application of TiO_2 coating via RF sputtering on the Ti6Al4V surface. This indicates that, in comparison to the surface without coating, the Ti alloy surface became hydrophobic.

However, the pre-sputtered Ti6Al4V surface's contact angle value dropped to 79.538 during laser remelting. Figure 4 shows that after laser remelting, the surface wettability of the titanium alloy increased and the CA value dropped. This may be explained by the fact that the water droplets are stored in the groove structure at a specific depth as a result of laser track overlapping, which facilitates their spreading across the titanium alloy surface and results in a reduced contact angle (33).

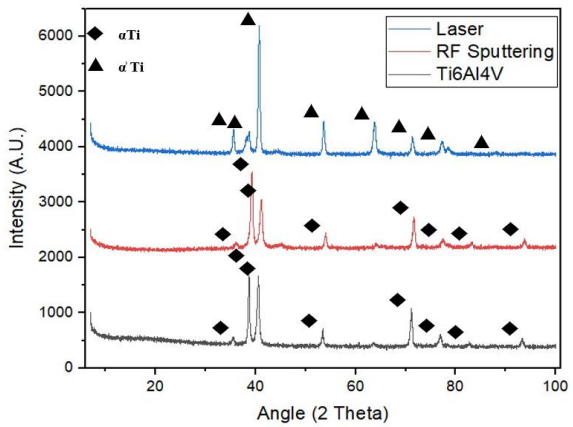


Figure 2. The XRD analysis of the substrate (Ti6Al4V), RF sputtering of TiO_2, and laser remelting of pre-sputtered substrate

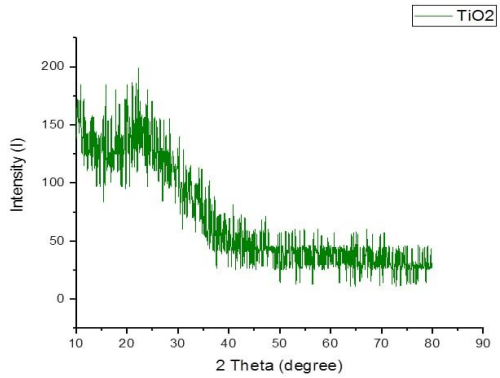


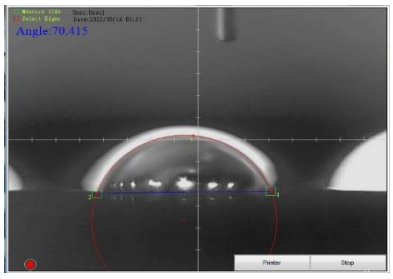
Figure 3. GXRD curve of sputtering TiO_2 on the surface of glass

Surface wettability plays a crucial role during the insertion of an implant into the body. Additionally, the process governs the adsorption of the proteins, which is followed by cell attachments. It appears that on hydrophilic surfaces as opposed to hydrophobic ones, the cells will be attacked and grow (33).

Surface roughness was obtained for the specimens of Ti6Al4V alloy and after RF sputtering with TiO_2 and laser surface remelting as shown in Table 3 and Figure 5. It shows the surface roughness profiles of the specimens. It can observe that the condition with ceramic nano coating using the magnetron RF sputtering reduces surface roughness of the Ti6Al4V alloy to (0.344 µm) and improved integrity compared to the substrate condition after grinding with abrasive paper (1.522 µm). This attributed to the filling the groves that generated due to the grinding paper with the ceramic coating materials (TiO_2) and that agree with reported data (35). On the other hand, the surface roughness increases after laser surface remelting of pre-sputtered Ti6Al4V to (4.581 µm). Because, the surface texture that generated after overlapping of laser track results in the increasing of roughness values. As well as, the laser processing parameters used in this investigation by Naji et al. (33). It is important to note that increasing the surface roughness to some extent was preferred for some of the biological tests, such as bioactivity, that improve osseointegration.

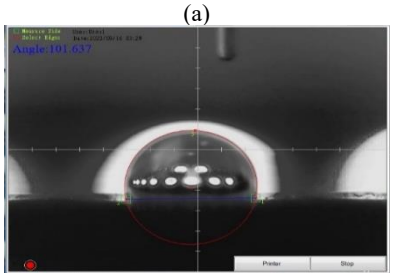
TABLE 2. The FWHM and crystalline size of the three specimens

ITEM	FWHM	Area	Crystallite size(A°)	Micro Strain [%]
substrate (Ti6Al4V)	0.2642	413.09	296	0.39288
After RF sputtering of Ti6Al4V	0.4607	661.22	152	0.75423
After Laser surface treatment of pre- sputtered Ti6Al4V	0.6648	135.64	101	1.16361



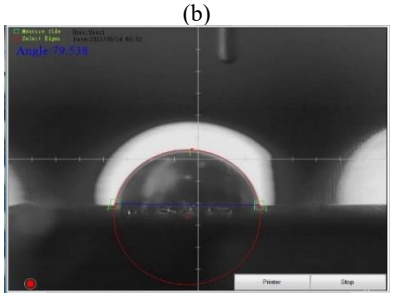
70.415

(a)



101.637

(b)



79.538

(c)

Figure 4. The water contact angle measurements of (a) Ground substrate (Ti6Al4V) (b) After RF sputtering of Ti6Al4V (c) After Laser surface treatment of pre- sputtered Ti6Al4V

3. 3. Microstructure after Bioactivity

The surface bioactivity is an aspect of high importance for ensuring subsequent bone implant integration in terms of in vivo conditions, thus three distinctive samples (Ti6Al4V alloy surface, RF sputtering of TiO_2 on the surface of Ti6Al4V, and laser surface remelting of pre-sputtered Ti6Al4V with TiO_2) were specified via immersion in the SBF for 2 and 4 weeks. The (FE-SEM)

TABLE 3. Surface roughness for Ti6Al4V samples

Samples	R_a (µm)	R_z (µm)	R_q (µm)
Ground substrate (Ti6Al4V)	1.522	9.96	1.978
After RF sputtering of Ti6Al4V	0.344	2.796	0.448
After Laser surface treatment of pre- sputtered Ti6Al4V	4.581	17.596	5.258

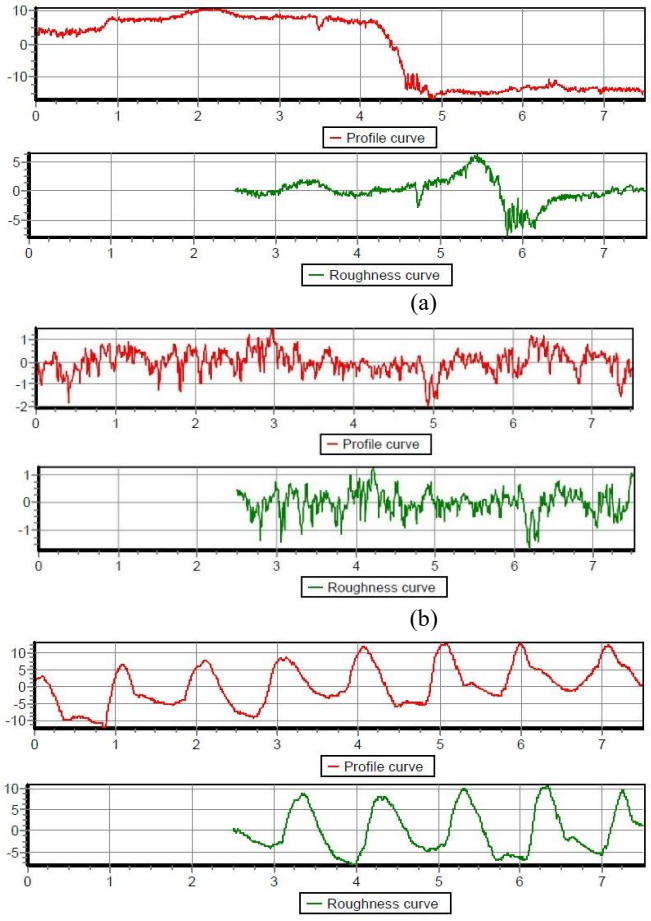


Figure 5. Surface roughness profilometer of titanium alloy samples at the condition: (a) Ground substrate (Ti6Al4V) (b) After RF sputtering of Ti6Al4V (c) After Laser surface treatment of pre- sputtered Ti6Al4V

analysis visually assesses the morphological features of all cases investigated in this study. The FE-SEM micrograph clearly shows the apatite formation on the surfaces of the samples. Figure 6(a) indicates the surface of titanium alloy Ti6Al4V covered with a thin layer of hydroxyapatite with random distribution. while Figure 6(b) shows that almost the TiO_2 nano-coating surface had been covered with apatite in semi-equal distribution. The surface will be denser and thicker because of an increase in the nucleation of the apatite, indicates a creating a layer of the apatite on the surface related to mineralized bio ceramic material types.

Also, on the surface of the laser- treated of pre-sputtered samples in Figure 6(c), there is a homogeneity in apatite formation and some aggregation highlight the success of the synthesis process in achieving a uniform distribution. These results suggest that bone formation depends on the type of materials and surface nature. It is significant that after laser surface remelting, the apatite formation was better than that of both Ti6Al4V surface and sputtered surface with TiO_2. The existence of overlapping regions after laser surface treatment promotes the growth, density, and cell survival rate. Moreover, Figure 7 at high magnification shows that all region on the surface was covered with bone structure. Which means the surfaces after laser treatment have high activation energy and nucleation sites for bone formation.

This indicates that osteointegration may response differently to different surface structure. In addition, the surface energy and surface roughness would be essential factors for facilitating the cells adhesion and proliferation. It's important to note this study deals with laser-treated surfaces after ceramic coating which means responding and absorbing the laser beam with high efficiency.

Figure 8 shows the microscopic morphology and the corresponding chemical composition obtained from EDS analysis on the titanium alloy after laser treatment. The SEM-elemental mapping in Figure 8 illustrates the existence of bone elements Ca, P, and O. Highlighting

the successful osteointegration process on the surface of the sample. Also, the XRD pattern confirms the formation of apatite on the surface of all specimens in Figure 9. The HA peaks were identified according to JCPDS card No 96-230-0274.

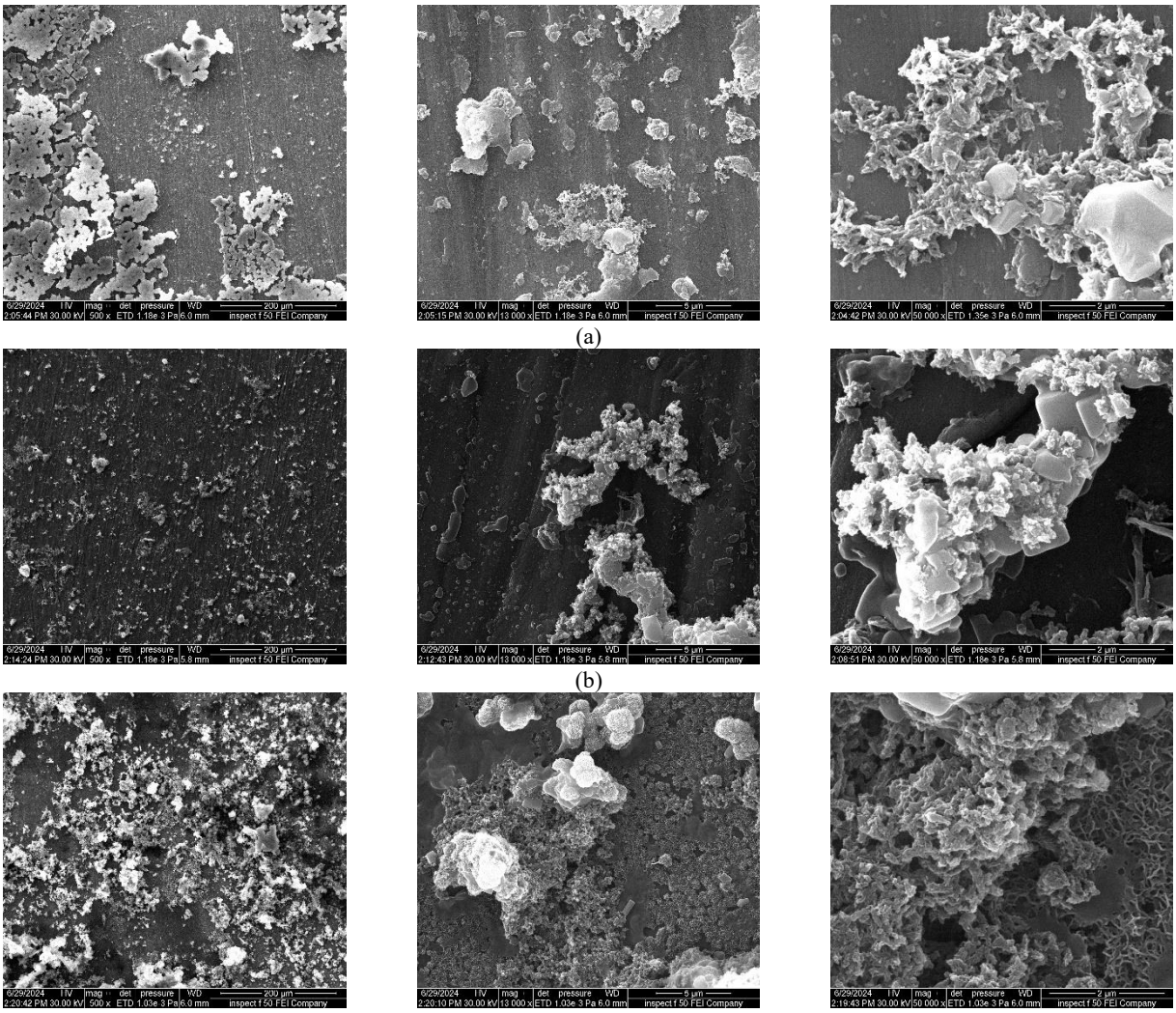


Figure 6. FE-SEM images of bone formation on three different samples for 7 days (a) Ti6Al4V substrate (b) RF sputtered surface of Ti6Al4V with TiO_2 (c) Laser surface remelting of pre-sputtered Ti6Al4V with TiO_2

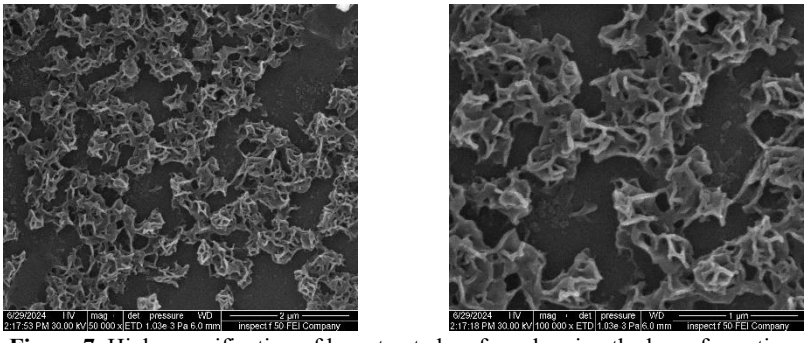


Figure 7. High magnification of laser treated surface showing the bone formation

3. 4. EDS Analysis

3. 5. Tafel Results Discussion

To evaluate the effect of TiO_2 sputtering and laser treatment of pre-sputtered surface on the corrosion resistance of the titanium alloy (Ti6Al4V), the electrochemical corrosion performance of the samples in the SBF solution was tested. Also, the electrochemical corrosion performance was conducted for all samples after immersion in SBF for 4 weeks to simulate the investigated samples in human body as implants. The results are shown in Figure 10. The open circuit potential (OCP), corrosion current density (I_{corr}), corrosion potential (E_{corr}), and inhibition efficiency were calculated and summarized in Table 4. The values of OCP were nobler with (-0.128 and -0.241) volt when sputtered the Ti6Al4V with TiO_2 and surface laser remelting of pre-sputtered surface if compared with the OCP value for untreated Ti6Al4V (-0.515) volt.

The Tafel curve in Figure 10(a) demonstrates the improvement in corrosion behavior after using both two methods of surface engineering for Ti6Al4V. A reduced current density and a larger corrosion potential confirmed this in Table ٤. So, the values of corrosion rate decreased from (3.255×10^{-2}) to (3.892×10^{-3}) and (1.079×10^{-2}) mm/y. Due to the protection that comes from the surface treatment, which creates a noble transition compared to the untreated substrate. Moreover, these methods exhibited a higher inhibition efficiency 88.04% and 66.85% respectively. This enhancement is attributed to the application of a ceramic layer from titania which considered excellent corrosion resistance. It's important to note the improvement after sputtering is higher than the improvement following laser remelting treatment of pre- sputtered surface. The reason for this is that overlapping occurs during laser treatment, as shown in

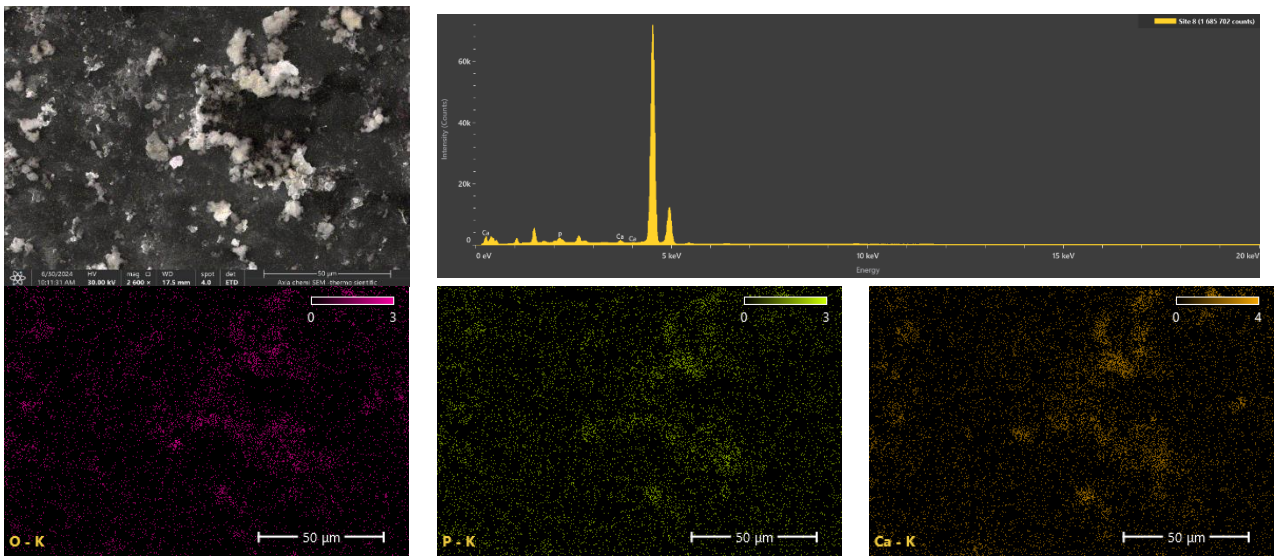


Figure 8. EDS analysis and SEM-elemental mapping after immersion in SBF solution of the laser remelted pre-sputtered Ti6Al4V exhibit the existence of HA on the surface

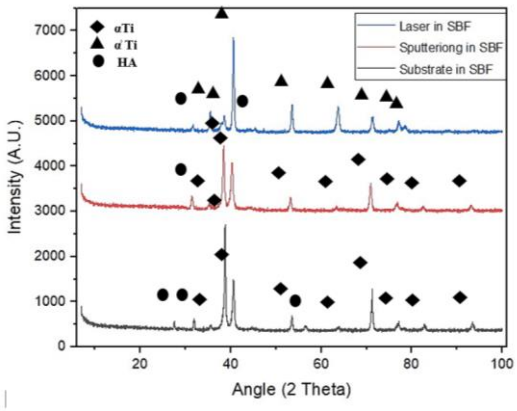
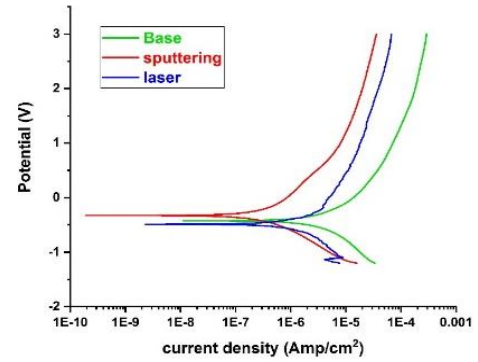


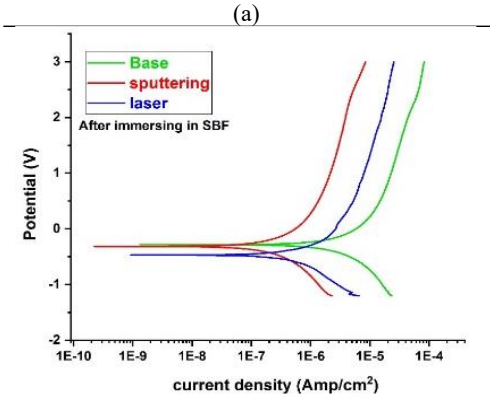
Figure 9. XRD analysis of the three cases after immersion in SBF solution for 2 weeks

TABLE 4. Electrochemical corrosion results

ITEM	E corr. (volt)	I corr. (Amp.)	Corr. Rate (mmpy)	βc	βa	OCP (volt)	η%
Base	-0.429	1.440×10^{-6}	3.255×10^{-2}	0.206	0.204	-0.515	---
Base in SBF	-0.278	8.659×10^{-7}	1.957×10^{-2}	0.212	0.204	0.290	39.87%
Sputtering	-0.319	1.722×10^{-7}	3.892×10^{-3}	0.188	0.206	-0.128	88.04%
Sputtering in SBF	-0.323	6.540×10^{-8}	1.478×10^{-3}	0.206	0.185	0.119	95.45%
Laser after sputtering	-0.493	4.774×10^{-7}	1.079×10^{-2}	0.212	0.210	-0.241	66.85%
Laser after sputtering in SBF	-0.490	2.131×10^{-7}	4.817×10^{-3}	0.206	0.184	0.162	85.20%



(a)



(b)

Figure 10. Anticorrosion measurements on the Ti6Al4V alloy of the three different samples (a) before and (b) after immersion in SBF solution for 4 weeks

Figure 1, resulting in an increase in surface roughness in the overlapping area and subsequently an increase in the surface area exposed to corrosion solution. Furthermore, the generation of residual stresses as well as the small grain sizes results due to rapid solidification rate after laser treatment suggested the cause of a reduction in corrosion resistance slightly compared to the case of sputtering.

While the OCP, Tafel curve results for the three investigated specimens after immersion in SBF solution for 4 weeks demonstrate more enhancement in corrosion behavior as shown in Figure 10(b). The OCP values was 0.119 and 0.162 compared to the value of the substrate 0.290. The corrosion rate also decreased after immersion in SBF solution to (1.478×10^{-3}) and (4.817×10^{-3}), respectively compared to the substrate corrosion rate after immersion in SBF (1.957×10^{-2}) mm/y., which attributed to the formation of HA film of the surface of the samples play as a barrier against direct contact to the solution. Which means the reduction of corrosion of implants due to the osseointegration of bones. The inhabitation efficiency also increases to 95.45% and 85.20% as well as for the substrate become 39.87% after immersion in SBF solution compared to the values before immersion. These results reflect the performance of implants with the progress of time.

3. 6. EIS Analysis EIS is a method that is especially sensitive to minor system changes and facilitates for the characterization of material properties and electrochemical process mechanisms. The Nyquist graphs (see Figure 11 and Table 5) demonstrate the electrochemical corrosion performance of the samples in the SBF solution to find out the impact of TiO_2 sputtering and laser treatment of the pre-sputtered surface on the titanium alloy's (Ti6Al4V) corrosion resistance (Figure 11 (a, c, e)). In order to replicate the examined materials in the human body as implants, an electrochemical corrosion performance assessment was additionally carried out on all samples following a 4-week immersion in SBF (Figure 11(b, d, f)).

It is obvious from the diagrams that the base metal's (Ti6Al4V) impedance response was 9858 Ω (Figure 11(a)). The solution resistant is denoted by R_1 = 207 Ω, and the double layer resistant by R_2 = 9651 Ω. While after sputtering of nano ceramic layer type titania oxide, the value of the impedance became 255733 Ω (Figure 11(c)). The solution resistant is denoted by R_1 = 217 Ω, the TiO_2 ceramic layer resistant is denoted by R_2= 242943Ω. High adhesion, a strong atomic bond, and a low pore rate are characteristics of this layer. R_3= 12573Ω for the double layer resistant. As long as R_2 is linked to R_3, this means that there are pores in the layer, but within an acceptable limit. After laser remelting of pre-sputtered with TiO_2 of the base metal, the value of impedance response was56741 Ω (Figure 11(e)). The solution resistant is denoted by R_1 = 200 Ω, and there were two separated

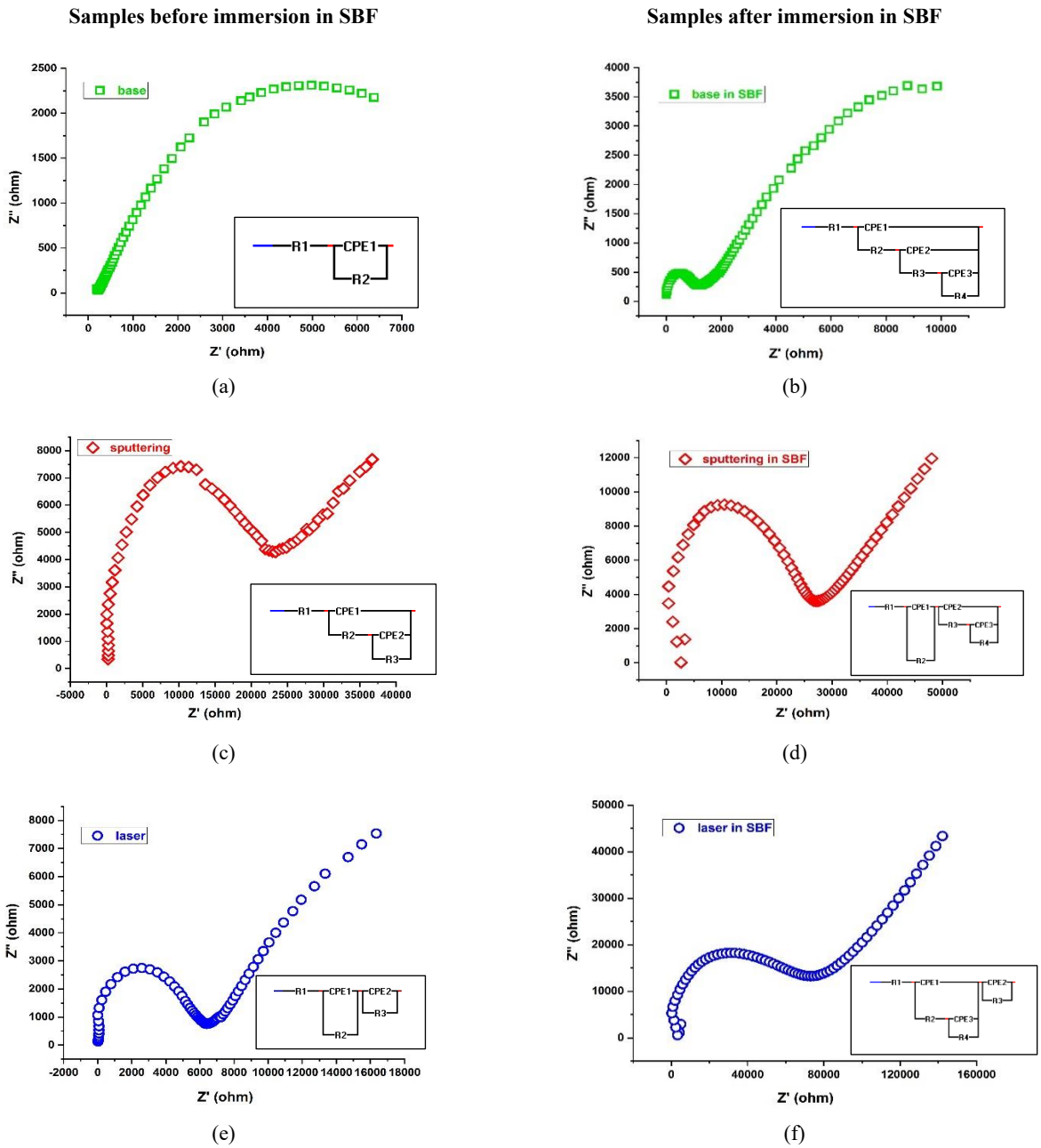


Figure 11. Nyquist plot of Ti6Al4V implant alloy in SBF solution (a) Base metal (Ti6Al4V) (b) Base metal in SBF (c) After sputtering of TiO_2 (d) After sputtering of TiO_2 in SBF (e) Laser after sputtering (f) Laser after sputtering in SBF

TABLE 5. EIS analysis for the investigated samples

Samples type	Total Impedance(ohm)	R1 (ohm)	R2(ohm)	R3(ohm)	R4(ohm)
Base	9858	207	9651	----	-----
Base in SBF	40062	232	34816	3576	1438
Sputtering	255733	217	242943	12573	----
Sputtering in SBF	1.172472e+6	228	1E06	1.639E-05	8334
Laser after sputtering	56741	200	50611	5930	----
Laser after sputtering in SBF	158263	214	1.0012E05	4.9012E04	8917

layers formed R_2=50611Ω and R_3= 5930 Ω (the double layer resistant). The R_2 corresponds to the effect of laser beam on the pre-sputtered surface.

The impedance spectra for Ti6Al4V implant alloys after immersion for 4 weeks in SBF solution (after apatite formation) are comprised of three semicircle loops (Figure 11 (b)). The solution resistant is denoted by R_1 = 232 Ω, the HA layer resistant by R_2 =34816 Ω, the titanium oxide film formed as a result of sintering the HA layer in the furnace at 400°C after immersion in SBF solution by R_3= 3576 Ω, and the double layer resistant by R_4 = 1438 Ω. These findings show that the apatite layer (HA), which is thought to act as a barrier that improves the metal's resistance to corrosion, emerges as time passes, causing the rate of corrosion in the (SBF) solution to decrease. When compared to the resistance value prior to immersion, the titanium alloy's resistance rose to (40062 Ω) after immersion. The EIS results for the titanium alloy's sputtered surface with titanium revealed varying levels of corrosion resistance (see Figure 11 (d)). The overall impedance increases to 1.172472e+6 Ω, indicating that the surface of Ti6Al4V is difficult for the SBF solution to attack. The apatite layer that developed on top of the titanium layer (R_2=1E06), which has a high density and is strongly bound to the surface, was primarily responsible for the high resistance. The impedance value after laser treatment of the surface sputtered with titanium after 4 weeks of immersion in the solution was 158263Ω Figure 11(f). Likewise, the resistance resulting from the apatite layer (R_2=1.0012E05Ω) is greater than the resistance to the laser effect on the material (R_3= 4.9012E04 Ω). We conclude that in all the mentioned cases, the resistance values increase over time as a result of the compatibility and osseointegration provided by the mentioned materials.

4. CONCLUSIONS

1. A uniform remelting surface can be achieved using Nd- YAG laser after sputtering of the ceramic layer from TiO_2 on the surface of the base metal (Ti6Al4V). The laser – Ti6Al4V interaction increased after sputtering of ceramic layer.
2. XRD results reveal the formation of martensite phase α' after laser treatment due to the rapid solidification instead of α phase.
3. Biocompatibility increases after both sputtering and laser treatment of the Ti6Al4V surface.
4. For certain biological tests, as bioactivity, which seemed to indicate a method of improved osseointegration, it was preferable to increase the surface roughness to a certain degree. For laser surface remelting of pre-sputtered Ti6Al4V, the roughness was 4.581μm.
5. EIS technique reflects the electrochemical corrosion of the investigated specimens. RF sputtering of TiO_2 and laser surface remelting increase the values of impedance to the high level. The highest one was after sputtering of ceramic layer on the surface of Ti6Al4V.

5. REFERENCES

1. Wiatrowski A, Mazur M, Obstarczyk A, Kaczmarek D, Pastuszek R, Wojcieszak D, et al. Influence of magnetron powering mode on various properties of TiO2 thin films. Mater Sci Pol. 2018;36:748-60. 10.2478/msp-2018-0099
2. Sobieszczyk S. Surface modifications of Ti and its alloys. Advances in Materials Science. 2010;10:29-42. 10.2478/v10077-010-0003-3
3. Souza JC, Apaza-Bedoya K, Benfatti CA, Silva FS, Henriques B. A comprehensive review on the corrosion pathways of titanium dental implants and their biological adverse effects. Metals. 2020;10(9):1272. 10.3390/met10091272
4. Xue T, Attarilar S, Liu S, Liu J, Song X, Li L, et al. Surface modification techniques of titanium and its alloys to functionally optimize their biomedical properties: thematic review. Frontiers in bioengineering and biotechnology. 2020;8:603072. 10.3389/fbioe.2020.603072
5. Avcu E, Avcu YY, Baştan FE, Rehman MAU, Üstel F, Boccaccini AR. Tailoring the surface characteristics of electrophoretically deposited chitosan-based bioactive glass composite coatings on titanium implants via grit blasting. Progress in organic coatings. 2018;123:362-73. 10.1016/j.porgcoat.2018.07.021
6. Hanawa T. Biocompatibility of titanium from the viewpoint of its surface. Science and Technology of Advanced Materials. 2022;23(1):457-72. 10.1080/14686996.2022.2106156
7. Abd-Elaziem W, Darwish MA, Hamada A, Daoush WM. Titanium-Based alloys and composites for orthopedic implants Applications: A comprehensive review. Materials & Design. 2024;241:112850. 10.1016/j.matdes.2024.112850
8. Quinn J, McFadden R, Chan C-W, Carson L. Titanium for orthopedic applications: an overview of surface modification to improve biocompatibility and prevent bacterial biofilm formation. IScience. 2020;23(11). 10.1016/j.isci.2020.101745
9. Xu Y, Gao J, Huang Y, Rainforth WM. A low-cost metastable beta Ti alloy with high elastic admissible strain and enhanced ductility for orthopaedic application. Journal of Alloys and Compounds. 2020;835:155391. 10.1016/j.jallcom.2020.155391
10. Bandyopadhyay A, Mitra I, Goodman SB, Kumar M, Bose S. Improving biocompatibility for next generation of metallic implants. Progress in materials science. 2023;133:101053. 10.1016/j.pmatsci.2022.101053
11. Willis J, Li S, Crean SJ, Barrak FN. Is titanium alloy Ti-6Al-4 V cytotoxic to gingival fibroblasts—A systematic review. Clinical and experimental dental research. 2021;7(6):1037-44. 10.1002/cre2.444
12. Ektessabi A, Otsuka T, Tsuboi Y, Yokoyama K, Albrektsson T, Sennerby L, et al. Application of micro beam PIXE to detection of titanium ion release from dental and orthopaedic implants. International Journal of PIXE. 1994;4(02n03):81-91. 10.1142/S0129083594000118
13. Tuikampee S, Chaijareenont P, Rungsiyakull P, Yavirach A. Titanium surface modification techniques to enhance osteoblasts

and bone formation for dental implants: a narrative review on current advances. Metals. 2024;14(5):515. 2075-4701/14/5/515

14. Kligman S, Ren Z, Chung C-H, Perillo MA, Chang Y-C, Koo H, et al. The impact of dental implant surface modifications on osseointegration and biofilm formation. Journal of clinical medicine. 2021;10(8):1641. 10.3390/jcm10081641
15. Li J, Zhou P, Attarilar S, Shi H. Innovative surface modification procedures to achieve micro/nano-graded Ti-based biomedical alloys and implants. Coatings. 2021;11(6):647. 10.3390/coatings11060647
16. Al-Zubaidi SM, Madfa AA, Mufadhal AA, Aldawla MA, Hameed OS, Yue X-G. Improvements in clinical durability from functional biomimetic metallic dental implants. Frontiers in Materials. 2020;7:106. 10.3389/fmats.2020.00106
17. Zhang LC, Chen LY. A review on biomedical titanium alloys: recent progress and prospect. Advanced engineering materials. 2019;21(4):1801215. 10.1002/adem.201801215
18. Sánchez-López JC, Rodríguez-Albelo M, Sánchez-Pérez M, Godinho V, López-Santos C, Torres Y. Ti6Al4V coatings on titanium samples by sputtering techniques: Microstructural and mechanical characterization. Journal of Alloys and Compounds. 2023;952:170018. 10.1016/j.jallcom.2023.170018
19. Liu X, Chu PK, Ding C. Surface modification of titanium, titanium alloys, and related materials for biomedical applications. Materials Science and Engineering: R: Reports. 2004;47(3-4):49-121. 10.1016/j.mser.2004.11.001
20. García-Cabezón C, Godinho V, Salvo-Comino C, Torres Y, Martín-Pedrosa F. Improved corrosion behavior and biocompatibility of porous titanium samples coated with bioactive chitosan-based nanocomposites. Materials. 2021;14(21):6322. 10.3390/ma14216322
21. Frutos E, Karlik M, Polcar T. The role of α ″orthorhombic phase content on the tenacity and fracture toughness behavior of Ti-22Nb-10Zr coating used in the design of long-term medical implants. Applied Surface Science. 2019;464:328-36. 10.1016/j.apsusc.2018.09.017
22. Prosolov KA, Lastovka VV, Khimich MA, Chebodaeva VV, Khlusov IA, Sharkeev YP. RF magnetron sputtering of substituted hydroxyapatite for deposition of biocoatings. Materials. 2022;15(19):6828. 10.3390/ma15196828
23. Garg R, Gonuguntla S, Sk S, Iqbal MS, Dada AO, Pal U, et al. Sputtering thin films: Materials, applications, challenges and future directions. Advances in colloid and interface science. 2024;330:103203. 10.1016/j.cis.2024.103203
24. Jafari S, Mahyad B, Hashemzadeh H, Janfaza S, Gholikhani T, Tayebi L. Biomedical applications of TiO2 nanostructures: recent advances. International journal of nanomedicine. 2020:3447-70. 10.2147/IJN.S249441
25. Kumaravel V, Nair KM, Mathew S, Bartlett J, Kennedy JE, Manning HG, et al. Antimicrobial TiO2 nanocomposite coatings for surfaces, dental and orthopaedic implants. Chemical Engineering Journal. 2021;416:129071. 10.1016/j.cej.2021.129071
26. Geng Y, McCarthy É, Brabazon D, Harrison N. Ti6Al4V functionally graded material via high power and high speed laser surface modification. Surface and Coatings Technology. 2020;398:126085. 10.1016/j.surfcoat.2020.126085
27. Kolarovszki B, Ficsor S, Frank D, Katona K, Soos B, Turzo K. Unlocking the potential: laser surface modifications for titanium dental implants. Lasers in Medical Science. 2024;39(1):162. 10.1007/s10103-024-04076-1
28. Sypniewska J, Szkodo M. Influence of laser modification on the surface character of biomaterials: titanium and its alloys—a review. Coatings. 2022;12(10):1371. 10.3390/coatings12101371
29. Jażdżewska M, Majkowska-Marzec B, Ostrowski R, Olive J-M. Influence of surface laser treatment on mechanical properties and residual stresses of titanium and its alloys. Advances in Science and Technology Research Journal. 2023;17(6). 10.12913/22998624/172981
30. Bash A, Resen AM, Atiyah A, Jasim K. Effect of Yb-YAG Laser Parameters on the Operating Regime of Plasma Sprayed NiCrAlY Premixed Coatings. Progress in Color, Colorants and Coatings. 2024;17(2):145-58. 10.30509/pccc.2023.167180.1240
31. Wang C, Tian P, Cao H, Sun B, Yan J, Xue Y, et al. Enhanced biotribological and anticorrosion properties and bioactivity of Ti6Al4V alloys with laser texturing. ACS omega. 2022;7(35):31081-97. 10.1021/acsomega.2c03166
32. Sun X, Lin H, Zhang C, Huang R, Liu Y, Zhang G, et al. Improved osseointegration of selective laser melting titanium implants with unique dual micro/nano-scale surface topography. Materials. 2022;15(21):7811. 10.3390/ma15217811
33. Naji T ABM, Resen A. . Laser Scanning Speed Influences on Assessment of Laser Remelted Commercially Pure Titanium Grade 2. International Journal of Engineering Transactions A: Basics. 2024 37(1):178-86. 10.5829/ije.2024.37.01a.16
34. Melero HC, Sakai RT, Vignatti CA, Benedetti AV, Fernández J, Guilemany JM, et al. Corrosion resistance evaluation of HVOF produced hydroxyapatite and TiO2-hydroxyapatite coatings in Hanks' solution. Materials Research. 2018;21(2):e20170210. 10.1590/1980-5373-mr-2017-0210
35. Alibash M, Resen AM, Kareem A, Muhi Abdulsahib Y. Improving the Hot Corrosion Behavior of Plasma-sprayed MCrAlY by RF Sputtering of TiO2 Nano-Coating and Laser Remelting Treatment. Progress in Color, Colorants and Coatings. 2024;17(1):27-38. 10.30509/pccc.2023.167142.1221

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

در زمینه ارتوپدی، آلیاژ Ti6Al4V به دلیل ویژگی‌هایی مانند چقرمگی بالا، خواص مکانیکی، مقاومت در برابر خوردگی بالا، چقرمگی شکست، زیست سازگاری و استحکام بالا، اغلب به عنوان یک زیست ماده مورد استفاده قرار می‌گیرد. با این حال، این آلیاژ دارای مشکلات خاصی مانند خوردگی و آزاد شدن عنصر وانادیوم به مایعات بدن است. روش‌های عملیات سطحی با فراهم کردن زیست فعالی و اثر سدگری، باعث بهبود اتصال استخوانی و زیست سازگاری می‌شوند. در این مطالعه، رفتار بیولوژیکی اکسید تیتانیوم اسپاترینگ (TiO2)با استفاده از روش اسپاترینگ مگنترون فرکانس رادیویی (RF)و همچنین ذوب مجدد لیزری سطح Ti6Al4V از پیش اسپاترینگ شده با استفاده از لیزر Nd: YAG بررسی شد. ریزساختار، فازها، ترشوندگی، توپوگرافی، ترکیب، زیست فعالی سطوح اصلاح شده با استفاده از SEM، XRD، EDS، زبری، سختی، خوردگی الکتروشیمیایی و آنالیز طیف‌سنجی امپدانس الکتروشیمیایی (EIS)مشخص شد. نتایج نشان می‌دهد که رفتار زیرلایه تحت تأثیر مثبت کاربرد هر دو تکنیک مهندسی سطح، شامل پاشش پوشش سرامیکی و اصلاح سطح با عملیات ذوب مجدد لیزری، قرار گرفته است. این بهبود به دلیل تشکیل سطوح زیست فعال و مقاوم‌تر است. اصلاح سطح دو عملکرد انجام می‌دهد. عملکرد اول به عنوان مانعی عمل می‌کند که از رسیدن مایعات بدن به ایمپلنت جلوگیری می‌کند و همچنین زبری و تخلخل سطحی مشابه زبری سطح استخوان را برای رشد بافت اطراف فراهم می‌کند.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1152-1160

International Journal of Engineering

Journal Homepage: www.ije.ir

Evaluation of Efficiency of Associated Petroleum Gas Utilization under Conditions of Low Atmospheric Humidity

A. V. Mysin*[a], M. V. Bykova[b], E. E. Soloveva[c]

[a] *Empress Catherine II Saint Petersburg Mining University, Drilling Department, Saint Petersburg, Russia*

[b] *Empress Catherine II Saint Petersburg Mining University, Research Center "Assessment of anthropogenic transformation of ecosystems", Saint Petersburg, Russia*

[c] *Empress Catherine II Saint Petersburg Mining University, Department of Geoecology, Saint Petersburg, Russia*

P A P E R I N F O

Paper history:
Received 17 April 2025
Received in revised form 24 June 2025
Accepted 13 July 2025

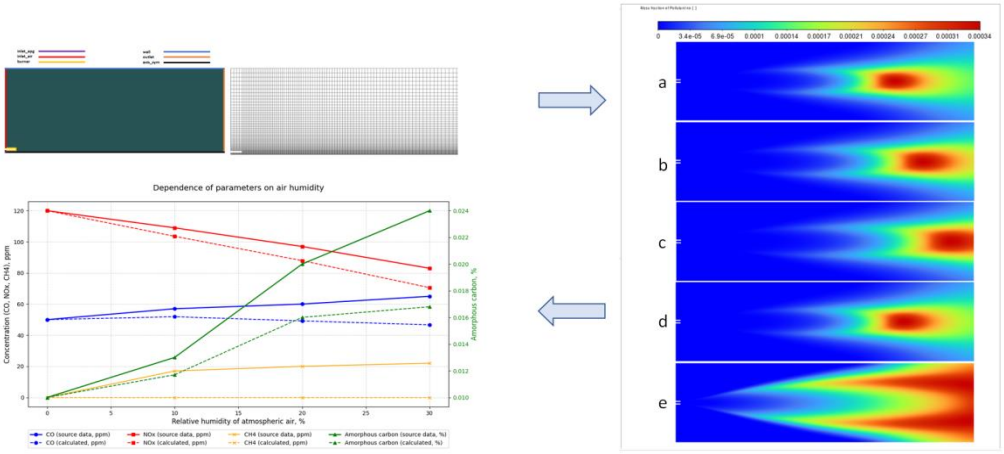
Keywords:
Associated Petroleum Gas
Numerical Modeling
Amorphous Carbon
Water Vapor
Carbon Dioxide
Nitrogen Oxides

A B S T R A C T

The study is devoted to the analysis of changes in the composition and amount of emissions during the utilization of associated petroleum gas (APG) under conditions of low humidity of the atmospheric air. An analysis of modern methods for assessing pollutant emissions and thermal effects during flaring of APG is presented. For a more complete assessment of the APG flaring process, determining the composition of emissions, flare stability when changing the flow rate and composition of gas due to the introduction of water vapor, a calculation was performed in ANSYS Fluent, which allows simulating the dynamics of liquids and gases (CFD). Numerical studies were performed based on the solution of equations of conservation of mass, momentum, energy and mass transfer of particles. Within the framework of the study, the composition, completeness of combustion (efficiency) were determined by adding water vapor to the combustion zone as one of the ways to reduce the formation of amorphous carbon. A one-step model was used to predict the formation of amorphous carbon. According to the modeling results, it was found that with a water vapor content of about 30%, a reduction in the formation of amorphous carbon, nitrogen oxides and carbon dioxide was observed. The greatest distance of the combustion zone from the torch was observed with the introduction of 10% water vapor. As a result of the study, it was established that the achievement of maximum effective utilization of APG in conditions of low atmospheric humidity occurs with the introduction of 30% water vapor into the combustion zone of the flare burner.

***doi**: 10.5829/ije.2026.39.05b.09*

Graphical Abstract



* Corresponding Author Email: *mysin_av@pers.spmi.ru* (A. V. Mysin)

Please cite this article as: Mysin AV, Bykova MV, Soloveva EE. Evaluation of Efficiency of Associated Petroleum Gas Utilization under Conditions of Low Atmospheric Humidity. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1152-60.

NOMENCLATURE

R_i	Rate of formation of chemical reaction products	u	Average speed (m/s)
S_i	Rate of product formation taking into account the addition of the dispersed phase	u'	Pulsation speed (m/s)
N	The total amount of chemical components in the system	k_{eff}	Effective thermal conductivity coefficient
J_i	Diffusion flux of the i- component arising due to the concentration gradient	h	Enthalpy
p	Pressure (Pa)	Y_M	Contribution of fluctuating expansion in compressible turbulence to the total dissipation rate
T	Temperature (K)	$C_1\ C_2\ C_{1\varepsilon}\ C_{3\varepsilon}$	Constants
k	Kinetic energy of turbulence (J/kg)	$S_k\ S_\varepsilon$	User-defined quantities based on initial conditions
G_k	Kinetic energy of turbulent flow	**Greek Symbols**	
G_b	Kinetic energy of a turbulent flow arising from buoyancy	τ	Shear stress tensor
Y_{soot}	Mass fraction of amorphous carbon (%)	ρ	Density (kg/m^3)
R_{soot}	Amorphous carbon formation rate (kg/m^3·s)	ω	Dissipation of turbulent energy (s^{-1})
ε	Specific energy dissipation (W/kg)	$\sigma_k, \sigma_\varepsilon$	Turbulent Prandtl numbers
ν	Kinematic viscosity	μ_t	Turbulent dynamic viscosity

1. INTRODUCTION

Over the past two years, oil production in the world has remained at a consistently high level. In the US, it is 12.9 million barrels per day, in Saudi Arabia and Russia, about 10.5 million barrels per day. This in turn indicates high volumes of recoverable associated petroleum gas. To achieve sustainable development goals, it is necessary to utilize associated petroleum gas with a high degree of efficiency. According to the data presented in Figure 1, it can be noted that in the US, almost all associated petroleum gas is processed or used. In Russia, there is also a significant volume of gas utilization; but in Iraq and Nigeria, there is a low level of utilization (1, 2).

Integrated approaches to solve the problem of increasing the efficiency of APG utilization, taking into account the development of technologies, have led to the creation of a large number of different methods.

The most promising and frequently used include the following: processing into liquid hydrocarbons; injection into the reservoir; power generation; compression and transportation; obtaining LPG; combustion. The last of the above methods still occupies the main share among all those used.

Among the countries of the world, Russia occupies one of the leading positions in hydrocarbon production and is a leader in associated petroleum gas flaring. About 15% of the world's volume of flared gas is accounted for by Russia (3-5).

Large deposits, where the bulk of hydrocarbons are produced, are located in the northern regions of the country and are the main supplier of pollutants into the atmosphere.

However, Russia is not the only one influencing the Arctic region of the planet: the United States of America also considers the Arctic territories as a potential expansion of its borders, including for gaining access to new hydrocarbon deposits.

The Sixth Assessment Report of the United Nations (UN) Intergovernmental Panel on Climate Change (IPCC) for 2021 confirmed that the observed global warming is the most pronounced in the Arctic (6).

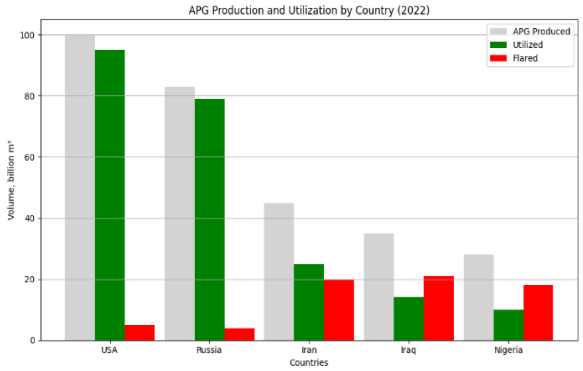


Figure 1. APG production and utilization volumes (3)

This is reflected not only in changes in physical factors such as increased Arctic land and ocean surface temperatures, increased precipitation, reduced snow cover, melting permafrost, and reduced sea ice thickness and extent (Figure 2), but also in increased intensity and frequency of extreme events associated with rapid sea ice loss, melting of the Greenland ice sheet, and wildfires (7).

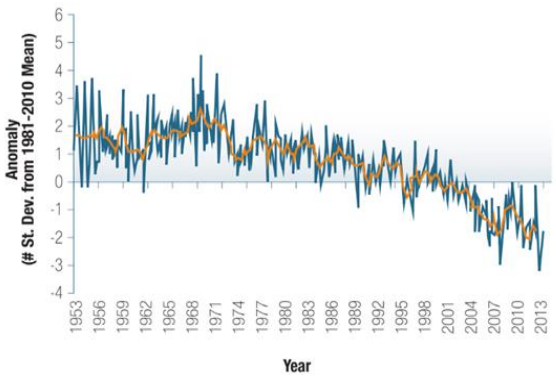


Figure 2. Arctic sea ice extent anomalies (deviations from average) – 1953-2013 (1, 2)

Transformation of the cryosphere due to global warming affects ecosystems across the Arctic, altering productivity, seasonality, distribution and interactions of species in terrestrial, coastal and marine ecosystems (3, 4). This in turn affects the cycling of carbon and greenhouse gases.

One of the causes of climate change is the emission of greenhouse gases or "short-lived climate pollutants", which include amorphous carbon (soot), methane, ozone, hydrogen-containing compounds of fluorine and carbon (8, 9). The main process that results in the largest emission of methane and black carbon into the atmosphere is the combustion process. In the industrial sector, a significant contribution to the release of pollutants into the atmosphere is made by flare flaring of associated petroleum gas (10-12).

According to the Circum-Arctic Resources Assessment Report published by the United States Geological Survey (USGS), the total undiscovered oil and gas resources in the Arctic region account for 22% of the world's total undiscovered oil and gas resources (2). The potential increase in the number of oil fields being developed in low humidity conditions, such as the Arctic region, will certainly exacerbate the problem of global warming. Thus, assessing and reducing greenhouse gas emissions from APG flaring in low humidity conditions is a pressing issue. It should be noted that the main difficulties encountered in quantifying pollutants released into the atmosphere are related to the changing composition of emissions from APG utilization at a flare unit, due to the component composition of the hydrocarbons produced.

The aim of the study is to assess the composition of pollutant emissions during the utilization of associated petroleum gas in conditions of low atmospheric humidity based on numerical modeling.

2. MATERIALS AND METODS

Currently, one of the most widely used methods for assessing emissions of pollutants into the atmosphere during the utilization of APG is satellite monitoring. Earth remote sensing data (hereinafter referred to as ERS) allows for prompt monitoring of the status and number of facilities utilizing APG.

Zhizhin et al. (13) presented an assessment of emissions based on remote sensing with subsequent processing of space images based on the VIIRS Nightfire algorithm. This algorithm relies on the detection of a source signal in the near, mid, and far IR ranges and allows for the recognition and description of “hot spots” on the Earth’s night surface (14). The method recognizes flares as point, stable, high-temperature sources of thermal IR radiation, which allows them to be separated from industrial facilities or natural phenomena. The decoding of daytime images in the near to far IR range by this algorithm is complicated by “noise” from sunlight. In medium-resolution photographs, “hot spots” are not recognized.

A similar algorithm for finding “hot spots” is based on satellite images from the Landsat 8 satellite (15). It is published on the official website of the US Geological Survey and can be implemented using a program in the Python programming language and ArcGIS software tools. The algorithm identifies thermal spots by converting four pre-processed raster images into a vector image representing groupings of pixels with higher brightness that can be interpreted as thermal spots (16).

Another method implemented on the TROPOMI spectrometer (Sentinel-5P satellite) allows indirectly assessing the volumes of APG utilization based on data on the concentrations of some chemicals in the vertical column of the troposphere (17). The TROPOMI spectrometer takes pictures in the ultraviolet, visible, near infrared and mid-infrared ranges of the electromagnetic spectrum (18). Obviously, such approaches require high-quality space images with sufficient resolution. In addition, based on the assessment results, it is possible to obtain the total volume of APG utilization of only some chemicals.

Mathematical methods are also used to assess pollutant emissions. Their advantage lies in the wide variety of implementation methods, for example, using mathematical software, MathCAD or MATLAB, or other similar programs (19). In addition, software from Integral, well known among environmental engineers, allows one to estimate gas emissions from flares (20-22).

For a more complete assessment of the process of utilization of exhaust gases in the flare (completeness of combustion, absence of formation of aldehydes, acids and other intermediate components), determination of the emission composition (including the presence of amorphous carbon) and, accordingly, the stability of the flare when changing the flow rate and composition of the gas, it is recommended to perform calculations in programs that allow modeling the dynamics of liquids and gases (CFD). Three-dimensional mathematical modeling implemented in the ANSYS Fluent program can most clearly implement this task. This software allows you to solve complex and diverse three-dimensional problems about the flow of liquid and gas. ANSYS Fluent implements a number of mathematical models that can be used with the Navier-Stokes equations to describe chemical or physical processes, such as the combustion process.

3. NUMERICAL SIMULATION

The CFD modeling method is based on solving the equations of conservation of mass, momentum, energy

and the equation of mass transfer of particles. Additionally, turbulence models (e.g. SST; RSM) are solved in scalar equations. These equations describe such quantities as: kinetic energy of turbulence (k); dissipation of turbulent energy (ω); specific energy dissipation (ε).

When the mass conservation equations are solved for chemical species, the computational code can predict the local mass of each species by solving the diffusion equation for the i-th species. This mass conservation equation takes the following general form (Equation 1) (23):

$$\frac{\partial}{\partial t}(\rho Y_i) + \nabla \cdot (\rho v Y_i) = -\nabla \cdot J_i + R_i + S_i \quad (1)$$

This equation is solved for n-1 chemical components in the system.

The basic equations that determine the flow field (Equations 2-4):

$$\frac{\partial}{\partial x_i}(\rho u_i) = 0 \quad (2)$$

$$\frac{\partial}{\partial x_i}(\rho u_i u_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}\left[\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}\delta_{ij}\frac{\partial u_k}{\partial x_k}\right)\right] + \frac{\partial}{\partial x_j}\left(-\overline{\rho u_i' u_j'}\right) \quad (3)$$

$$\frac{\partial}{\partial x_i}\left[u_i\left(\rho\left(h + \frac{1}{2}u_j u_i\right)\right)\right] = \frac{\partial}{\partial x_j}\left[k_{eff}\frac{\partial T}{\partial x_j} + u_i(\tau_{ij})_{eff}\right] \quad (4)$$

The calculation uses a "realizable" k-ε turbulence model. The standard formulation of the k-ε model is well known and has been applied to many problems since it was developed by Launder and Spalding (24-26). The variables to be carried forward are k, the kinetic energy of turbulence, and ε, the specific dissipation, which determines the scale of turbulence.

According to the ANSYS Fluent user manual, the "realizable" k-ε model satisfies certain mathematical constraints on the Reynolds stresses consistent with the physics of turbulent flows. The following expressions are used to calculate the flow velocities (Equations 5 and 6):

$$\frac{\partial}{\partial t}(\rho k) = \frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_j}\right] + G_k + G_b - \rho\varepsilon - Y_M + S_k \quad (5)$$

$$\frac{\partial}{\partial t}(\rho\varepsilon) = \frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_\varepsilon}\right)\frac{\partial \varepsilon}{\partial x_j}\right] + \rho C_1 S\varepsilon - \rho C_2 \frac{\varepsilon^2}{k+\sqrt{v\varepsilon}} + C_{1\varepsilon}\frac{\varepsilon}{k}C_{3\varepsilon}G_b + S_\varepsilon \quad (6)$$

The numerical study of the combustion process of associated petroleum gas was carried out under the condition of complete preliminary mixing with atmospheric air. The geometry of the model under study is shown in Figure 3. To reduce the time spent on the calculation, a two-dimensional axisymmetric formulation was used.

The computational domain is a combustion chamber with a diameter of 400 mm and a length of 1500 mm. The axial line is the x-axis, relative to which the symmetry condition is adopted. To improve the quality of the solution, the finite element mesh was adapted (Figure 4). As can be seen from Figure 4a, the mesh thickening region is performed in the areas of intense chemical reactions at high temperatures. The boundary conditions adopted in the study of the combustion process for the computational model are shown in Figure 4b. The chamber walls (wall), according to the conditions of Garréton and Simonin (27), were assumed to be impermeable and prevent slipping. At the inlet, the velocity and concentration profiles were considered uniform. The turbulence intensity was set equal to 5% for the inlet air and gas flows, respectively. In the outlet channel, zero diffusion flows were assumed for all variables. The effect of viscosity is not taken into account; there are no reactions with the burner material.

The combustion of associated gas was modeled using the Species Transport module of the Fluent solver and the Eddy dissipation model.

The eddy dissipation model uses the dissipation rate of eddies containing reactants and products to determine the combustion reaction rate. The eddy dissipation method in its purest form calculates reaction rates based on turbulent mixing rates. An implicit assumption is that chemical kinetic rates are much higher than turbulent

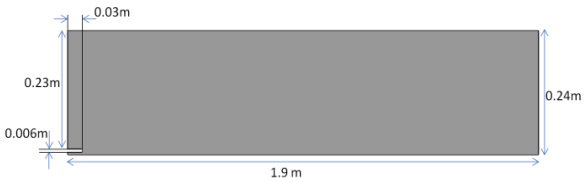


Figure 3. Torch burner

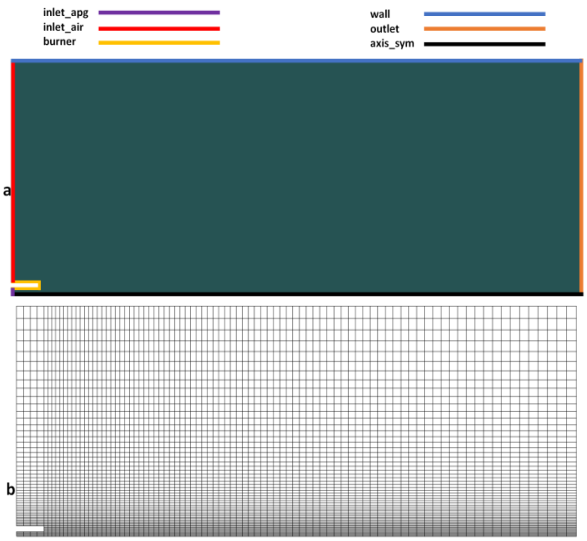


Figure 4. Calculation model: a - boundary conditions; b - grid

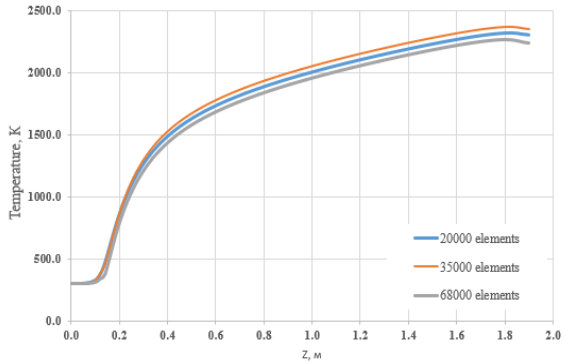


Figure 5. Results of grid convergence testing

mixing rates, i.e., turbulent mixing is the rate-limiting process. This assumption is valid for many combustion processes encountered in practice.

In the calculations, the convergence criterion is set at the level of 10^{-6} for the mass fraction of combustion components, the momentum equation and the continuity equation, and at the level of 10^{-4} for the convergence of the energy conservation equation (28).

Since the grid size is important when using numerical modeling (increasing the number of elements does not always lead to an increase in the accuracy of the solution), a grid independence test was performed. Three different grids were used for this: 1 - 20,000 elements, 2 -35,000 elements, 3 -68,000 elements. The calculation results are shown in Figure 5. The convergence criterion was the change in temperature at different distances from the point of supply of gas, air and water vapor.

As can be seen from Figure 5 in the above, the grid size of 20,000 elements fully satisfies the set tasks of the conducted modeling.

4. RESULTS AND DISCUSSION

The study included determination of the emission composition (formation of amorphous carbon, nitrogen oxides), combustion completeness (efficiency) taking into account the addition of water vapor to the combustion zone (5%, 10%, 20%, 30%) as one of the ways to reduce the formation of amorphous carbon.

4. 1. Analysis of the Combustion Temperature Field The temperature contours during the utilization of APG with the addition of water vapor are shown in Figure 6. It can be seen from the figure that the distribution of the combustion temperature under operating conditions with a mass fraction of water vapor (5%, 10%, 20%, 30%) is approximately the same.

The high temperature region is in the axial direction, and the range remains virtually unchanged, indicating stable combustion. It is also worth noting that the peak

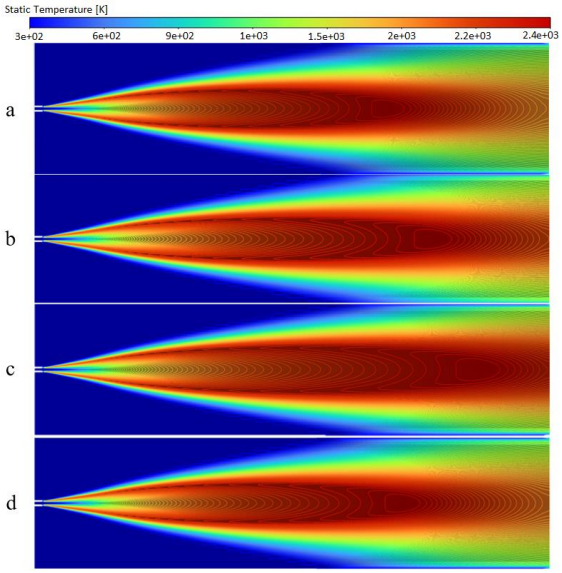


Figure 6. Temperature distribution in the combustion zone with different mass fractions of water vapor (a – 30%; b – 20%, c – 10%; d – 5%)

temperature increases slightly, and the position of the peak area shifts forward under the condition of adding 10% water vapor by mass.

4. 2. Formation of Products as a Result of Utilization of Associated Petroleum Gas

4. 2. 1. Amorphous Carbon Models of amorphous carbon formation can be classified depending on their complexity and the required computational costs. In the work of Kennedy (29), three categories are distinguished: empirical, semi-empirical and detailed models. One of the most well-known empirical models was proposed by Khan et al. (30) to predict emissions of amorphous carbon (soot) in diesel engines and, neglecting particle growth and oxidation rate, allows estimating soot formation as a function of pressure, temperature and equivalence coefficient of unburned gases. A more accurate description of the physics of the process can be obtained using semi-empirical models that take into account the chemical processes involved in soot formation, for example, expressing particle nucleation and surface growth based on the concentration of PAHs.

Zhu et al. (31) used a single-stage model of soot formation, which allows for a fairly accurate prediction of emissions resulting from combustion of associated petroleum gas. For this model, the equation for the transfer of the mass fraction of soot is solved (Equation 7):

$$\frac{\partial}{\partial t}(\rho Y_{soot}) + \nabla \cdot (\rho \vec{v} Y_{soot}) = \nabla \left(\frac{\mu_t}{\sigma_{soot}} \nabla Y_{soot} \right) + R_{soot} \quad (7)$$

The calculation results showed (see Figure 7) that with an increase in the mass fraction of water vapor in the fuel-

air mixture, the formation of amorphous carbon decreases.

4. 2. 2. CO_2 In this work, CO_2 formation was determined through stoichiometric coefficients, enthalpy of formation, and parameters controlling the reaction rate. From the CO_2 formation graphs (Figure 8), it can be concluded that as the water vapor content increases, the combustion efficiency increases and CO_2 formation decreases. That is, when water vapor was not added to the fuel-air mixture, the mass fraction of CO_2 was about 0.145. When the mole fraction of water vapor was 5%, 10%, 20%, and 30%, the mass fraction of CO_2 was 0.144, 0.143, 0.142, and 0.141, respectively.

4. 2. 3. NO_x The main reaction pathways, each with unique characteristics, responsible for the formation of NO_x in combustion processes are:

1. Thermal NO_x, formed as a result of the combination of atmospheric nitrogen and oxygen at high temperatures;
2. Fast-acting NO_x, formed as a result of the reaction of hydrocarbon fragments of the fuel with atmospheric nitrogen at an early stage of the flame front.

Thermal Formation of NO_x:

This mechanism is classified as thermal, due to the strong dependence of the rate of NO_x formation at high temperature (32, 33).

The rate of NO_x formation is determined by the effect of temperature and the residence time of nitrogen at this temperature. Combustion at temperatures below (1300 °C) forms significantly small-scale concentrations of thermal NO_x (34). Usually, at temperatures above this temperature (1600 °C), nitrogen (N_2) and oxygen (O_2) in the air decompose during combustion, taking an atomic form, or enter into successive reactions.

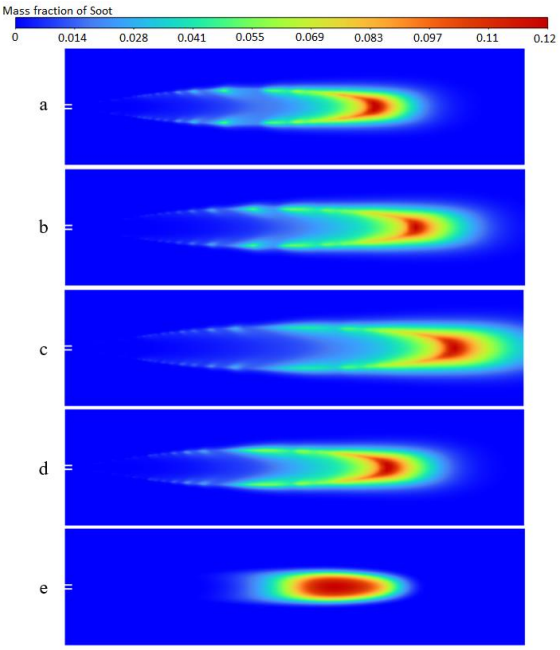


Figure 7. Formation of amorphous carbon at different mass fractions of water vapor (a – 30%; b – 20%, c – 10%; d – 5%; e – 0%)

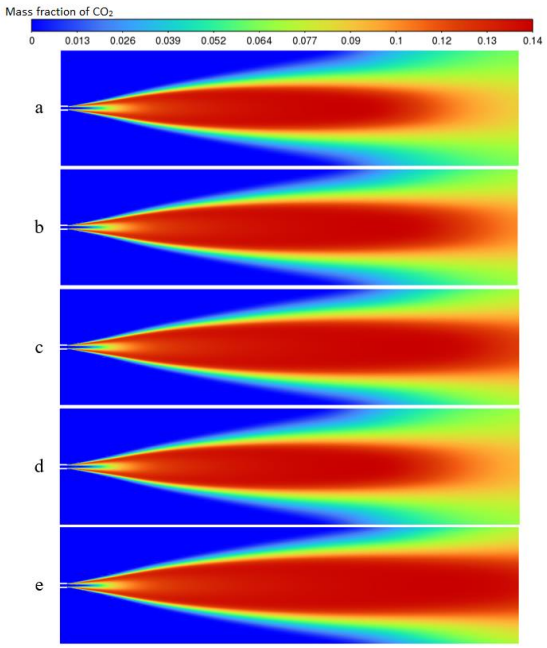


Figure 8. Formation of CO_2 at different mass fractions of water vapor (a – 30%; b – 20%, c – 10%; d – 5%; e – 0%). (Compiled by the authors)

Rapid NO_x formation:

Rapid NO_x formation is proportional to the number of carbon atoms present per unit volume and does not depend on the identity of the original hydrocarbon (35).

The amount of HCN formed increases with increasing concentration of hydrocarbon radicals, which in turn increases with increasing equivalence factor. As the equivalence factor increases, rapid NO_x formation first increases, then reaches a peak, and then decreases due to oxygen deficiency. In this paper, mechanisms 1 and 2 of NO_x formation are used, which allow fairly accurate prediction of emissions from APG utilization.

As can be seen from Figure 9, with an increase in the mass fraction of water vapor in the fuel-air mixture, the content of NO_x pollutants decreases.

According to the simulation results, it was found that with the highest water vapor content (30%), a reduction in the formation of amorphous carbon, nitrogen oxides and carbon dioxide was observed (Figure 10). The greatest distance of the combustion zone from the torch, i.e. a high altitude of pollutant formation, was observed with the introduction of 10% water vapor. In addition, the introduction of 30% water vapor slightly reduces the altitude of the temperature distribution in the combustion zone.

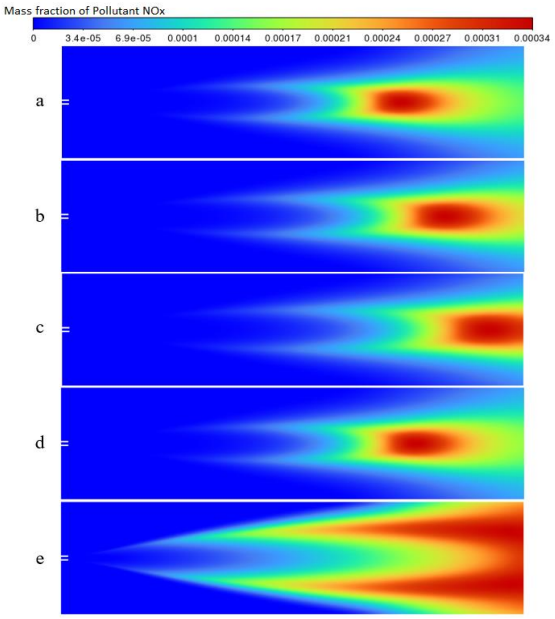


Figure 9. Formation of NOx at different mass fractions of water vapor (a – 30%; b – 20%, c – 10%; d – 5%; e – 0%)

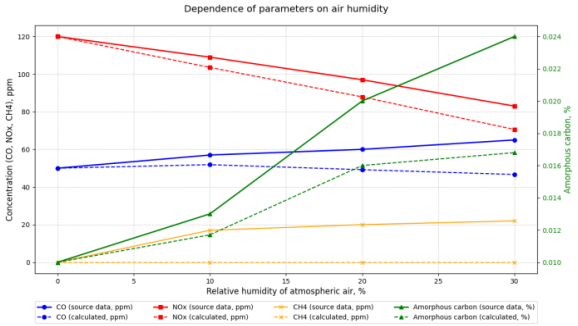


Figure 10. Formation of products during combustion of associated petroleum gas under low humidity conditions without the addition of water vapor (initial) and with different mass fractions of water vapor (calculated)

Thus, the modeling results confirm the fact that the introduction of water vapor into the flare combustion zone makes it possible to reduce the negative impact of APG utilization on the environment, since the completeness, and therefore the quality, of APG combustion increases and the formation of pollutants entering the atmosphere decreases.

The choice of a specific direction for solving the problem of utilization of associated petroleum gas in conditions of low atmospheric humidity can be made on the basis of a comparative analysis of existing methods (Table 1).

However, in the context of modern energy transformation and ESG-oriented agenda, the last three methods of APG utilization are the most promising, but due to the poorly developed infrastructure and remoteness of the hydrocarbon deposits being developed, the method being studied is the most optimal.

TABLE 1. Comparative analysis of the main methods of utilization of associated petroleum gas

Method	Sustainability	Technological complexity	Economic efficiency
Flaring of APG	Low	Low	Low
Flaring of APG (+ 30% water vapor)	Average	Low	Low
Injection into the reservoir	High	High	Low
Electricity generation	Average	High	High
Processing into liquid hydrocarbons	High	High	Average
Compression and transportation	High	Average	Average

5. CONCLUSIONS

Currently, hydrocarbon deposits characterized by low atmospheric humidity, such as those located in the Arctic, are promising regions with significant reserves of hydrocarbon raw materials. However, back in 2021, the UN confirmed that global warming has the greatest impact on the Arctic. Therefore, the tasks associated with reducing emissions of pollutants into the atmosphere, in particular greenhouse gases from the utilization of associated petroleum gas, as the main contributor to "short-lived climate pollutants", are an urgent problem that must be addressed.

As part of this work, an analysis was made of the quantitative change in the composition of greenhouse gases from the utilization of associated petroleum gas for atmospheric air conditions with low humidity, taking into account the addition of water vapor to the combustion zone. Mathematical models were constructed using ANSYS Fluent software.

Before carrying out the calculations, the geometry of the model under study was formed, grid convergence was tested and, in accordance with it, a computational grid was constructed.

The efficiency of fuel-air mixture combustion and pollutant formation were assessed for different mass fractions of water vapor.

6. ACKNOWLEDGMENT

The research was carried out within the state assignment of Ministry of Science and Higher Education of the Russian Federation (FSRW-2023-0002).

7. REFERENCES

1. Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, et al. Summary for policymakers. 2021. 10.1017/9781009157896.001
2. AMAP E. Arctic climate change update 2021: Key trends and impacts. Summary for policy-makers. Tromsø Norway; 2021.
3. Bank W. Global gas flaring reduction partnership (GGFR). 2022.
4. Litvinenko VS, Petrov EI, Vasilevskaya DV, Yakovenko AV, Naumov IA, Ratnikov MA. Assessment of the role of the state in the management of mineral resources. Записки Горного института. 2023(259 (eng)):95-111. 10.31897/PMI.2022.100
5. Cherepovitsyn AE, Tsvetkov PS, Evseeva OO. Critical analysis of methodological approaches to assessing sustainability of arctic oil and gas projects. Записки Горного института. 2021;249:463-78. 10.31897/PMI.2021.3.15
6. Babyr N. Topical themes and new trends in mining industry: Scientometric analysis and research visualization. International Journal of Engineering, Transactions B: Applications. 2024;37(2):439–51. 10.5829/IJE.2024.37.02B.18
7. Secretariat UNFCoCC. United Nations framework convention on climate change: UNFCCC; 1992.
8. Smirnova E, Saychenko L. Hydrodynamic Modeling and Evaluation of Partial Substitution of Cushion Gas During Creation of Temporary Underground Gas Storage in an Aquifer. International Journal of Engineering Transactions A: Basics. 2024;37(07):1221-30. 10.5829/ije.2024.37.07a.02
9. Blinov P, Salakhov K, Nikishin V, Kuchin V. Development of Cement Slurry Composition with Self-healing Properties. International Journal of Engineering, Transactions A: Basics. 2025;38:236-46. https://doi.org/10.5829/ije.2025.38.01a.21
10. Semenova T. Value improving practices in production of hydrocarbon resources in the Arctic regions. Journal of Marine Science and Engineering. 2022;10(2):187. 10.3390/jmse10020187
11. Blinov PA, Sadykov MI, Gorelikov VG, Nikishin VV. Development and research of backfill compounds with improved elastic and strength properties for oil and gas well lining. Записки Горного института. 2024(268 (eng)):588-98.
12. Buslaev G, Morenov V, Konyaev Y, Kraslawski A. Reduction of carbon footprint of the production and field transport of high-viscosity oils in the Arctic region. Chemical Engineering and Processing-Process Intensification. 2021;159:108189. 10.1016/j.cep.2020.108189
13. Zhizhin M, Matveev A, Ghosh T, Hsu F-C, Howells M, Elvidge C. Measuring gas flaring in Russia with multispectral VIIRS nightfire. Remote Sensing. 2021;13(16):3078. https://doi.org/10.3390/rs13163078
14. Elvidge CD, Zhizhin M, Baugh K, Hsu F-C, Ghosh T. Methods for global survey of natural gas flaring from visible infrared imaging radiometer suite data. Energies. 2015;9(1):14. https://doi.org/10.3390/en9010014
15. Landsat U. Landsat 8-9 collection 2 (c2) level 2 science product (l2sp) guide. United States Geological Survey: Asheville, NC, USA. 2022:1-42.
16. Anejionu OC, Blackburn GA, Whyatt JD. Detecting gas flares and estimating flaring volumes at individual flow stations using MODIS data. Remote Sensing of Environment. 2015;158:81-94. https://doi.org/10.1016/j.rse.2014.11.018
17. Veefkind JP, Aben I, McMullan K, Förster H, De Vries J, Otter G, et al. TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. Remote sensing of environment. 2012;120:70-83. https://doi.org/10.1016/j.rse.2011.09.027
18. Ialongo I, Stepanova N, Hakkarainen J, Virta H, Gritsenko D. Satellite-based estimates of nitrogen oxide and methane emissions from gas flaring and oil production activities in Sakha Republic, Russia. Atmospheric Environment: X. 2021;11:100114. https://doi.org/10.1016/j.aeaoa.2021.100114
19. Palaev A, Fuming Z. A Leak Detection Method for Underground Polyethylene Gas Pipelines Using Simulation Software Ansys Fluent. International Journal of Engineering Transactions B: Applications. 2024;37(8):1615-21. 10.5829/ije.2024.37.08b.14
20. Nwosisi M, Oguntoke O, Taiwo A. Dispersion and emission patterns of NO2 from gas flaring stations in the Niger Delta, Nigeria. Modeling earth systems and environment. 2020;6(1):73-84. https://doi.org/10.1007/s40808-019-00658-z
21. Strizhenok AV, Ivanov AV. Monitoring of air pollution in the area affected by the storage of primary oil refining waste. Journal of Ecological Engineering. 2021;22(1):60-7. https://doi.org/10.12911/22998993/128873
22. Mysin A, Kovalevskiy V, Kirkin A, editors. Ensuring environmental safety of massive explosions in the combined development of coal deposits in Kuzbass. IOP Conference Series: Earth and Environmental Science; 2021: IOP Publishing.
23. Anderson JD, Wendt J. Computational fluid dynamics: Springer; 1995.
24. Noor M, Wandel AP, Yusaf T, editors. Detail guide for CFD on the simulation of biogas combustion in bluff-body mild burner. Proceedings of the 2nd International Conference on Mechanical Engineering Research (ICMER 2013); 2013: University of Southern Queensland.
25. Paul C, Haworth DC, Modest MF. A simplified CFD model for spectral radiative heat transfer in high-pressure hydrocarbon–air combustion systems. Proceedings of the Combustion Institute. 2019;37(4):4617-24. 10.1016/j.proci.2018.08.024
26. Shi Y, Zhong W, Chen X, Yu A, Li J. Combustion optimization of ultra supercritical boiler based on artificial intelligence. Energy. 2019;170:804-17. 10.1016/j.energy.2018.12.172
27. Garréton D, Simonin O, editors. Aerodynamics of steady state combustion chambers and furnaces. ASCF Ercoftac Cfd Workshop; 1994.
28. Perrone D, Castiglione T, Klimanek A, Morrone P, Amelio M. Numerical simulations on Oxy-MILD combustion of pulverized coal in an industrial boiler. Fuel Processing Technology. 2018;181:361-74. 10.1016/j.fuproc.2018.09.001
29. Kennedy IM. Models of soot formation and oxidation. Progress in Energy and Combustion Science. 1997;23(2):95-132.
30. Khan I, Greeves G, Probert D. Prediction of soot and nitric oxide concentrations in diesel engine exhaust. Air Pollution Control in Transport Engines C. 1971;142:205-17.
31. Zhu Q, Niu Y, Chen X. Impact of water vapor on the soot formation in the laminar methane/air diffusion flame. J Saf Environ. 2017;17:174-7. 10.1016/j.combustflame.2013.12.017
32. Zeldvich YB. The oxidation of nitrogen in combustion and explosions. J Acta Physicochimica. 1946;21:577.
33. Chen Z, Li Z, Zhu Q, Jing J. Gas/particle flow and combustion characteristics and NOx emissions of a new swirl coal burner. Energy. 2011;36(2):709-23. 10.1016/j.energy.2010.12.037
34. Fenimore CP, editor Formation of nitric oxide in premixed hydrocarbon flames. Symposium (international) on combustion; 1971: Elsevier.
35. Zhao L, Wang T, editors. Simulation of combustion and thermal-flow inside a pyroscrubber. ASME International Mechanical Engineering Congress and Exposition; 2009.

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

این مطالعه به تجزیه و تحلیل تغییرات در ترکیب و میزان انتشار گازهای گلخانه ای در طول استفاده از گاز نفتی مرتبط در شرایط رطوبت پایین هوای جوی اختصاص داده شده است .تجزیه و تحلیل روش های مدرن برای ارزیابی انتشار آلاینده ها و اثرات حرارتی در طول شعله ور شدن گاز نفتی مرتبط ارائه شده است .برای ارزیابی کامل تر فرآیند شعله ور شدن گاز نفتی مرتبط ، تعیین ترکیب انتشار گازهای گلخانه ای ، ثبات شعله ور شدن هنگام تغییر سرعت جریان و ترکیب گاز به دلیل معرفی بخار آب ، محاسباتی در Ansys Fluent انجام شد که امکان شبیه سازی دینامیک مایعات و گازها را فراهم می کند .مطالعات عددی بر اساس حل معادلات حفظ جرم ، حرکت ، انرژی و انتقال جرم ذرات انجام شد .در چارچوب مطالعه ، ترکیب ، کامل بودن احتراق) کارایی (با اضافه کردن بخار آب به منطقه احتراق به عنوان یکی از راه های کاهش تشکیل کربن بی شکل تعیین شد .یک مدل یک مرحله ای برای پیش بینی تشکیل کربن بی شکل استفاده شد .با توجه به نتایج مدل سازی ، مشخص شد که با محتوای بخار آب حدود 30٪ ، کاهش تشکیل کربن بی شکل ، اکسید نیتروژن و دی اکسید کربن مشاهده شد .بیشترین فاصله منطقه احتراق از شعله ور ، یعنی بیشترین ارتفاع تشکیل آلاینده ، با معرفی 10 ٪ بخار آب مشاهده شد .در نتیجه این مطالعه مشخص شد که دستیابی به حداکثر استفاده موثر از گاز نفتی مرتبط در شرایط رطوبت کم اتمسفر با وارد کردن 30 ٪ بخار آب به منطقه احتراق مشعل رخ می دهد.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1161-1175

International Journal of Engineering

Journal Homepage: www.ije.ir

Optimized Deep Learning Model for Pomegranate Disease Detection: A Convolutional Neural Network Long Short-Term Memory Approach

P. Sahebgouda*[a], S. Maradithaya[b], A. Jadhav[c]

[a] *Department of Computer Science and Engineering, BLDEA's V. P. Dr. P. G. Halakatti College of Engineering and Technology, (Affiliated to Visvesvaraya Technological University, Belagavi-590018), Vijayapur, Karnataka, India*
[b] *Department of Computer Science and Engineering, M.S. Ramaiah Institute of Technology, (Affiliated to Visvesvaraya Technological University, Belagavi-590018), Bengaluru, Karnataka, India*
[c] *Department of Electrical and Electronics Engineering, BLDEA's V.P.Dr.P.G.Halakatti College of Engineering and Technology, (Affiliated to Visvesvaraya Technological University, Belagavi-590018) Vijayapur, India*

PAPER INFO

Paper history:
Received 26 March 2025
Received in revised form15 June 2025
Accepted 17 June 2025

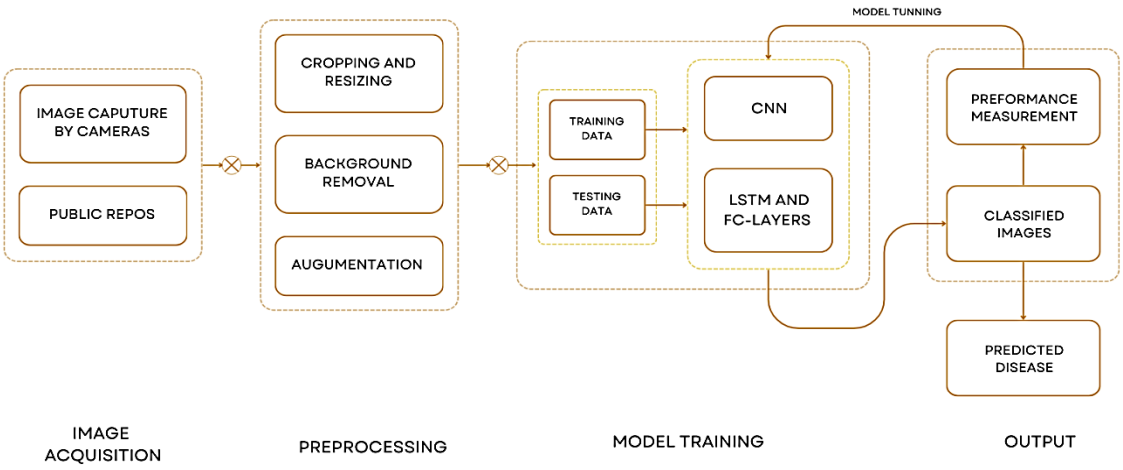
Keywords:
Pomegranate
Disease Detection
Deep Learning
Convolutional Neural Network
Long Short-Term Memory
Machine Learning

ABSTRACT

Pomegranate is a high-value fruit globally recognized for its nutritional benefits and applications in traditional medicine and cosmetics. India is a key player in the global pomegranate market, but the industry faces challenges such as diseases that affect crop productivity and economic losses for farmers. This study proposes a novel approach to pomegranate disease detection using a hybrid Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) model. The proposed model leverages CNNs for effective feature extraction and LSTMs for sequential data handling, achieving superior performance compared to traditional methods and other deep learning techniques. Experimental results demonstrate high accuracy, recall, precision, and F1 score. The Proposed model achieved an accuracy of 98.53% and loss of 0.0677. The study also explores the limitations of transfer learning approaches such as VGG16 and ResNet50, and larger models like AlexNet, which did not perform well in this context. The findings suggest that the hybrid CNN-LSTM model offers a scalable and adaptable solution for agricultural disease detection, with potential applications for various crops.

***doi**: 10.5829/ije.2026.39.05b.10*

Graphical Abstract



System Architecture of the Proposed CNN-LSTM Model

*Corresponding author email: cse.sahebgouda@bldeacet.ac.in (P. Sahebgouda)

Please cite this article as: Sahebgouda P, Maradithaya S, Jadhav A. Optimized Deep Learning Model for Pomegranate Disease Detection: A Convolutional Neural Network Long Short-Term Memory Approach. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1161-75.

NOMENCLATURE

x_t	Current input	h_t	Hidden state
$\tilde{C}_t$	Candidate cell state	h_{t-1}	Previous hidden state
C_{t-1}	Old cell state	O_t	Output gate's activation
W_i, W_f, and W_o	Weight matrices	b_i, b_f, and b_o	Biases associated with the respective gates

1. INTRODUCTION

Pomegranate (*Punica granatum*) is one of the most valuable fruits available. It is globally recognized as a "Super-food" owing to its nutritious characteristics (1). The leathery red skin encases vibrant, ruby-red arils makes up for a delicious sweet food which also offers a variety of health benefits. Pomegranate fruit is rich in vitamins and dietary fiber (2).

Beyond its use as a food, pomegranate has all found many applications in various sectors, from traditional medicine to cosmetics. The fruit's extracts are utilized in the pharmaceutical industry for their potential in treating various medical conditions, such as cardiovascular diseases and cancer (3). Moreover, the aesthetic and therapeutic properties of pomegranate make it a good ingredient in skincare and beauty products.

India is one of the leading countries in pomegranate production (4). Contributing significantly to the global pomegranate market. India provides a suitable environment for pomegranate cultivation. According to recent statistics, our country has witnessed an increase in pomegranate production, making it a key player in the global supply chain (5).

Pomegranates are grown on more than 500,000 hectares worldwide, with a projected 6.54 million metric tons produced globally as of 2021–2022. The United States, Spain, China, Iran, and Turkey are other important producers. With 3.22 million metric tons, India leads the world in production, mostly from states like Telangana (2.34%), Maharashtra (54.85%), Gujarat (21.28%), Karnataka (9.51), and Andhra Pradesh (8.82%).

However, the thriving pomegranate industry faces many formidable challenges like drought, climate change, etc. With diseases being a major factor contributing to yield losses. Pomegranate orchards are susceptible to various diseases caused by fungi, bacteria, and viruses. These diseases not only affect the fruit's quality but also pose a threat to the overall health of the trees, potentially leading to long-term economic losses for farmers. One of the major disease is the Bacterial blight, which affects all parts of the plant but is more destructive on fruits (6). Other common diseases include Alternaria, Anthracnose, etc. affecting both plant and fruit. The impact of these diseases extends beyond individual crops, affecting the livelihoods of farmers and the overall economy of regions dependent on pomegranate cultivation.

Traditional methods of disease detection in agriculture rely on visual inspection and manual monitoring which are inherently time-consuming and often lack the precision required for early disease detection. Advent of Machine learning techniques, provides us an opportunity to implement the new approaches to disease detection in pomegranate cultivation.

These new machine learning techniques enable us to build automation disease detection systems which are more efficient and accurate. Deep Learning techniques, which a subset of machine learning, is becoming increasingly popular in research as graphics processing technology advances and low-cost GPUs become more widely available. Within the realm of image processing, a deep learning technique known as Convolutional Neural Networks (CNNs) have emerged as a particularly effective deep learning technique.

This research proposes a hybrid model i.e. combining CNN with LSTM layer. This paper also compares other state-of-the-art models to the proposed model.

1. 1. Overview of Pomegranate Fruit Diseases

Pomegranate crop is susceptible to many diseases that pose significant challenges to the health of fruit and crop productivity. Improper Agricultural practices and harsh climate can promote diseases. Figure 1 shows the diseased pomegranate fruits with corresponding disease name below it. Below are few common diseases that affect pomegranate.

1) Bacterial Blight (*Xanthomonas axonopodis*)

Bacterial blight in pomegranates is primarily caused by the bacterium *Xanthomonas axonopodis pv. punicae*. The disease manifests itself as watery lesions on leaves, stems and fruits. These lesions expand and merge, forming large necrotic areas. On fruits, the spots are dark brown to black and can be slightly raised with a cracked appearance. In severe cases, fruits can crack open, leading to secondary infections and significant economic losses. The epidemiology of bacterial blight indicates that it is favored by warm and humid conditions.

Management of bacterial blight involves the use of copper-based bactericides, cultural practices to reduce humidity around plants, and pruning to remove infected plant part.

2) Alternaria Fruit Rot (*Alternaria alternata*)

Alternaria fruit rot disease is caused by the fungus *Alternaria alternata*. This disease is characterized by the appearance of small, black spots on the surface of the fruit, which eventually enlarge and merge, covering large areas of the fruit. Internally, the rot progresses, leading to blackened seeds and a mushy consistency of the arils.

The infection can start at the blossom end or through wounds on the fruit surface.

The development of *Alternaria* disease is typically favored by high humidity and warm temperatures. The fungal spores are spread by wind, rain, and insects, and the disease can be exacerbated by poor orchard hygiene and inadequate ventilation. Effective management includes maintaining good orchard sanitation, ensuring proper spacing and pruning to improve air circulation, and applying fungicides during periods of high risk. Post-harvest handling practices, such as rapid cooling and proper storage, can also help reduce the incidence of *Alternaria* rot.

3) Anthracnose (*Glomerella cingulata*)

Anthracnose of pomegranate is a disease caused by the fungus *Glomerella cingulata*. This disease is distinguished by the presence of dark, sunken lesions on the fruit, which can coalesce to cover larger areas. The affected tissues may exude a pinkish spore mass under humid conditions. Internally, the rot can extend into the fruit, causing decay and making the fruit unmarketable.

Anthracnose thrives in warm, wet conditions, similar to other fungal diseases. The spores are disseminated by rain splash, wind, and contaminated tools. The disease can also spread from infected plant debris. Effective management strategies include the removal and destruction of infected plant material, improving air circulation through pruning, and applying fungicides during critical periods of infection risk. Additionally, avoiding overhead irrigation and ensuring good drainage can help reduce the prevalence of anthracnose.

4) Cercospora (*Cercospora punicae*)

Cercospora fruit spot is caused by the fungus *Cercospora punicae*. Pomegranate fruits that have been infected develop small, circular to irregular spots that range from dark brown to black. As the disease progresses, these spots can grow and coalesce, often resulting in premature fruit drop. The disease can also affect leaves and stems, causing similar spotting and potential defoliation.

The fungus thrives in warm, wet conditions and is spread by rain splash and wind. It can survive on plant debris and in soil, contributing to ongoing infection cycles if not properly managed. Controlling Cercospora fruit spot involves cultural practices such as removing and destroying infected plant debris, improving air circulation through pruning, and applying appropriate fungicides. Crop rotation and avoiding overhead irrigation can also reduce the disease's spread.

Correct Identification of diseases at early stages can help farmers to mitigate the losses.The dataset mentioned in Figure 1 is used from Mendeley Data (7). The proposed system in this project can be used by farmers for accurate detection of diseases. The main advantage of employing a CNN-LSTM hybrid model for pomegranate disease detection lies in its integration of effective spatial feature extraction (CNN) with temporal pattern recognition (LSTM), resulting in enhanced accuracy, improved precision, and more resilient classification of fruit diseases relative to the use of CNN or LSTM independently.The proposed method uses image as input instead of Csv file which will helpful for algorithm to learn and obtain good accuracy.

Figure 1. Various Diseases Affecting Pomegranate

2. LITERATURE SURVEY

Several studies have explored the use of machine learning and image processing techniques for pomegranate disease detection and classification. This section explores the findings from the literature survey, which summarizes information from 24 research publications (Table 1).

This section is divided into three subsections each following a theme of machine learning like traditional techniques, deep learning techniques and emerging technologies.

2. 1. Traditional Machine Learning Techinques

Earlier attempts to detect pomegranate disease relied heavily on traditional machine learning techniques such as KNN, k-means, and SVM. For example, Tejal Deshpande et al. proposed a system that uses k-means clustering for image segmentation and grades infected areas based on color and size. Color and size of the infected area are not reliable indicators of disease (8). Their system can only detect healthy and bacterial blight. This limits their system's ability to perform multiclass classification tasks.

Several studies used K-means clustering for image segmentation, followed by SVM for classification. Bhange and Hingoliwala (9) proposed a system that uses K-means and SVM to classify pomegranate diseases with accuracy of 82%. Gaikwad et al. (10) applied similar techniques for multi-fruit disease detection. Their system showed remarkable accuracy.

Kantale and Thakare (11) utilized the Ada-Boost Ensemble algorithm to classify pomegranate diseases,

leveraging a dataset of 190 images. Their model achieved an impressive accuracy of 92.90%. However, despite its high accuracy, the Ada-Boost algorithm has some drawbacks, such as its sensitivity to noise and high computational demands, especially when training on large datasets. These limitations can hinder its practical application in various agricultural scenarios.

Neural networks have also been applied in pomegranate disease classification. Dhakate and Ingole (12) developed a system that uses Artificial Neural Network as a classifier and showed an accuracy around 90% with a dataset of 500 images. Patil (13) have focused on PNN (Probabilistic neural network) for disease classification and achieved an accuracy of 80.30% with a dataset of 166 images, which is divided into 100 for training and 66 or testing. These systems are in their early stages and require image preprocessing, such as noise reduction to be effective.

Otsu Threshold combined with k-means can be used for image segmentation before Image Classification. This approach was used by Pawar and Jadhav (14) along with multi layer neural network, this system achieved an accuracy of 90% on 40 test images. The study was limited to a small dataset. It suffers similar drawbacks as discussed above.

Reddy et al. (15) developed an ANN based approach for detecting diseases in grapes, apple and pomegranates. The system achieved 93% accuracy. The system uses color texture analysis for feature extraction and employs k-means clustering to segment an image. However, the authors did not provide details on the ANN architecture and hyperparameters.

A simple histogram equalization is applied on extracted ROI to enhance the features. These features are used by Multiclass SVM to classify the leaf disease. Using this approach the authors achieved 98.07% accuracy (16).

2. 2. Deep Learning Methods With the advent of deep learning, field of image processing has significantly elevated the accuracy and efficiency of disease detection.

Sharath et al. (17) have proposed a system for detecting diseases in fruits like pomegranate, orange, grapes, and papaya using convolutional neural networks (CNN). The system achieved around 91% accuracy in detecting diseases like citrus canker, greening, etc in grapes, and bacterial blight, borers, cercospora in pomegranates (18). Pawar et al. (19) have proposed a leaf disease system for multiple plants using 15 layer CNN model. Their system could detect diseases in 10 different plants with upto 93% accuracy. The authors did not report specific accuracy metrics for each plant species.

A newer more complex variant of CNN called Alexnet. AlexNet architecture has 8 layers, this model was employed by Wakhare et al. (20); their model achieved an accuracy of 97.6%. They have used a dataset of 1245 pomegranate leaf images. Their research was limited to detecting only two diseases which are bacterial blight and Alternaria (21).

An advanced region based CNN model called faster-RCNN was used in pomegranate disease, detection and classification in a research by Javeriya (22). Faster R-CNN is 250 times faster than RCNN. The model was designed to identify two common diseases—anthracnose and bacterial blight—using a custom dataset of annotated images. The authors acknowledged that low image resolution could lead to detection failures.

2. 3. Emerging Technologies Newer methods like transfer learning have emerged. In transfer learning, the pretrained models are utilized to fine tune on the pomegranate disease dataset. Commonly used pretrained models are VGG16, Inception-V3, Resnet, DenseNet. These models are the variants of CNN.

Nirgude and Rathi (21) compared three CNN based architectures they are Resnet-18, Resnet-50 and Inception-V3. They collected 1,493 images of the pomegranate fruits and the leaves at different stages of disease development over six months. They acknowledged the need to address the "Black Box" problem often associated with deep learning models. This could be handled by merging the Grad-CAM model. The authors did not provide specific implementation details to recreate the experiment.

Al Ansari (23) used deep learning techniques such as AlexNet and VGG-16 to demonstrate the ability to automatically diagnose diseases in pomegranate plants using leaf images. AlexNet achieved an accuracy of 89.57%, while VGG-16 achieved a higher accuracy of 95.23%. However, there are notable drawbacks to this study. The models were trained for only 12 epochs, which might not be sufficient for them to fully learn the features necessary for accurate disease diagnosis. For example a research paper by Qi et al. (24) used MSRCR defogging algorithm and image normalization techniques to enhance the images. They used Canny SLIC algorithm for precise segmentation of diseased region. A focal loss function was used to improve the DenseNet169 architecture, resulting in an recognition accuracy of 98.98% for three types of grape, outperforming traditional DenseNet's 97.95% accuracy. Additionally, the computational complexity and processing time associated with the proposed model are not discussed, which could be crucial for real-time applications.

Vasumathi and Kamarasan (7) proposed an architecture which combined CNN-LSTM model to classify a set of about 6519 pomegranate fruit images. Features like fruit color, number of spots, shape, size of diseased area, etc where extracted manually into a csv dataset. Their model achieved 98.17% on this dataset. The binary classification of normal or abnormal fruit limits its practical use.

TABLE 1. Literature Survey

Sl.No.	Part	Feature Extraction	Classifier	Accuracy	Critical Analysis
(1)	Both leaf and fruit	K- means clustering	Classification based on color and area	Not Mentioned	This system calculates the percentage of infection based on the size of the infected area. This is better than manual techniques but has lower accuracy (8).
(2)	Fruit	K- means clustering	SVM	82	Bhange and Hingoliwala (9) proposed a web based tool which uses Image processing techniques to determine if a pomegranate fruit is infected with bacterial blight or not. The SVM used here is not suitable for multiclass classification (9).
(3)	Both leaf and fruit	GLCM, PSO	Ada-Boost Ensemble	92.9	Gaikwad et al. (10) used a very small dataset of 190 images of both fruit and leaf. There may be a chance of over fitting.
(4)	Both leaf and fruit	GLCM	Neural Network	90	Kantale and Thakare (11) constructed a classifier from basic Artificial Neural Network(ANN) and uses the basic back-propagation algorithm to train ANN. 500 images of leaf and fruit were used as a dataset .
(5)	Fruit	Standard deviation, entropy, variance, smoothness & skewness.	Probabilistic neural network(PNN)	80.3	PNN is a special class of ANN. This network is not widely used because they are slower than other methods (12).
(6)	leaf	Otsu thresholding, K-means clustering	Multi layer neural network	90	Dhakate and Ingole (12) paper did not provide details of the model.
(7)	Multiple fruits including pomegranate	ANN	ANN	90	Accuracy of model can be improved by better preprocessing and restricting to classification of single fruit type (16).
(8)	Leaf	K- means clustering	Multi-Class SVM	98.07	Simple histogram equalization is applied on extracted ROI to enhance the features . Those features are used by Multiclass SVM to classify the leaf disease (15).
(9)	Both leaf and fruit	Grabcut Segmentation and canny edge detection	CNN	91	Authors used both leaf and fruit images. Dimensions of images were limited to 432x288 to reduce processing time. Morphological processing is also used to clean the image (16).
(10)	Fruit	Statistical Analysis and DWT	CNN, SVM	96.93, 96	Sharath et al. (17)compared segmentation techniques like manual segmentation and k-means using CNN and SVM classifiers, concluding that manual segmentation combined with CNN yielded the highest accuracy.
(11)	Leavesof multiple plants	None	CNN	93	Sánchez et al. (18) trained a CNN with 15 layers to detect diseases in 10 different plants with a dataset size of 50000 images. Their system also suggests pesticide for the detected disease.
(12)	Leaf	None	Alexnet	98.07	Pawar et al. (19) developed a leaf disease detection system using AlexNet, achieving high accuracy. Comparisons with other CNN variants like ResNet50 and VGG16 showed AlexNet performed best.
(13)	Fruit	None	Multiple - ResNet50, ResNet18, InceptionV3	95.26, 98.37, 97	Wakhare et al. (20) compared 3 CNN based models. These models were pretrained on the ImageNet dataset. Authors have mentioned the “Black Box” problem of Deep Learning models, which can be solved by integrating these models with the Grad-CAM model.
(14)	Fruit	None	Faster-RCNN	Not Mentioned	Nirgude and Rathi (21) used a variant of RCNN called Faster-RCNN. Authors collected and annotated the images. They also provided the detailed architecture of the model. Drawback is they have not provided proper metrics.
(15)	Fruit	None	CNN-LSTM	98.19	Javeriya (22) combined two deep learning techniques CNN and LSTM to construct a binary classifier. This model has an accuracy of 92.9%. Authors used a dataset of 6519 images. This system can be further developed to multi-class classification.

Sl.No.	Part	Feature Extraction	Classifier	Accuracy	Critical Analysis
(16)	Fruit - Multiple Fruits	Canny SLIC algorithm	Multi-Scale Improved DenseNet	98.98	Al Ansari (23) used the MSRCR based defogging algorithm and Canny SLIC algorithm for image segmentation. Authors improved the DenseNet algorithm by using Focal loss function, focuses more on processing samples, that are difficult or hard for classification in training process.
(17)	Leaf	None	CNN	98.38	This paper provides the detailed architecture of the model. They did not mention any preprocessing step. Qi et al. (24) provide a detailed explanation of CNN, but fails to provide information on other aspects of the system like preprocessing, data collection, metrics, etc.
(18)	Fruit	None	CNN-LSTM	97.1	Vasumathi and Kamarasan (7) proposed a system combining CNN and LSTM with Dragonfly algorithm optimization, trained on a 6,500-image dataset for multiclass classification.
(19)	Leaf	None	CNN based Architectures	Not Mentioned	Vasumathi and Kamarasan (25) proposed a GAN-based data augmentation technique to expand dataset size, which was then used with transfer learning models to develop a classifier.

Vasumathi and Kamarasan (25) extended their work to a multi-class classification using similar hybrid CNN-LSTM model and used dragonfly optimization algorithm for weight optimization. The same csv dataset was used in training of their model and achieved accuracy of 97.1%. Extracting the features manually is a time consuming task, this also limits its use in automated disease detection.

In the study by Naseer, et al. (26) proposed a model to detect growth stages of pomegranate fruit. They utilized a dataset consisting of 5,857 images classified into five growth stages: Bud stage, Early-Fruit stage, Flower stage, Mid-growth stage and Ripe stage. The study utilized Synthetic Minority Over-sampling Technique (SMOTE) is used to address the class imbalance in the dataset. While the proposed method achieved a high accuracy of 98%, it also faced some drawbacks.

Nirmal et al. (27) studied supervised and unsupervised machine learning approaches for detecting diseases in pomegranate leaves. K-means, an unsupervised technique, to isolate the diseased regions of the leaves. The MobileNet model achieved a superior accuracy of 98.18%, while ResNet achieved 95.53% on the Mendeley database. This high resource demand limits the scalability of the approach, making it less feasible for deployment on large, diverse datasets.

3. RESEARCH GAP

The reviewed literature reveals significant advancements in detection and classification of pomegranate diseases using various machine learning and deep learning techniques. Despite the promising results, several research gaps persist. Many studies utilized relatively small datasets, which limits their ability to generalize, leading to overfitting issues. Larger and more diverse datasets are needed to enhance model robustness and performance. This requires collaborative efforts among researchers and agricultural institutions to collect and share extensive, high-quality image datasets. Moreover, leveraging data augmentation techniques can help mitigate the issue of limited datasets by artificially increasing the size, quality and variability of training data.

Traditional machine learning approaches like SVM and ANN have shown good accuracy with limited datasets, but training these models on large, diverse datasets can be computationally intensive. These approaches also require complex preprocessing techniques; without preprocessing, models deliver poor accuracy and often increase training time. Significant preprocessing, including cropping, noise reduction, and background removal, helps models to focus on the relevant or suitable parts of the fruit rather than the noise present in images. However, some research papers provide vague information on preprocessing techniques, limiting the replication of exact models. To address this, future research should include detailed descriptions of preprocessing steps and consider standardizing these procedures to facilitate replication and comparison across studies.

Many studies have used single deep learning approaches, which, while effective, are time-consuming to implement and train, especially for large models like AlexNet and ResNet. Combining multiple deep learning approaches, such as CNN and LSTM, can potentially reduce model size and training time while maintaining or even enhancing performance. For example, CNNs can be used for feature extraction, and LSTMs can process the extracted features to capture temporal dependencies, if any.

Furthermore, the robustness of these models to varying environmental conditions, such as lighting, occlusions, and background clutter, has not been

explicitly addressed, which can significantly impact the accuracy of disease detection in real-world scenarios. Future research should focus on developing models that are resilient to such variations. This can be achieved by training models on datasets that include a wide range of environmental conditions or by employing techniques like data augmentation to simulate these conditions during training.

Finally, some models only provide binary classification, limiting their use in applications where multi-class classification is more appropriate. Developing models that can handle multi-class classification is crucial for practical applications. This involves not only designing appropriate network architectures but also ensuring that the training datasets are sufficiently diverse and well-labeled to cover multiple disease categories.

4. PROPOSED METHODOLOGY

This section discusses the preprocessing of the dataset, developing the model and training the model. For creating model and training model python programming language and many advanced libraries were leveraged. The libraries used are keras, tensorflow, numpy, rembg, pillow.

4. 1. Dataset Preprocessing Pakruddin and Hemavathy (28) have created a standard dataset for training deep learning models which is publicly available. This dataset contains 5099 images divided into 5 classes. The classes are Alternaria, Anthracnose, Bacterial Blight, Cercospora, Healthy. Below Table 2 shows the distribution of images in each class. The dataset contains high quality images collected from areas like Ballari, Bangalore, Bagalkote, etc.

1) Resizing

Fortunately the dataset has all the images in square format. Images are of 3120x3120 pixels. The images were reduced to 512x512 pixels in this step by the resize_images function from the library fastai sublibary vision.

2) Background Removal

Images have fruit in the center and leaves in the background. Removal of background can help the model to focus on the fruit rather than background. It also helps to speed up the machine learning process by reducing the epochs. Figure. 2 shows the output of the process.
For Removing background rembg library was used.

3) Augmentation

Computational techniques like scaling, rotating, shifting were used to generalize the model and reduce overfitting. The ImageDataGenerator class from tensorflow was used to augment only the training dataset. Table 3 shows the parameters used in augmentation.

Figure 2 shows the stages background removal and augmentation of a single image.

4. 2. Proposed Model Vasumathi and Kamarasan (7) proposesd a CNN+LSTM model that utilizes an image dataset for pomegranate fruit disease detection. In this suggested architecture, convolutional neural networks (CNNs) are used here to extract features from input image samples, while long short-term memory networks (LSTMs) utilize these features to model the temporal connections between the data points. This combined approach allows the model to effectively capture both local patterns, such as edges and corners, and global patterns, such as objects and scenes. Moreover, by incorporating LSTMs, the model is capable of understanding the temporal dynamics and dependencies that may exist between different parts of the image over time.

This integrated methodology enables the model to comprehensively understand the data and produce highly accurate predictions. The model can provide precise and reliable classification results by combining CNNs' spatial feature extraction strengths with LSTMs' temporal

TABLE 2. Curated Collection of Labeled Mendeley Dataset Details

Sl. No	Class Name	Image Quantity
1	Alternaria	886
2	Anthracnose	1166
3	Bacterial Blight	966
4	Cercospora	631
5	Healthy	1450

TABLE 3. Different Parameters used in Augmentation

Sl. No	Parameter	Value
1	Rotation range	40
2	Width shift range	0.2
3	Height shift range	0.2
4	Zoom range	0.2
5	Horizontal flip	True
6	Fill mode	Nearest

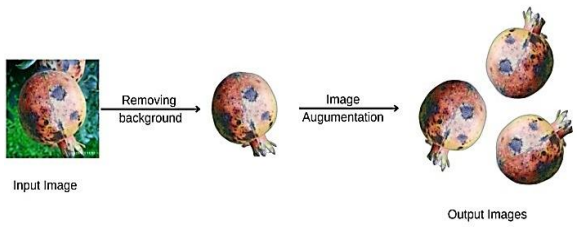


Figure 2. Background Removal and Augmentation

modeling capabilities. Convolutional neural network (CNN) and a long short-term memory (LSTM) network are briefly described as follows.

4. 1. 1. Convolutional Neural Network A simple neural network often lacks the capability to learn complex features required for sophisticated image recognition tasks such as disease classification. Deep learning architectures, such as Convolutional Neural Networks (CNNs), are more suitable for these tasks (23). CNNs are a specific type of multilayer perceptron network that operate on the principle that complex features can be constructed by refining lower-level features at each successive layer (29-31). A CNN typically consists of multiple convolutional layers, multiple pooling layers, and fully connected (FC) layers. An example of a CNN architecture with its layers is illustrated in Figure 3. The *CNN optimal structure* is usually found by optimizing *hyperparameters* like: activation functions, learning rate and fully connected layer size.

Convolution operations apply these filters across the input using a parameter called "stride," which determines the step size for the filter application, ensuring the output dimensions are integer values (32).

Mathematically, the convolution operation is expressed as Equation 1.

$$F(i,j) = (I * K)(i,j) = \sum\sum I(i+m, j+n)K(m,n) \quad (1)$$

where i here is the input matrix, K is a 2D kernel (filter) given by of size m×n, and F is the resulting 2D feature map from the convolution operation denoted by I∗K (33).

To introduce non-linearity into the feature maps, a Rectified Linear Unit (ReLU) layer is used (34). The ReLU activation function computes the output by applying a threshold at zero, which can be mathematically represented as Equation 2.

$$f(x) = max(0, x) \quad (2)$$

The pooling layer, often following convolutional layers, performs downsampling of the input dimensions to reduce the number of parameters and computational load. Max pooling is one of the most common pooling technique used, where the maximum value within a defined region of the input is selected to form the pooled feature map (35). Finally, the fully connected (FC) layer acts as a classifier, making decisions based on features extracted by the convolutional and pooling layers.

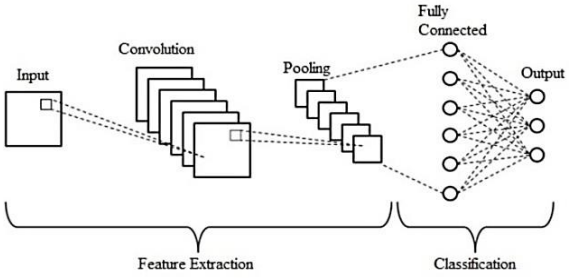


Figure 3. A typical architecture of the CNN

4. 1. 2. Long Short-Term Memory Long Short-Term Memory (LSTM) networks are an advanced type of Recurrent Neural Network (RNN) designed to address the problem vanishing and exploding gradients that are associated with traditional RNNs (36). LSTMs introduce memory blocks that replace conventional RNN units, allowing the network to retain long-term states and effectively connect previous information to current data (37). An LSTM network includes three primary gates: the input gate, forget gate, and output gate, which control the flow of information within the network. Figure 4 shows the internal architecture of LSTM network, depicting how these gates function. Figure 5 illistrates the system architecture of the proposed CNN-LSTM model.

The input gate determines which part of the input to be added to the cell state. This process is defined by the following Equations 3, 4 and 5.

$$i_t = \sigma(W_i \cdot (h_{t-1}, x_t) + b_i) \quad (3)$$

$$\tilde{C}_t = \tanh(W_i \cdot (h_{t-1}, x_t) + b_i) \quad (4)$$

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \quad (5)$$

Here, i_t is calculated by passing the previous hidden state h_{t-1} and the current input x_t through a sigmoid layer to

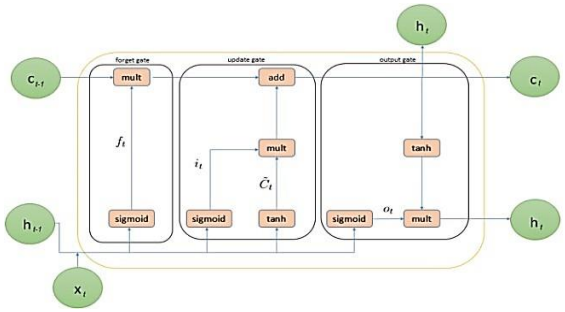


Figure 4. Internal Structure of Long Short-Term Memory

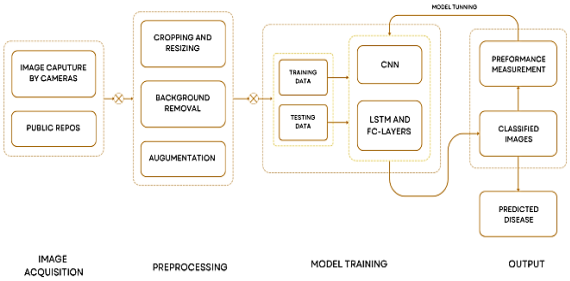


Figure 5. System Architecture of the Proposed CNN-LSTM Model

decide which information to add. The new candidate cell state $\tilde{C}_t$ is obtained by passing h_{t-1} and x_t through a tanh layer. The current cell state C_t is a combination of the old cell state C_{t-1}, modified by the forget gate ft and the new candidate cell state $\tilde{C}_t$.

The forget gate allows selective information passage using a sigmoid function, which can be expressed as Equation 6.

$$f_t = \sigma(W_f \cdot (h_{t-1}, x_t) + b_f) \quad (6)$$

The output gate determines the necessary states for continuation using the inputs h_{t-1} and x_t, as shown in the following Equations 7 and 8.

$$O_t = \sigma(W_o \cdot (h_{t-1}, x_t) + b_o) \quad (7)$$

$$ht = O_t \tanh(C_t) \quad (8)$$

where O_t is the output gate's activation, and h_t is the final hidden state output. W_i, W_f, and W_o represent weight matrices, while b_i, b_f, and b_o are the biases associated with the respective gates.

4. 1. 3. Combined CNN-LSTM Model The convolutional layers here are used to extract spatial features from images. Each layer applies a set of kernels (filters) to the input image, performing an element wise multiplications and summations. This operation is called as convolution which highlights specific features in the input image. The first convolutional layer (Conv2D) is having 32 kernels with a kernel size of 3x3 which uses the ReLU activation function. This layer processes the input images of shape (256, 256, 3). All subsequent convolutional layers all use ReLU activation function with each layer doubling the number of kernels. Number of kernels used in each convolution layer is given in Table 4.

After the convolutional layers, the model incorporates a Long Short-Term Memory (LSTM) layer which captures temporal dependencies and sequence information. This information is useful for understanding the progression of disease symptoms over sequences of image features. Proposed model uses 128 LSTM units, the input to this layer is reshaped by a reshape layer. The reshape layer preceding the LSTM layer ensures that the input is in the correct format, transforming the 2D feature maps into a tensor suitable for LSTM processing. This transformation is critical for leveraging the LSTM's ability to process sequential data, as it enables the model to understand the progression of disease symptoms over a series of images.

The dense layers, also called as Fully Connected (FC) layers, perform the final classification based on the features extracted and processed by the preceding layers. The first dense layer in this model has 1024 units with ReLU activation function, which will introduce non-linearity. The subsequent dropout layer has a dropout value of 0.2 which helps in preventing overfitting by randomly setting a fraction of input units to 0 during training. The final dense layer has 5 units one for each class with softmax activation. The last layer does the classification of input into one of the five disease categories. Detailed Layer information of proposed model is given Table 4.

The model training utilizes 60% of the dataset, with 20% used for validation and the remaining 20% for testing. Training is performed with a batch size of 32 for 210 epochs. The model achieved a training accuracy of

TABLE 4. Layer Information of Proposed CNN-LSTM Model

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 254, 254, 32)	896
max_pooling2d (MaxPooling2D)	(None, 127, 127, 32)	0
conv2d_1 (Conv2D)	(None, 125, 125, 16)	4624
max_pooling2d_1 (MaxPooling 2D)	(None, 62, 62, 16)	0
conv2d_2 (Conv2D)	(None, 60, 60, 32)	4640
max_pooling2d_2 (MaxPooling 2D)	(None, 30, 30, 32)	0
conv2d_3 (Conv2D)	(None, 28, 28, 64)	18496
max_pooling2d_3 (MaxPooling 2D)	(None, 14, 14, 64)	0
conv2d_4 (Conv2D)	(None, 12, 12, 128)	73856
max_pooling2d_4 (MaxPooling 2D)	(None, 6, 6, 128)	0
reshape (Reshape)	(None, 36, 128)	0
lstm (LSTM)	(None, 36, 256)	394240
flatten (Flatten)	(None, 9216)	0
dense (Dense)	(None, 1024)	9438208
dropout (Dropout)	(None, 1024)	0
dense_1 (Dense)	(None, 5)	5125

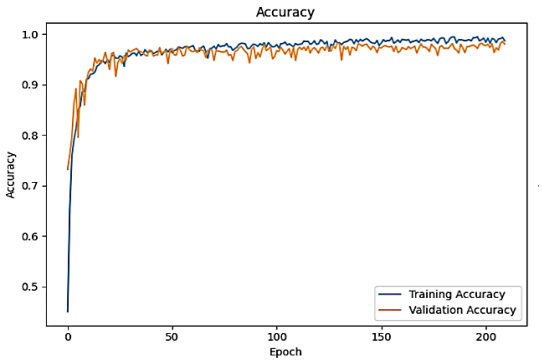


Figure 6. Training and Validation Accuracy through each Epoch

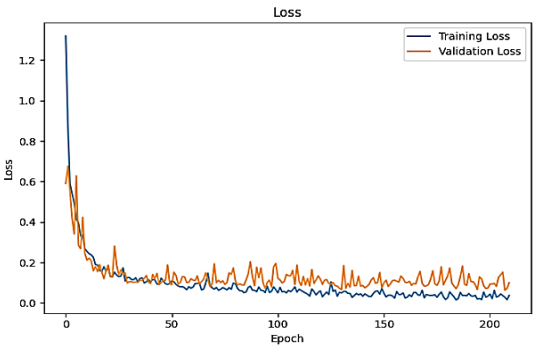


Figure 7. Training and Validation Loss through each Epoch

98.82% with a loss of 0.0430 and a validation accuracy of 98.14% with a loss of 0.0535. The training and validation accuracy and loss for each epoch are depicted in Figures 6 and 7.

4. 3. Other Models Along with the proposed model 4 more models were trained for comparing results. Additional models VGG16, Alexnet, ResNet50. Same data with same ratio was used to train these other models

1) Alexnet

AlexNet marked a significant milestone in the realm of deep learning. Its architecture revolutionized the computer vision field by demonstrating the power of Convolutional Neural Networks (CNNs) on large-scale image classification tasks. With five convolutional layers followed by three dense layers, it achieved remarkable accuracy, winning the ILSVRC in 2012 (38). Table 5 shows the summary of the alexnet model.

This model achieved 99.61% training accuracy 93.03% validation accuracy. The training and validation accuracy and loss for each epoch are shown in Figures 8 and 9.

2) VGG16

The VGG16 model, introduced by the Visual Geometry Group(VGG), University of Oxford, represents another significant advancement in field deep learning, particularly for tasks related to image classification (38). With its 16-layer design, comprising 13 convolutional layers and then 3 Fully Connected (FC) layers, VGG16 demonstrated impressive capabilities in feature extraction and classification.

In many instances, researchers adopted transfer learning techniques, where pretrained VGG16 models were fine-tuned for specific tasks by freezing the weights of the convolutional layers and adding new fully connected layers for task-specific classification (39).

Table 6 shows a summary of the VGG16 model used in this study. It outlines the different layers and how they are organized.

This model achieved 99.80% training accuracy and 95.88 %validation accuracy. The training and validation accuracy for each epoch are shown in Figures 10 and 11.

TABLE 5. Layer Information of AlexNet

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 55, 55, 96)	34944
batch_normalization (Batch Normalization)	(None, 55, 55, 96)	384
max_pooling2d (MaxPooling2D)	(None, 27, 27, 96)	0
conv2d_1 (Conv2D)	(None, 27, 27, 256)	2973952
batch_normalization_1(Batch Normalization)	(None, 27, 27, 256)	1024
conv2d_2 (Conv2D)	(None, 27, 27, 384)	885120
batch_normalization_2 (Batch Normalization)	(None, 27, 27, 384)	1536
conv2d_3 (Conv2D)	(None, 27, 27, 384)	1327488
batch_normalization_3 (Batch Normalization)	(None, 27, 27, 384)	1536
conv2d_4 (Conv2D)	(None, 27, 27, 256)	884992
batch_normalization_4(Batch Normalization)	(None, 27, 27, 256)	1024
max_pooling2d_1 (MaxPooling2D)	(None, 13, 13, 256)	0
flatten (Flatten)	(None, 43264)	0
dense (Dense)	(None, 4096)	177213440
dense_1 (Dense)	(None, 4096)	16781312
dense_2 (Dense)	(None, 5)	20485

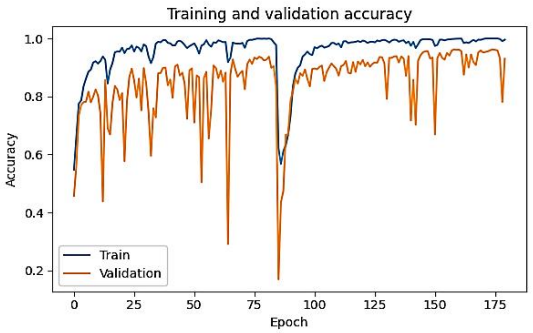


Figure 8. AlexNet Model Accuracy through each Epoch

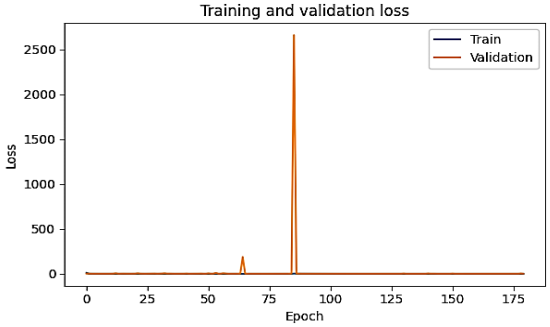


Figure 9. AlexNet Model Loss through each Epoch

3) ResNet50

ResNet50, an advancement of the ResNet architecture pioneered by Microsoft Research, exemplifies the ongoing refinement of deep learning frameworks (38). Distinguished by its 50-layer structure, ResNet50 presents a heightened level of complexity, enabling nuanced feature extraction across diverse datasets and domains.

In the conducted research, a customized version of ResNet50 was utilized, integrating four supplementary layers before the training phase. This augmentation aimed to further enhance the model's capacity to extract and leverage hierarchical features from input data, thereby potentially improving its performance on specific tasks. Table 7 shows the model summary of the ResNet50 model.

This model achieved 99.08% training accuracy and 74.88% validation accuracy. The training and validation accuracy for each epoch are shown in Figures 12 and 13.

TABLE 6. Layer Information of VGG16

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	((None, 224, 224, 3))	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten_1 (Flatten)	(None, 25088)	0
dense_1 (Dense)	(None, 10)	250890

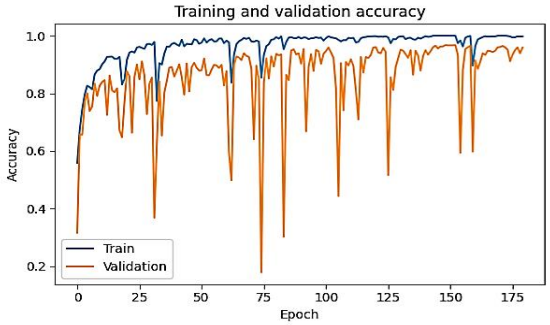


Figure 10. VGG16 Model Accuracy through each Epoch

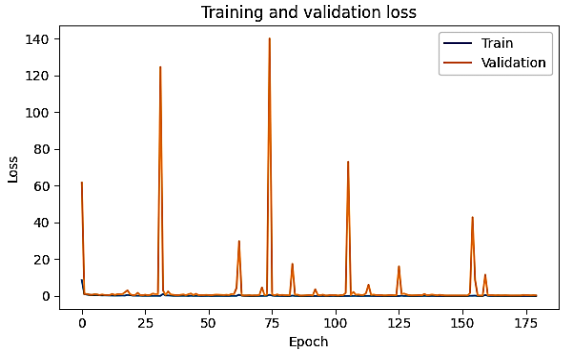


Figure 11. VGG16 Model Loss through each Epoch

5. RESULTS AND DISCUSSION

The proposed CNN+LSTM hybrid model was compared to various models as discussed above using some of the important performance metrics. Metrics used are Classification Accuracy, Precision, Recall, F1-Score and Specificity.

TABLE 7. Layer Information of ResNet50

Layer (type)	Output Shape	Param #
resnet50 (Functional)	(None, 7, 7, 2048)	23587712
global_average_pooling2d (GlobalAveragePooling2D)	(None, 2048)	0
batch_normalization (Batch Normalization)	(None, 2048)	8192
dense_2 (Dense)	(None, 256)	524544
batch_normalization_1 (BatchNormalization)	(None, 256)	1024
dense_3 (Dense)	(None, 128)	32896
dropout (Dropout)	(None, 128)	0
batch_normalization_2 (BatchNormalization)	(None, 128)	512
dense_4 (Dense)	(None, 64)	8256
dropout_1 (Dropout)	(None, 64)	0
batch_normalization_3 (BatchNormalization)	(None, 64)	256
dense_5 (Dense)	(None, 5)	325

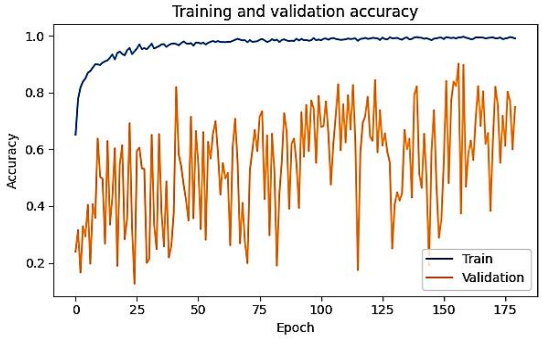


Figure 12. ResNet50 ModelAccuracy through each Epoch

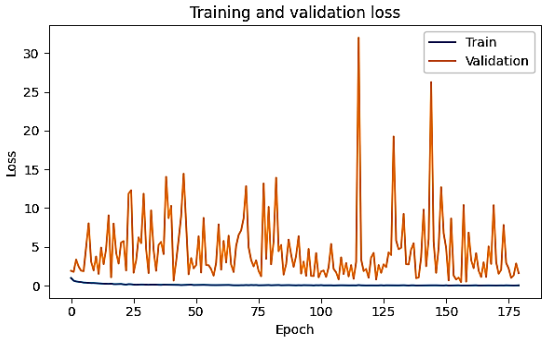


Figure 13. ResNet50 Model Loss through each Epoch

5. 1. Classification Accuracy This is one of the critical metrics for measuring performance. It refers to the percentage of correct predictions made by the classification model. Equation 9 is used for the calculation of this metric.

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ Predictions} \quad (9)$$

Figure 14 shows the accuracies of all trained models. The proposed model achieved the highest test accuracy of 98.8% with a test loss of 0.804. Resnet50 performed the worst among all the models.

5. 2. Precision Precision is the ratio of correct predictions(true positives) to total predictions. Equation 10 is used for the calculation of precision.

$$Precision = \frac{True\ Positives}{(True\ Positives + False\ Positives)} \quad (10)$$

Figure 15 shows the proposed model achieved precision of 0.99 which is highest compared to other models.

5. 3. Recall Recall is also known as sensitivity. It measures the ratio of true positives correctly classified as Positive to the total number of positive predictions. Equation 11 is used for the calculation of this metric.

$$Recall = \frac{True\ Positives}{(True\ Positives + False\ Negatives)} \quad (11)$$

Figure 16 shows the recall values of all trained models. The proposed model has the achieved highest recall value.

5. 4. F1 Score F1 Score is a harmonic mean of precision and recall. Equation 12 is used for the calculation of this metric.

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (12)$$

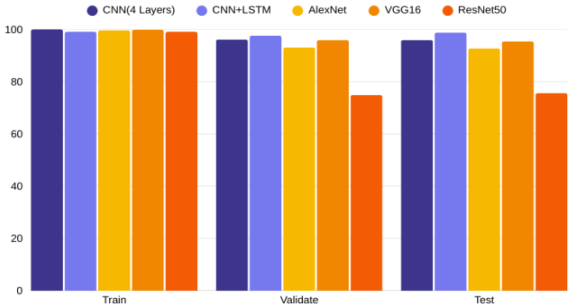


Figure 14. Test Accuracy of all Trained Models

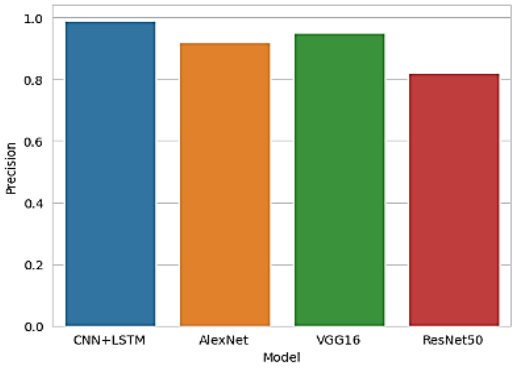


Figure 15. Highest Precision of CNN LSTM Model

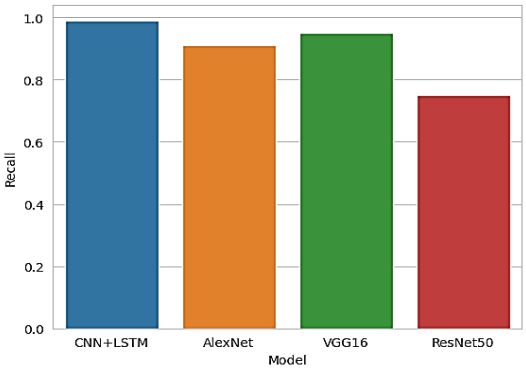


Figure 16. Recall of all Trained Models

Figure 17 shows the F1 Scores of all trained models. It shows that the proposed model has outperformed all other models.

5. 5. Confusion Matrix Confusion Matrix provides a detailed breakdown of the actual versus predicted classification. Figure 18 shows the confusion matrix of Proposed Model.

From the above metrics, it can be observed that the proposed CNN+LSTM model outperformed the other CNN based model Alexnet and state-of-art transfer learning models ResNet and VGG16. Also it is noted that the proposed model uses fewer layers (CNN as well as fully connected) compared to other models.

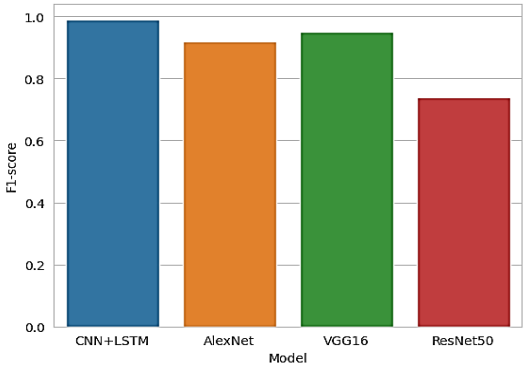


Figure 17. F1-Score of all Trained Models

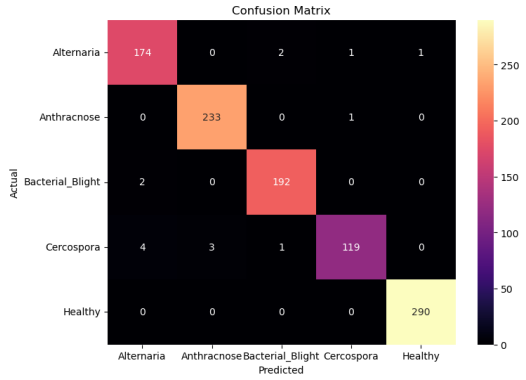


Figure 18. CNN+LSTM Model outperformed the other CNN based model

6. CONCLUSION

This study proposed a novel approach for detecting pomegranate fruit disease detection by using a hybrid model. By combining the feature extraction of CNN with sequential handling of LSTM, the proposed model demonstrates higher performance compared to other models that were considered for this study.

Experiment results show that the proposed model CNN-LSTM archives higher accuracy and precision, significantly improving the disease detection and classification capabilities. This advancement can help the agriculture and food industry in reducing economic costs.The study provides a comprehensive evaluation of various machine learning models, highlighting the efficacy of hybrid models in agricultural disease detection. Similar approaches could benefit other crops which offer scalable and adaptable solutions.

However there are limitations also exist. The dataset, while extensive, is region-specific and limited to few varieties of pomegranate fruit. Expanding the dataset to include other varieties and other locations can increase the robustness of the model. Additionally, integrating advanced image preprocessing techniques and exploring other hybrid combinations may yield better results.

The study also adds the ineffectiveness of transfer learning approaches like VGG16 and ResNet50 and the performance issues of larger models like AlexNet indicating that larger architectures may not always perform better in this context.

In conclusion, the CNN-LSTM hybrid model marks a significant advancement in agricultural disease detection, offering a promising tool for improving the sustainability and productivity of pomegranate cultivation and potentially revolutionizing agricultural practices.

Future Work:The following are some potential avenues for further research, building on the encouraging findings of this study: investigation of further hybrid architectures, such as CNN-GRU, Transformer-based networks, or lightweight attention-based architectures, for use in real-time disease detection in the field using mobile devices or drones.

7. REFERENCES

1. Khan N, Fahad S, Naushad M, Faisal S. Pomegrantes economics and medicinal aspects in the world. Available at SSRN 3597891. 2020. 10.2139/ssrn.3597891
2. Mena P, Gironés-Vilaplana A, Moreno DA, García-Viguera C. Pomegranate fruit for health promotion: Myths and realities. Funct Plant Sci Biotechnol. 2011;5:33-42. http://globalsciencebooks.info/Online/GSBOnline/images/2011/FPSB_5(SI2)/FPSB_5(SI2)33-42o.pdf
3. Ismail T, Sestili P, Akhtar S. Pomegranate peel and fruit extracts: a review of potential anti-inflammatory and anti-infective effects. Journal of ethnopharmacology. 2012;143(2):397-405. https://doi.org/10.1016/j.jep.2012.07.004
4. Ganeshkumar D, Bhattacharyya K, Narsimhaiah L, Yumnam SS. Performance of pomegranate export from India. Economic Affairs. 2016;61(4):575-80. https://www.researchgate.net/publication/329871322
5. Hosamani S, Alur A, editors. Pomegranate-an Indian scenario. III International Symposium on Pomegranate and Minor Mediterranean Fruits 1089; 2013.

6. Heber D. Pomegranate ellagitannins. Herbal Medicine: Biomolecular and Clinical Aspects 2nd edition. 2011. https://pubmed.ncbi.nlm.nih.gov/22593938/
7. Vasumathi M, Kamarasan M. An effective pomegranate fruit classification based on CNN-LSTM deep learning models. Indian Journal of Science and Technology. 2021;14(16):1310-9. https://doi.org/10.17485/IJST/v14i16.432
8. Deshpande T, Sengupta S, Raghuvanshi K. Grading & identification of disease in pomegranate leaf and fruit. International Journal of Computer Science and Information Technologies. 2014;5(3):4638-45. 10.13140/RG.2.2.12639.84643
9. Bhange M, Hingoliwala H. Smart farming: Pomegranate disease detection using image processing. Procedia computer science. 2015;58:280-8. 10.1016/j.procs.2015.08.022
10. Gaikwad SA, Deore KS, Waykar MK, Dudhane PR, Sorate G. Fruit disease detection and classification. International Research Journal of Engineering and Technology. 2017;4:1151-4.
11. Kantale P, Thakare S. Pomegranate disease classification using Ada-Boost ensemble algorithm. Int J Eng Res Technol. 2020;9:612-20. 10.17577/IJERTV9IS090216
12. Dhakate M, Ingole A, editors. Diagnosis of pomegranate plant diseases using neural network. 2015 fifth national conference on computer vision, pattern recognition, image processing and graphics (NCVPRIPG); 2015: IEEE. 10.1109/NCVPRIPG.2015.7490056
13. Patil J. Pomegranate fruit diseases detection using image processing techniques: a review. Information Technology in Industry. 2021;9(2):115-20. https://doi.org/10.17762/itii.v9i2.310
14. Pawar R, Jadhav A, editors. Pomogranite disease detection and classification. 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI); 2017: IEEE. 10.1109/ICPCSI.2017.8392162
15. G. Nagarjuna Reddy LG, P. Eshwar, P. Sarvothama Reddy , K. Madhu Sudhan Reddy. ANN based Fruit Disease Detection using Texture Analysis. International Journal of Emerging Technologies and Innovative Research. 2023;10(4):e187-e93. http://www.jetir.org/papers/JETIR2304425.pdf
16. Madhavan MV, Thanh DNH, Khamparia A, Pande S, Malik R, Gupta D. Recognition and classification of pomegranate leaves diseases by image processing and machine learning techniques. Computers, Materials & Continua. 2021;66(3):2939-55.
17. Sharath D, Kumar SA, Rohan M, Suresh K, Prathap C, editors. Disease detection in plants using convolutional neural network. 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT); 2020: IEEE. 10.1109/ICSSIT48917.2020.9214159
18. Sánchez MG, Miramontes-Varo V, Chocoteco JA, Vidal V, editors. Identification and classification of botrytis disease in pomegranate with machine learning. Science and Information Conference; 2020: Springer.
19. Pawar S, Shedge S, Panigrahi N, Jyoti A, Thorave P, Sayyad S, editors. Leaf disease detection of multiple plants using deep learning. 2022 international conference on machine learning, big data, cloud and parallel computing (COM-IT-CON); 2022: IEEE. 10.1109/COM-IT-CON54601.2022.9850899
20. Wakhare PB, Kandalkar JA, Kawtikwar RS, Kalme SA, Patil RV. Development of automated leaf disease detection in pomegranate using Alexnet algorithm. Curr Agric Res J. 2023;11(1):177-85. http://dx.doi.org/10.12944/CARJ.11.1.15
21. Nirgude V, Rathi S. A robust deep learning approach to enhance the accuracy of pomegranate fruit disease detection under real field condition. Journal of Experimental Biology and Agricultural Sciences. 2021;9(6):863-70. 10.18006/2021.9(6).863.870
22. Javeriya S. Faster-RCNN based deep learning model for pomegranate diseases detection and classification. LC International Journal of STEM (ISSN: 2708-7123). 2021;2(3):114-20. 10.36647/CIML/02.02.A002
23. Al Ansari MS. A Machine Learning Approach to Pomegranate Leaf Disease Identification. International Journal on Recent and Innovation Trends in Computing and Communication ISSN. 2023:2321-8169. 10.17762/ijritcc.v11i9.9597
24. Qi Y, Zhichao C, Lan L, editors. Fruit Disease Identification Based on Improved Densenet Fusion Defogging Algorithm. 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP); 2021: IEEE. 10.1109/ICCWAMTIP53232.2021.9674088
25. Vasumathi M, Kamarasan M. An lstm based cnn model for pomegranate fruit classification with weight optimization using dragonfly technique. Indian J Comput Sci Eng. 2021;12(2):371-84. 10.21817/indjcse/2021/v12i2/211202051
26. Naseer A, Amjad M, Raza A, Munir K, Samee NA, Alohali MA. A novel transfer learning approach for detection of pomegranates growth stages. IEEE access. 2024;12:27073-87. 10.1109/ACCESS.2024.3365356
27. Nirmal M, Jadhav PP, Pawar S. Pomegranate leaf disease detection using supervised and unsupervised algorithm techniques. Cybernetics and Systems. 2023:1-12. 10.1080/01969722.2023.2166192
28. Pakruddin B, Hemavathy R. Development of a Pomegranate Fruit Disease Detection and Classification Model using Deep Learning. Indian Journal of Agricultural Research. 2024;58. 10.17632/b6s2rkpmvh.1
29. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural networks. Pattern recognition. 2018;77:354-77. 10.1016/j.patcog.2017.10.013
30. Hochreiter S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. 1998;6(02):107-16. 10.1142/S0218488598000094
31. Schmidhuber J. Deep learning in neural networks: An overview. Neural networks. 2015;61:85-117. 10.1016/j.neunet.2014.09.003
32. Hassan SM, Maji AK. Plant disease identification using a novel convolutional neural network. IEEE access. 2022;10:5390-401. 10.1109/ACCESS.2022.3141371
33. Devi EA, Gopi S, Padmavathi U, Arumugam SR, Premnath S, Muralitharan D, editors. Plant disease classification using CNN-LSTM techniques. 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT); 2023: IEEE. 10.1109/ICSSIT55814.2023.10061003
34. Shafik W, Tufail A, Liyanage CDS, Apong RAAHM. Using a novel convolutional neural network for plant pests detection and disease classification. Journal of the Science of Food and Agriculture. 2023;103(12):5849-61. https://doi.org/10.1002/jsfa.12700
35. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Communications of the ACM. 2017;60(6):84-90. 10.1145/3065386
36. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014. https://www.researchgate.net/publication/265385906
37. Yan Q, Yang B, Wang W, Wang B, Chen P, Zhang J. Apple leaf diseases recognition based on an improved convolutional neural network. Sensors. 2020;20(12):3535. https://doi.org/10.3390/s20123535
38. He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. 10.1109/CVPR.2016.90

39. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, et al. Backpropagation applied to handwritten zip code recognition. Neural computation. 1989;1(4):541-51. 10.1162/neco.1989.1.4.541

COPYRIGHTS

©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

انار میوه‌ای با ارزش بالا است که به دلیل فواید تغذیه‌ای و کاربردهایش در طب سنتی و لوازم آرایشی، در سطح جهانی شناخته شده است. هند بازیگر کلیدی در بازار جهانی انار است، اما این صنعت با چالش‌هایی مانند بیماری‌هایی که بر بهره‌وری محصول تأثیر می‌گذارند و ضررهای اقتصادی برای کشاورزان روبرو است. این مطالعه رویکرد جدیدی را برای تشخیص بیماری‌های انار با استفاده از یک مدل ترکیبی شبکه عصبی کانولوشن (CNN)و حافظه کوتاه‌مدت بلند (LSTM)ارائه می‌دهد. مدل پیشنهادی از CNNها برای استخراج ویژگی مؤثر و از LSTMها برای مدیریت داده‌های متوالی استفاده می‌کند و در مقایسه با روش‌های سنتی و سایر تکنیک‌های یادگیری عمیق، به عملکرد برتر دست می‌یابد. نتایج تجربی، دقت، یادآوری، دقت و امتیاز F1 بالایی را نشان می‌دهد. مدل پیشنهادی به دقت ۹۸.۵۳٪ و تلفات ۰.۰۶۷۷ دست یافت. این مطالعه همچنین محدودیت‌های رویکردهای یادگیری انتقالی مانند VGG16 و ResNet50 و مدل‌های بزرگ‌تر مانند AlexNet را که در این زمینه عملکرد خوبی نداشتند، بررسی می‌کند. یافته‌ها نشان می‌دهد که مدل ترکیبی CNN-LSTM یک راه‌حل مقیاس‌پذیر و سازگار برای تشخیص بیماری‌های کشاورزی با کاربردهای بالقوه برای محصولات مختلف ارائه می‌دهد.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1176-1190

International Journal of Engineering

Journal Homepage: www.ije.ir

Comparative Performance Evaluation of Savonius Vertical Axis Wind Turbines: Two and Three-bladed Profiles

A. Sadeghi[a,b], T. B. Gorji*[a], M. Gorji-bandpy[c]

[a] *Department of Mechanical Engineering, Mazandaran Institution of Technology, Babol, Iran*
[b] *Aerodynamic Laboratory of Mechanical Engineering, Babol Noshirvani University of Technology, Babol, Iran*
[c] *Faculty of Mechanical Engineering, Babol Noshirvani University of Technology, Babol, Iran*

PAPER INFO

Paper history:
Received 08 April 2025
Received in revised form 29 May 2025
Accepted 15 June 2025

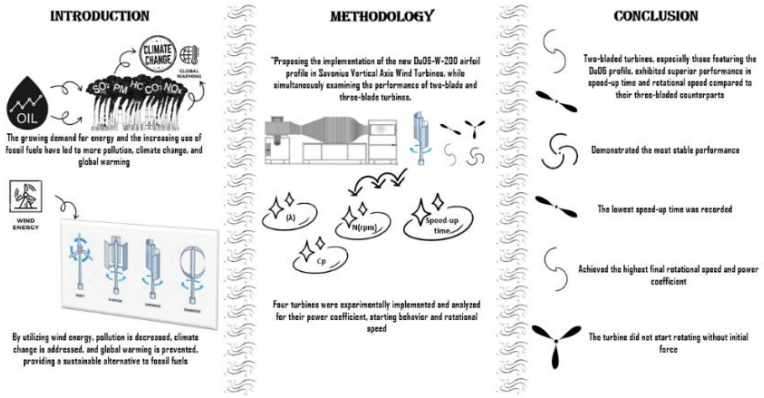
Keywords:
Wind Turbine
Semi-circular
Du06-W-200 Profile
Power Coefficient
Speed-up Time
Rotational Speed

ABSTRACT

In recent decades, the environmental impact of energy production has led to growing global efforts to transition from polluting fossil fuels to renewable energy sources, attracting considerable attention from organizations and researchers. This transition is essential due to escalating environmental concerns and the limited availability of fossil fuels. Concurrently, the deployment of wind turbines worldwide has expanded significantly. This study aims to propose an optimal Savonius vertical axis wind turbine design suited for urban applications. Two turbine models featuring different blade numbers and profiles were designed, constructed, and experimentally tested. The turbines comprised double- and triple-blade configurations, utilizing semicircular and Du06-W-200 blade profiles. Performance evaluation focused on key parameters including power coefficient, final rotational speed, power output, tip speed ratio, and time to reach maximum rotational speed. The results demonstrated that the semicircular double-blade turbine delivered the highest final rotational speed of 1433 RPM, approximately 33% higher than the three-bladed semicircular turbine and over 160% higher than the Du06-profiled turbines. It also achieved the highest power coefficient among all turbines tested. The two-blade turbine (Du06-W-200) profile exhibited the fastest acceleration, reaching its final speed in only 11 seconds, twice as fast as the two-bladed semicircular turbine and three times faster than the three-bladed semicircular configuration. While the Du06-profiled turbines showed slightly lower tip speed ratios, the three-bladed semicircular turbine exhibited the most stable power output. These findings indicate that Savonius turbines with the tested configurations are well-suited for urban environments, though their effectiveness depends strongly on local geographical and climatic conditions.

***doi**: 10.5829/ije.2026.39.05b.11*

Graphical Abstract



*Corresponding Author Email: gorji.tahereh@mit.ac.ir (T. B. Gorji)

Please cite this article as: Sadeghi A, Gorji TB, Gorji-bandpy M. Comparative Performance Evaluation of Savonius Vertical Axis Wind Turbines: Two and Three-bladed Profiles. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1176-1190.

NOMENCLATURE			
A	Cross-sectional area (m^2)	P_{den}	Power density (W)
C_P	Power coefficient	R	Radius of the blade (m)
E_K	Kenetic energy (J)	t	Time (s)
K_m	Mass flow rate (kg/s)	v	Velocity (m/s)
m	Air mass (kg)	**Greek Symbols**	
N	Rotational speed (rpm)	λ	Linear speed of the blade tip to the wind speed
P	Power (W)	ρ	Air density (kg/m^3)
P_0	Wind power (W)	ω	Angular velocity (rad/s)

1. INTRODUCTION

Climate change is driving the world to seek cleaner energy sources. Renewable energy has become increasingly important as efforts intensify to move away from fossil fuels (1-3). This shift has created strong demand for improved technologies that can capture or generate energy more efficiently, contributing to the development of a cleaner future. Multiple factors are driving this transition to renewable energy. Scientific studies confirm the need to alleviate the adverse environmental impacts associated with fossil fuel–based power generation. Fossil fuels emit the most greenhouse gases, which have the most detrimental effects on the environment (4-6). Furthermore, hormonal activity can be disrupted by various substances originating from fossil fuels, which can be found in food, air, water, and manufactured products (7, 8). additionally, the finite nature of fossil fuel reserves and concerns over energy security have increased focus on renewable energy. By harnessing the power of nature, renewable energy offers a long-term solution that is not subject to the same resource limitations as fossil fuels. Moreover, renewable energy can be generated locally, reducing dependence on imported energy sources and enhancing energy self-sufficiency (9, 10).

Wind turbines are now widely recognized as symbols of renewable energy and are becoming increasingly significant in the global shift toward a sustainable future (11, 12). By harnessing wind power, these impressive structures are capable of generating clean and abundant electricity without the greenhouse gas emissions associated with traditional energy sources. While wind turbines are implemented in various designs (13), the principal aim of wind energy conversion is the efficient production of electrical power. The variety of wind turbine designs reflects the industry's evolving requirements and ongoing advancements in wind energy technology. The incorporation of renewable energy sources in the manufacturing sector is crucial for enhancing sustainability and reducing carbon emissions. However, obstacles such as high costs, limited technology, and the need for skilled workers remain. Policy frameworks, financial incentives, and technological advancements are key to facilitating this transition. The potential for cost savings, increased energy efficiency, and improved business reputation present a compelling case for this change, despite considerable challenges. The sector can overcome these barriers and achieve global sustainability goals through supportive regulations, continuous innovation, and stakeholder engagement. For broader adoption, further investigation into scalable, industry-specific solutions is required (14, 15).

Different types of wind turbines have emerged to optimize performance, efficiency, and adaptability under varying wind conditions and geographical settings. Each design has its own set of advantages, limitations, and applications, catering to specific requirements and maximizing the potential of wind power (16).

Traditionally, there are two primary types of wind turbines: vertical axis wind turbines (VAWTs) and horizontal axis wind turbines (HAWTs) (17). AWTs, which are among the most common and recognizable wind turbine designs, feature a horizontal rotor shaft oriented in line with the wind direction. Typically consisting of three blades, HAWTs are renowned for their high energy conversion efficiency and ability to generate substantial power. They are widely employed in large-scale wind farms and marine-based projects due to their scalability and cost-effectiveness (18). AWTs, on the other hand, have a rotor shaft positioned vertically. VAWTs can operate regardless of wind direction, making them versatile and suitable for diverse wind conditions. They are often utilized in urban and small-scale applications where space constraints and aesthetic considerations are important (19, 20). Within the category of VAWTs, the Savonius wind turbine is notable for its distinctive S-shaped blades and suitability for low wind speeds and turbulent conditions (21). The Darrieus type features vertically oriented airfoil-shaped blades that provide high torque at low rotational speeds (22). The Savonius wind turbine has attracted significant attention due to its unique design and potential applications. Named after its inventor, Finnish engineer Sigurd J. Savonius, these turbines offer specific advantages that distinguish them from other wind turbine designs (23). The distinguishing feature of Savonius wind turbines lies in their S-shaped blades, which rotate around a vertically oriented axis. This design allows

Savonius turbines to operate effectively in a variety of wind conditions, including low wind speeds and turbulent environments. Consequently, they offer promising possibilities for utilizing wind energy in urban and small-scale settings, where traditional horizontal-axis wind turbines may face limitations. One of the key advantages of Savonius wind turbines is their simplicity and solidarity. Their design comprises curved blades that capture wind energy by employing the drag force rather than lift. This simplicity of construction makes Savonius turbines relatively easy to manufacture, install, and maintain, thereby enhancing their cost-effectiveness and accessibility. Additionally, Savonius wind turbines exhibit excellent self-starting capabilities, allowing them to initiate rotation even at low wind speeds. This feature ensures continuous power generation, making them suitable for regions with moderate or unpredictable wind conditions.

Various researches have been conducted to improve the performance of Savonius wind turbines. The two primary areas of literature research on increasing Savonius turbine efficiency are exterior additional design and interior structural design as shown in Figure 1 (24).

Shame et al. (8) numerically analyzed seasonal wind energy potential on Zanzibar's coastal island using half-hourly wind data collected at 10 m height. A Weibull distribution model, fitted using the standard deviation method, was employed to evaluate wind speed, wind direction, and wind energy density (wED). The analysis indicated the presence of strong winds (4.5–7 m/s) persisting for approximately 15 hours per day, with the highest wED (23.3 GWh/m²) observed during the winter season. The POLARIS P62-1000 wind turbine was identified as the most suitable option, based on its high capacity factor (>58%) and power output (589 kW), thereby confirming the region's viability for wind farm development. Bozsik et al. (6) simulated the impact of climate change on household-scale photovoltaic (PV) systems in Hungary using hourly climate projections from the Meteonorm database under RCP2.6, RCP4.5, and RCP8.5 scenarios. A PV system comprising 18 LG monocrystalline panels and a Fronius inverter was modeled in PVsyst software. The study analyzed global horizontal irradiance (GHI), air temperature, and resulting energy yield (En) over the 21st century. Results revealed that while GHI increases in all scenarios, higher temperatures negatively affected PV efficiency, especially under RCP4.5 and RCP8.5. Partial correlation analysis confirmed that temperature rise diminishes power output when isolating the effect of GHI. The findings highlight the need for thermal management in PV systems and recommend further comparative studies across PV technologies and seasonal impacts to improve future projections.

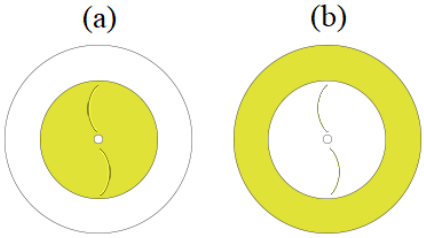


Figure 1. graphical abstract of Savinus Wind Turbine: (a): interior design area (b): exterior design area

The fundamental objective of external design studies is to improve turbine performance by reducing negative torque and increasing positive torque (25). Utilizing concepts derived from the field of exterior design, improvements to Savonius VAWTs through additional external features involve various flow augmentation systems. These include different types of deflectors, guide-box tunnels, shielding structures, conveyors, and other configurations, which have been extensively discussed in several studies (26-30).

The interior design performance of Savonius turbines is influenced by multiple factors, including blade shape and configuration, height and aspect ratio, number of blades, number of stages, twist angle, shaft diameter, installation of end plates, overlap ratio, material, weight, rotor diameter, and many other features (31). Wenehenubun et al. (32) experimentally compared Savonius turbines with different numbers of blades based on tip speed ratio (TSR), torque, and power coefficient relative to wind speed. Additionally, a simulation study was conducted to analyze pressure distribution. The three-bladed Savonius turbine was found to be the most efficient model, achieving the highest tip speed ratio of 0.555 at a wind speed of 7 m/s. Four-bladed turbine configurations generated the greatest actual torque within the wind speed range of 0 to 10 m/s. Gumilar et al. (33) performed a numerical study on an L-shaped Savonius wind turbine with the objective of assessing power capture as a function of wind speed and rotational speed. They also evaluated power relative to blade count. The maximum power output reached 643.51 W at a wind speed of 20 m/s and a rotor rotation speed of 50 rad/s. Two-bladed turbines captured the highest power, particularly at a wind speed of 20 m/s. Morever, Prabowoputra and Prabowo (34) emphasizes a direct relationship between aspect ratio and the performance of this rotor type, with optimal performance observed at overlap ratios between 0 and 0.3, while acknowledging the influence of additional factors. Modified blade shapes demonstrate performance improvements ranging from 8 to 25%. Additionally, the integration of multi-stage configurations enhances rotor performance in accordance with other factors. According to the literature, the optimal blade count for balancing performance and stability is determined to be two or three blades. n artificial neural network was utilized by Al-Shammari et al. (35) to predict the optimal blade shape design, aiming to

increase the power coefficient at low wind speeds within the TSR range of 0.2 to 1.2. Their study demonstrated that the blade-shaped Savonius wind turbine outperforms the semi-circular blade rotor within the TSR range of 0.6 to 1.1. Furthermore, the maximum power coefficient improved by 55% compared to the conventional semi-circular blade-shaped Savonius wind turbine. Tartuferi et al. (36) conducted a numerical study aimed at improving the aerodynamic efficiency of a Savonius wind rotor by utilizing two distinct approaches. The first approach involved designing blades with innovative airfoil shapes, while the second incorporated a novel curtain system capable of autonomously aligning with the wind direction. The results demonstrated that both proposed approaches significantly enhanced the energy performance of Savonius wind turbine. Hosseini Imeni et al. (37) performed a computational analysis of airfoil-bladed Savonius wind turbines to increase efficiency and reduce entropy generation through the incorporation of an innovative airfoil-shaped blade design. Nine key design parameters were identified, and twelve distinct design points were developed for evaluation. The findings revealed that the power coefficient could be increased by up to 12.89% at a TSR of 0.8 compared to conventional semi-circular blades. Shamsuddin and Kamaruddin (38) experimentally compared single-stage and two-stage Savonius wind turbines at flow speeds ranging from 5 to 9 m/s in a wind tunnel. The results indicated that the addition of an extra stage led to a substantial increase in the power coefficient, with a peak improvement of 138% at a wind speed of 5 m/s. Moreover, the two-stage configuration was found to reduce the self-starting speed across all rotor angles, thereby significantly enhancing the turbine's self-starting capability. Ikonwa et al. (39) employed experimental datasets to evaluate various airfoil profiles aimed at improving the self-starting performance of small-scale, fixed-pitch, straight-bladed vertical axis wind turbines. The findings indicated that conventional symmetrical airfoils were unsuitable for self-starting, and no single asymmetrical airfoil was universally optimal. However, two asymmetrical airfoils, DU-06-W-200 and S 1210, exhibited superior aerodynamic performance. Notably, the DU-06-W-200 airfoil demonstrated a higher starting torque, positioning it as a promising solution to mitigate self-starting challenges in small-scale fixed-pitch straight-bladed vertical axis wind turbines. Zamani et al. (40) numerically investigated the flow behavior around vertical axis wind turbine (VAWT) blades. The DU-06-W-200 airfoil was selected for application in Darrieus-type VAWT blades. Aerodynamic coefficients were analyzed and compared with those of conventional straight blades over a broad range of angles of attack. The findings indicated that the DU-06-W-200 airfoil exhibited an increased lift-to-drag coefficient ratio (CL/CD) and improved self-starting capability. A summary of the aforementioned interior design studies is presented in Table 1.

TABLE 1. Summarized information of interior design literature review

Reference	Year	Methodologies	Results or contributions
(24)	2023	Influential Aerodynamic Design	Interior Structure Design Exterior Additional Design
(31)	2023	Interior Structure Design Features	Blade Shape, Aspect Ratio, Number of Blades, Number of Stages etc.
(32)	2015	Compared (2, 3 and 4) Number of Blades Rotors	Three Blade Highest Efficiency Four Blade Highest Torque
(33)	2019	Impacts of Number of Blades	Maximum Power Surpasses Previous Research By Two-bladed Turbine
(34)	2022	Observed Overlap Variation, Blade Shapes, Multiple Stages	The Optimal Number of Blades is Two and Three-Bladed Savonius Turbines
(35)	2020	Optimized the Design of Blade Shape Using ANSYS-CFX Software	Better Performance of Blade-shaped Rotors than Semi-circular Rotor
(36)	2015	Airfoil Shaped Blades and Certain System for Aligning Wind Direction	Sensibly Enhance the Energy Performance of the Savonius Wind Turbine
(37)	2022	Formulated Crucial Parameters of Airfoil-shaped Blade Designs	12.89% Enhancement in Comparison with a Conventional Savonius Rotor
(39)	2016	Examine Different Airfoil Profiles in Self-starting Performance	DU06-W-200 Profile Could be an Ideal Airfoil to Eliminate the Problem
(38)	2023	Self-starting of One and Two Staged Rotors were Discussed	Double-stage Rotor Outperformed In Terms of Self-starting Capabilities
(40)	2021	Observed Flow Behavior Around VAWT Including Du06-W-200 Airfoil	The Airfoil Equipped Improves Self-starting and Cut-in Decrease

The present study primarily investigates the performance of a Savonius turbine with respect to interior design modifications. The main objective is to refine and compare the interior configurations of a Savonius VAWT and to evaluate the corresponding power coefficient of the modified designs. Regarding the number of blades, no universally accepted optimal configuration has been established, as the interior designs of specific rotors may yield varying results depending on factors such as wind speed, rotor size, and intended applications. Some studies have suggested that two-bladed configurations may provide better efficiency (33), whereas other research supports the efficiency of three-bladed rotors (32). Therefore, the choice between two or three blades for a Savonius VAWT depends on specific design considerations and operating conditions. Consequently, it is common for researchers to conduct experiments and simulations to identify the configuration that best addresses the challenges of their particular application.

Furthermore, the aerodynamics of the turbine, especially blade shape, is considered one of the most influential factors affecting the efficiency of Savonius VAWTs (35). Efficiency can be characterized by various turbine attributes such as self-starting capability, power coefficient, and torque coefficient. Comparative studies have been conducted on airfoil-shaped blades versus semi-circular blades under different conditions (36). Additionally, the use of the airfoil profile DU06-W-200 in VAWTs to enhance self-starting capability has been explored in recent years (40). owever, based on the reviewed literature, no experimental investigation has been conducted on the application of the DU06-W-200 airfoil in a Savonius VAWT.

In this study, experiments are conducted on two- and three-bladed configurations of semi-circular and DU06-W-200 profile blades in a Savonius vertical axis wind turbine. The experiments involve measuring the rotational speed, the speed of the exiting wind flow from the turbine, and the time taken for each turbine configuration to reach its ultimate rotational speed. The results are then compared in terms of rotor rotational speed, power coefficient, and startup time. The methodology employed, including governing equations, the detailed experimental setup, and testing procedures, will be discussed in the following chapter, followed by the presentation of results and conclusions.

2. METHODOLOGY

This study employs a systematic methodology to investigate the performance of Savonius VAWTs through a combination of theoretical analysis and experimental validation. The approach is divided into two parts: the governing equations, which form the foundation of the turbine's aerodynamic behavior, and the detailed experimental setup designed to measure its performance metrics. The following sections will first present the governing equations underlying the theoretical framework, followed by a comprehensive description of the experimental setup used to validate the results.

2. 1. Governing Equations Primarily, wind turbines extract the kinetic energy present in the wind and convert it into electrical energy. The governing equations of wind turbines are presented as follows (41). The kinetic energy of a mass of air with mass (m) and velocity (v) is represented by Equation 1.

$$E_k = \frac{1}{2} m v^2 \tag{1}$$

At speed v (m/s), the mass of passing air through the cross-sectional area A (m^2) is expressed in Equation 2 as follows:

$$m = \rho A V t \tag{2}$$

In this equation, ρ represents air density (kg/m^3). Based on Equations 1 and 2, the wind power can be expressed as Equation 3.

$$P_0 = \frac{1}{2} \rho A V^3 \tag{3}$$

Therefore, the specific power or power density can be calculated using to Equation 4.

$$P_{den} = \frac{P_0}{A} = \frac{1}{2} \rho V^3 \tag{4}$$

Consequently, wind power density is directly proportional to the cube of the wind speed. The useful power extracted by the turbine blades from the wind energy is determined by calculating the difference between the input wind power and the output wind power of the turbine:

$$P = \frac{1}{2} * K_m * (V_0^2 - V^2) \tag{5}$$

$$K_m = \rho A \frac{(V + V_0)}{2} \tag{6}$$

In Equation 5, V_0 is the speed of the incoming flow to the turbine, V is the speed of the outgoing flow from the turbine, and Km is the mass flow rate obtained from Equation 6. In Equation 6, the variable A is the cross-sectional area swept by the turbine blade, also known as the effective cross-sectional area (A=H*D, where H represents the blade height and D is the turbine diameter). Therefore, based on the above equations, the mechanical

power extracted by the turbine is mathematically expressed as shown in Equation 7:

$$P = \frac{1}{2}\left[\rho A \frac{V + V_0}{2}\right]\left(V_0^2 - V^2\right) \quad (7)$$

by having the formula for the power coefficient in the Equation 8 (41):

$$C_p = \frac{1}{2}\left(1+\frac{V}{V_0}\right)\left[1-\left(\frac{V}{V_0}\right)^2\right] \quad (8)$$

The mechanical power extracted by the turbine is expressed in Equation 9:

$$P = \frac{1}{2}\rho A V^3 C_p \quad (9)$$

$$\lambda = \frac{R * \omega}{V} \quad (10)$$

From Equation 9, it can be concluded that variations in the effective cross-sectional area and wind conditions directly affect the output power of the turbine. Therefore, the control of wind turbine systems primarily relies on these two factors. The ratio of the linear speed of the blade tip to the wind speed, known as the tip speed ratio (λ), is defined by the following equation:

In Equation 10, R is the radius of the blade, V is upstream wind velocity at the entrance of the turbine and ω is the angular velocity of the rotor which is calculated by Equation 11, with N representing the maximum rotational speed of the rotor at the corresponding wind speed and $\pi = 3.14$.

$$\omega = \frac{2\pi N}{60} \quad (11)$$

The experimental setup involves measuring wind speed, rotor rotational speed, and power output to calculate the power coefficient (Cp) under various operating conditions. The significance of Cp in quantifying the efficiency of the wind turbine rotor in converting wind energy into mechanical power will also be emphasized. This analysis aims to contribute to the refinement and optimization of turbine design for practical applications (41).

2. 2. Experimental Setup

2. 2. 1. Wind Section This section provides a detailed explanation of the testing procedure and related considerations. The parameters examined in this study focus on key elements relevant to turbines intended for urban environments. These include measuring the wind speed exiting the turbine to determine the power coefficient, calculating the tip speed ratio to identify the optimal turbine rotational speed within a specified wind speed range, and determining the speed-up time, defined as the time required for the turbine to reach its maximum rotational speed within the relevant wind speed range. Subsequently, each parameter is analyzed in detail to assess its impact on selecting and optimizing the turbine's aerodynamic design. The experiments are conducted within a wind tunnel to simulate airflow conditions around the turbine. A wind tunnel provides a controlled environment that replicates airflow around objects, allowing researchers to study and analyze aerodynamic behavior. This controlled simulation enables the evaluation of aerodynamic forces, such as lift and drag, and their effects on the performance and stability of turbine models or prototypes. Wind tunnels are widely used in aerodynamic research, particularly in studies aiming to improve turbine performance by understanding airflow dynamics. The insights gained from such simulations contribute to the design refinement and optimization prior to real-world implementation. The primary objective of this study is to analyze and compare the performance of specific turbines in urban settings. The wind speed range covered in the experiments extends from 0 to 12 m/s, which approximates the average wind speed commonly observed in rural areas. However, this value can vary significantly depending on factors such as geographic location, local topography, and prevailing weather conditions. Generally, wind speeds in urban environments tend to be lower than in open areas due to the presence of buildings and other structures that create turbulence and obstruct wind flow. Nonetheless, the precise wind speed in urban areas depends on variables including city layout, building heights, and localized weather patterns.

2. 2. 2. Turbine Structure As discussed in the literature, the efficiency of Savonius VAWTs is directly influenced by various factors related to their interior structural design, including blade shape, aspect ratio, number of blades, and several other parameters (24, 31). Certain characteristics are held constant, while others are varied for comparative analysis. Specifically, all parameters except the number of blades and blade shape remain unchanged. This study aims to evaluate the power coefficient and self-starting behavior of a Savonius VAWT featuring a DU-06-W-200 airfoil profile blade compared to a conventional semi-circular bladed rotor, considering both two-bladed and three-bladed configurations. To this end, four prototype Savonius turbines were fabricated. Each turbine comprises three primary components: blades, shaft, and end plates. Two of these turbine prototypes and their components are illustrated in Figure 2.

As shown in Figure 2, two blade models have been examined: the conventional semi-circular bladed rotor (section a) and the DU-06-W-200 profile bladed rotor (section b). Both blade types have a diameter of 125 mm and a height of 300 mm. Additionally, an important parameter in wind turbine design is the overlapping ratio, which indicates the extent to which one blade overlaps or

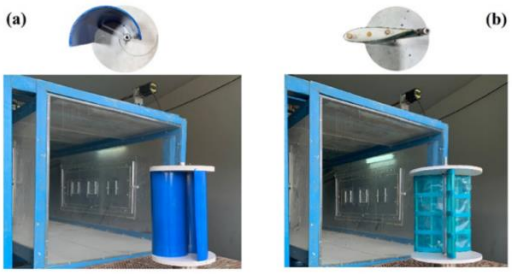


Figure 2. Samples of executed turbines: (a): Semi-circular blade and its related turbine (b): Du06-W-200 blade and its related turbine

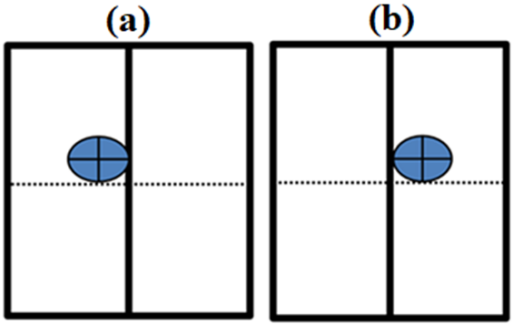


Figure 3. (a): measurement point for semi-circular bladed turbines (b): measurement point for Du06-W-200 profile bladed turbines

covers another. This ratio influences the smoothness of airflow around the blades. A higher overlapping ratio can promote smoother airflow, potentially improving turbine performance, whereas a lower overlapping ratio may cause increased turbulence, possibly reducing efficiency. The practical results of this research (42) revealed a significant effect of the overlap ratio on the aerodynamic performance of the Savonius wind turbine. Under high Reynolds number conditions, turbine prototypes without an overlap ratio demonstrated superior aerodynamic coefficients. Conversely, at lower Reynolds numbers, rotor configurations with a moderate overlap ratio yielded more favorable performance. Based on these findings and considering that the wind speeds examined range from low to average values, an overlap ratio of 32% was adopted for the conventional turbines. In contrast, the DU-06-W-200 profile blades were employed in a Savonius rotor for the first time in this study; therefore, an overlap ratio of zero was applied for these turbines, representing an initial step toward further investigation.

2. 2. 3. Experimental Testing Environment Experiments were conducted in an open-return wind tunnel with test section dimensions of 0.7 m × 0.7 m × 2 m at Babol Noshirvani University of Technology (NIT). The turbine was strategically positioned immediately downstream of the test section exit to minimize blockage effects and maintain flow uniformity. The exact positioning of the testing setup is illustrated in Figure 2.

2. 3. Testing Process This section provides a detailed description of the testing procedure and relevant considerations. Due to flow disturbances generated downstream of the turbine, the velocity varies at different coordinates along the rotor. Therefore, specific coordinates were selected to meaningfully define and measure the output wind speed. Additionally, since turbines with semi-circular blades rotate counterclockwise, whereas those with DU06-W-200 airfoil blades rotate clockwise, the wind speed measurement coordinates were adjusted accordingly to standardize and validate test results for both turbine types. The measurement points for both the semi-circular blade and airfoil blade turbines are illustrated in Figure 3. In Figure 3, the swept area of the turbine blades is depicted, with the shaft represented by a vertical line. All output wind speed measurements were conducted at these defined coordinates. Particular attention was given to minimizing experimental error by adopting a specific coordinate system aligned with the initial blade movement for each turbine. Although experiments were performed across various speed ranges, the coordinate system remained constant, ensuring consistent benchmarking of turbines. The initial blade movement coordinates utilized in this study are shown in Figure 4. Measurements were taken at discrete wind speeds ranging from 0 m/s to 12 m/s, increasing in 0.5 m/s increments. Prior to each trial, the flow was allowed to stabilize for a consistent period to ensure steady-state conditions. This approach ensured that each turbine design was tested under uniform and repeatable wind speed conditions.

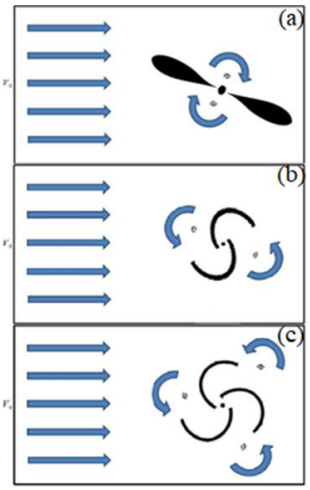


Figure 4. (a): starting coordinate of two-bladed Du06, (b): starting coordinate of two-bladed semi-circular, (c): starting coordinate of three-bladed semi-circular

The selection of these coordinates was intended to position the origin of movement at the point where the blade experiences the maximum pressure from the wind flow. By optimizing the starting conditions with higher initial pressure on the blades, the rotor is able to reach its maximum speed more quickly, thereby reducing the time required to achieve this speed. Additionally, as shown in Figure 4, the starting coordinates for the three-bladed turbine with the DU06-W-200 blade profile are not presented because this turbine did not start rotating without an external initial force. Consequently, measuring the speed-up time for this rotor was deemed illogical, and no specific starting coordinates were assigned for these turbines. Equally important, the wind speed entering the turbine is the most critical parameter influencing its performance. In this study, the wind speed range investigated extends from 0 to 12 meters per second.

2. 4. Instrument Precision and Uncertainty

Accurate characterization of Savonius wind turbine performance requires careful consideration of instrument precision and measurement uncertainty. The Marmonix MAN-745 anemometer employed in this study has a specified accuracy of ±(3% of reading + 0.20 m/s) and a resolution of 0.1 m/s, while the DT-2236C digital tachometer provides RPM measurements with a precision of ±(0.05% of reading + 1 digit). These uncertainties propagate into derived parameters such as the tip-speed ratio and power coefficient. All reported values incorporate these instrument limitations, with expanded uncertainties calculated at a 95% confidence level to ensure reliable and reproducible results.

3. RESULT AND DISCUSSION

As mentioned in introduction, the speed-up time refers to the duration required for the turbine to reach its final rotational speed. Therefore, a shorter speed-up time indicates a more optimal turbine performance, as the desired speed is achieved sooner to generate the required electrical energy. In this study, efforts were made to identify the highest torque or pressure generated by the wind force on the blades at the initiation phase when the blades are stationary. The starting points across all wind speed ranges are illustrated in Figure 4, which shows the starting movement coordinates for each turbine. It is important to note that different wind turbines exhibit varying wind speed thresholds for startup, meaning they begin rotating at different wind speeds. Consequently, the graphs depicting speed-up time start from different initial wind speeds. The speed-up time results for the turbines at various wind speeds are presented in Figure 5, while the design of experiment (DOE) layout, including key factor levels and a summary of experimental runs, is provided in Table 2. The graph related to the three-bladed turbine with the Du06-W-200 VAWT Airfoil profile is not visible because, as mentioned earlier, the turbine did not start rotating without an initial force.

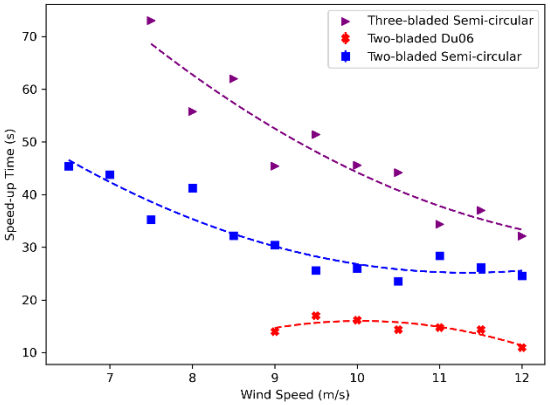


Figure 5. Speed-up time to wind speed graph

TABLE 2. DOE table for speed-up time experiment

Run	Blade profile	Blade number	Windspeed (m/s)	Speed-up time (s)
1	Du06	2	9	14
2	Du06	2	12	11
3	SC	2	6.5	45.4
4	SC	2	10.5	23.6
5	SC	2	12	24.6
6	SC	3	7.5	73
7	SC	3	12	32.2

Numerous laboratory experiments have been conducted on the starting behavior of wind turbines (10, 38, 43, 44). It has been concluded that the starting behavior of wind turbines is related to the rotor position at startup relative to the wind flow (38). However, evidence also suggests that the starting behavior depends on the incoming flow to the rotor and the initial force applied to the turbine. The results of these experiments indicate that the initial torque coefficient applied to the rotor is entirely a function of the turbine's position, while the turbine power coefficient depends on the relative velocity of the blade tip to the wind speed. Additionally, the significance of negative starting torque in turbines is noteworthy in many situations. Some studies suggest that the degree of blade overlap significantly influences the magnitude of the negative torque region in turbines (45, 46).

In this study, zero overlap was considered for the DU06 profile, preventing this turbine from moving in any rotor position. In other words, no initiation of movement occurred at any position or wind speed for this turbine, effectively resulting in a locked state. Consequently, no values appear in the speed-up time graph for the three-

bladed turbine with the DU06-W-200 VAWT Airfoil profile. This is because the turbine required an initial speed to start, and this initiation does not imply any subsequent decrease in speed over time. The turbine consistently operated at this initial speed throughout all experiments. As shown in Figure 5 and Table 2, the best performance is exhibited by the two-bladed DU06 turbine compared to the other turbines, followed by the two-bladed semi-circular turbine, and then the three-bladed semi-circular turbine. This performance ranking is influenced by speed-up time, where reaching the final speed faster is considered more optimal. The preference for the two-bladed semi-circular turbine over the three-bladed one is due to the lower weight of the rotor setup—an attribute that also affected the initiation of the turbine's movement. The two-bladed turbine, being lighter, required less starting torque and began rotating at a wind speed of 6.5 m/s, while the three-bladed rotor of the same model started at 7.5 m/s.

The DU06-W-200 VAWT profile, known for its rapid startup time, has demonstrated specialized performance in vertical-axis turbines. As noted in previous research (39), this profile exhibits unique efficiency in self-starting capabilities.

The Tip Speed Ratio (TSR), defined as the ratio of the tangential speed of the blade tip to the incoming wind speed, is a critical performance parameter in evaluating wind turbine efficiency. It governs the interaction between the rotor blades and the wind stream, significantly influencing the power coefficient and turbine behavior. Figure 6 and Table 3 depict the turbines' tip speed ratios at different wind speeds and key experimental runs.

Figure 6 and Table 3 illustrate the variation of the tip speed ratio (TSR, denoted as λ) with respect to free-stream wind speed for different blade configurations and profiles. It is evident from the results that the semicircular two-bladed configuration exhibits the highest TSR values across the entire wind speed range, peaking at approximately 1.65 at a wind speed of 12 m/s. This

TABLE 3. DOE table for λ

Run	Blade profile	Blade number	Windspeed (m/s)	Tip speed ratio
1	Du06	3	8	0.338
2	Du06	3	12	0.517
3	Du06	2	9	0.544
4	Du06	2	12	0.631
5	SC	3	7.5	0.836
6	SC	3	12	1.175
7	SC	2	6.5	1.021
8	SC	2	12	1.651

indicates superior rotational responsiveness under increasing wind speeds. In contrast, the Du06 three-bladed model shows a relatively moderate TSR profile, with maximum values around 0.52. The semicircular two-bladed turbine demonstrated the most efficient aerodynamic behavior in terms of rotational acceleration, while the Du06 three-bladed rotor exhibited the lowest TSR values, indicating greater resistance or inertia under the same conditions. These findings suggest that, while the semicircular two-bladed configuration may be more favorable in applications requiring rapid rotor acceleration, the Du06 configurations might offer benefits in terms of torque stability at lower rotational speeds. The trends also indicate a diminishing gain in TSR with increasing blade count, likely due to increased aerodynamic drag and interaction between blades.

Another performance parameter evaluated in this study is the power output of the turbine. Power output serves as the definitive metric for assessing a wind turbine's energy conversion efficiency and practical viability, as it represents the mechanical energy extracted from the wind (Equation 9). The power characteristics of the tested

Savonius turbine configurations are presented in

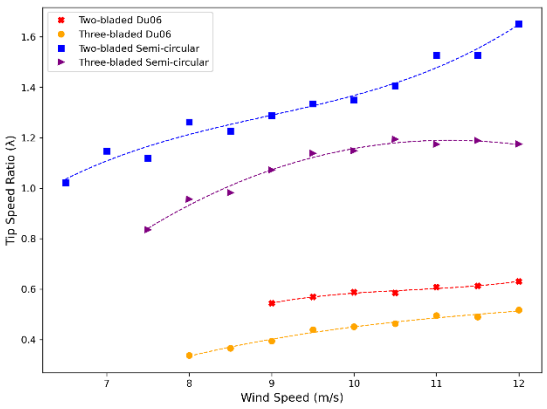


Figure 6. The tip speed ratio to wind speed graph

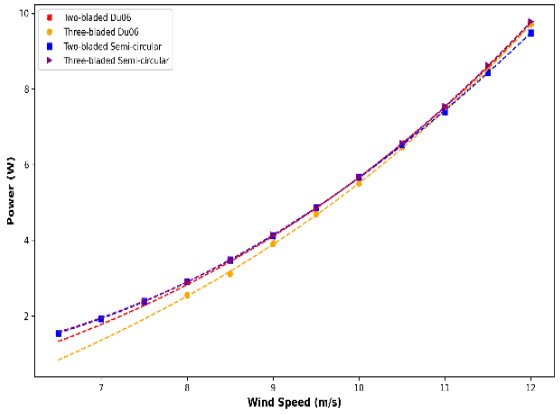


Figure 7. power to wind speed graph

Figure 7, while Table 4 provides a detailed summary of all experimental runs.

At wind speeds where all configurations generate power, the two-bladed Du06 generally produces slightly higher power than the three-bladed Du06. For example, at 9 m/s, the two-bladed Du06 outputs 4.11 W compared to 3.90 W for the three-bladed. This trend continues up to 12 m/s, where the two-bladed produces 9.74 W, marginally exceeding the three-bladed's 9.73 W. Similarly, the two- and three-bladed semi-circular turbines deliver comparable power outputs at corresponding wind speeds, with the three-bladed configuration showing a slight edge at the highest speed of 12 m/s (9.78 W versus 9.49 W). Overall, the two-bladed Du06 outperforms the three-bladed version by a small margin, while the semi-circular blades show near-equal performance between blade counts.

According to Figure 8 and Table 5, an upward trend in rotor speed with increasing wind speed is evident for all turbines across the wind spectrum up to 12 m/s. In general, higher rotational speeds were observed in the two-bladed semi-circular profile turbine compared to the others, with its rotational speed positioned relatively close to that of the three-bladed counterpart. This trend is also noticeable for the Du06 series, where the two-bladed and three-bladed turbines exhibit similar behavior. Additionally, two-bladed rotors generally demonstrated higher rotational speeds than three-bladed rotors, although the difference between them was more pronounced. These observations align well with the results of many previous studies (33, 47-49), indicating that two-bladed rotors tend to have higher rotational speeds. The higher rotational speed in two-bladed turbines compared to three-bladed ones is primarily attributed to their lower weight.

It is significant to point out that, aerodynamically, the semi-circular profile has shown greater effectiveness in terms of rotational speed compared to the Du06 profile. In Savonius turbines, which are highly drag-oriented, turbines with a semi-circular profile exhibit higher rotational speeds than those with the Du06 profile. The last parameter under consideration is turbine efficiency, defined as the ratio of the power coefficient to the tip speed ratio. The relationship between these two parameters (λ and the power coefficient) is presented in Figure 9 and Table 6.

TABLE 4. DOE table for power

Wind speed	P(Du06-2b)	P(Du06-3b)	P(SC-2b)	P(SC-3b)
6.5			1.55	
7			1.93	
7.5			2.39	2.39
8		2.56	2.90	2.90
8.5		3.12	3.48	3.48
9	4.11	3.90	4.13	4.13
9.5	4.85	4.69	4.86	4.85
10	5.65	5.50	5.67	5.67
10.5	6.56	6.47	6.53	6.56
11	7.52	7.41	7.41	7.54
11.5	8.58	8.50	8.45	8.62
12	9.74	9.73	9.49	9.78

TABLE 5. DOE table for the rotational speed

Run	Blade profile	Blade number	Windspeed (m/s)	Rotational speed(rpm)
1	Du06	3	8	195
2	Du06	3	12	449
3	Du06	2	9	354
4	Du06	2	12	547
5	SC	3	7.5	453
6	SC	3	12	1077
7	SC	2	6.5	480
8	SC	2	12	1433

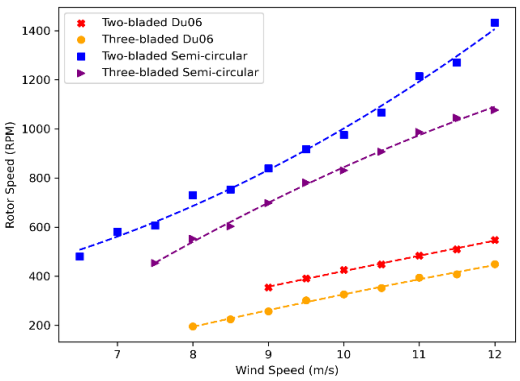


Figure 8. The rotational speeds to wind speed graph

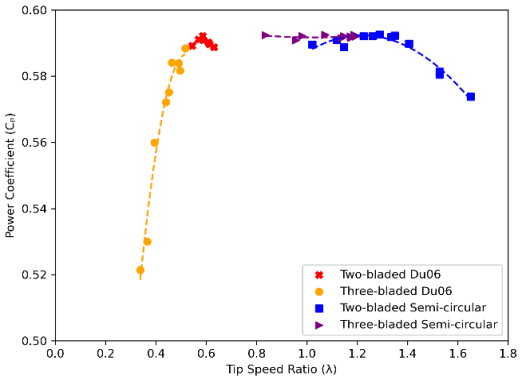


Figure 9. The coefficient power to lambda graph

TABLE 6. DOE table for the power coefficient to lambda

Run	Blade profile	Blade number	Power coefficient	Tip speed ratio
1	Du06	3	0.521	0.337
2	Du06	3	0.588	0.517
3	Du06	2	0.588	0.63
4	Du06	2	0.592	0.585
5	SC	2	0.580	1.526
6	SC	2	0.592	1.288
7	SC	3	0.590	0.956
8	SC	3	0.592	1.189

As demonstrated in Figure 6, two-bladed turbines achieve higher tip speed ratios compared to three-bladed turbines. One reason for the decrease in rotational speed with an increasing number of blades is the expansion of the drag area. However, this expansion also results in an increase in reverse torque relative to the total torque. In other words, while the drag force area increases positively with the addition of blades, the wind flow that generates lift force—which acts as a negative factor—increases simultaneously.

Generally, Savonius turbines are highly drag-based rotors, meaning that the main force driving rotational speed is the drag force, while the lift force partially acts as a negative force. The drag-based operation occurs as the curved blades of the Savonius turbine face the wind, causing air pressure imbalances that generate a net drag force on the concave part of the blade, which in turn causes the blades to rotate. Conversely, the lift force, acting on the convex part of the blade, tends to reverse the direction of rotation. The combined effect of these opposing forces negatively impacts the rotor's output. Therefore, adding more blades to the turbine does not necessarily result in an increase in angular velocity. Higher tip speed ratios are observed in the semi-circular turbines, due to their better aerodynamics for capturing drag forces. The three-bladed configurations exhibit an upward trend in TSR, while the two-bladed data show an ascending-descending pattern with a peak point. This indicates that three-bladed turbines demonstrate greater stability in terms of power coefficient, as airfoil efficiency depends on the turbine's utilization of input wind speed, and three-bladed turbines have a larger surface area for capturing wind compared to their two-bladed counterparts. Consequently, the output flow speed in three-bladed turbines is notably more stable than in two-bladed turbines. However, the analysis of power coefficients and optimal power coefficients in this study suggests that the difference in power output between the two configurations is minimal. On the other hand, two-bladed turbines achieve higher rotational speeds, primarily due to their lighter weight, resulting in higher speed but lower torque compared to three-bladed turbines. Thus, the tip speed ratio (λ) in two-bladed turbines is greater than in three-bladed turbines.

In this study, semi-circular blades demonstrated better speed and overall performance compared to the Du06 blades, although turbines with Du06 blades showed better speed-up times due to their specific aerodynamic profile. Table 7 summarizes the optimal parameters calculated for all turbine configurations.

According to the data shown in Table 7, the shortest speed-up time belonged to the two-bladed turbine with the Du06 profile, clocking in at approximately 11 seconds, about twice as fast as the two-bladed semi-circular turbine and three times faster than the three-bladed semi-circular turbine. In terms of final rotational speed, the two-bladed semi-circular turbine performed best, reaching 1433 rpm, which is roughly 33% faster than the three-bladed semi-circular turbine. This rotational speed exceeds that of the two- and three-bladed Du06 turbines by approximately 162 and 219%, respectively. Regarding the power coefficient, the two-bladed semi-circular turbine showed a slight advantage over the other turbines. The tip-speed ratio further emphasizes these performance differences: the two-bladed semi-circular turbine achieved the highest tip speed ratio of 1.65, indicating superior blade efficiency under the tested conditions. In comparison, the Du06 turbines had significantly lower tip speed ratios of 0.63 and 0.52 for the two- and three-bladed designs, respectively. Power output was relatively consistent across all configurations at their peak performance, with the three-bladed semi-circular turbine delivering the highest output at 9.78 W.

3. 1. Comparative Analysis The current study's findings align with and expand upon previous research on Savonius vertical axis wind turbines, particularly regarding blade profiles and blade counts. Ikonwa et al. (39) identified the DU-06-W-200 airfoil as a promising design due to its superior starting torque and aerodynamic performance, which corroborates our observation that turbines with the Du06 profile, especially the two-bladed configuration, exhibit faster speed-up times and competitive tip speed ratios. This confirms the effectiveness of the DU-06-W-200 airfoil in enhancing turbine self-starting and rotational

TABLE 7. Optimal parameters for each turbine

Blade profile	Cp	λ	Rotation speed (rpm)	Speed-up time (s)	Power (W)
SC-2b	0.5925	1.65	1433	23.6	9.48
SC-3b	0.5920	1.19	1077	32.17	9.78
Du06-2b	0.5921	0.63	547	11	9.73
Du06-3b	0.5884	0.52	449	-	9.73

performance. In contrast, Wenehenubun et al. (32) concluded that three-bladed Savonius turbines demonstrate higher tip speed ratios and efficiency at lower wind speeds, with four-bladed configurations producing maximum torque in the 0–10 m/s range. Our results partially agree with this, as the three-bladed semi-circular turbine exhibited the highest power output and the most stable performance, even though the two-bladed Du06 turbine outperformed others in speed-up time and tip speed ratio. This suggests that blade number and profile interact complexly to affect different performance metrics, with the three-bladed design favoring power stability and magnitude, and the two-bladed Du06 profile excelling in acceleration and tip speed. Additionally, Gumilar et al. (33) reported that two-bladed turbines achieved the highest power capture at higher wind speeds (20 m/s). Although our experimental wind speeds were lower (up to 12 m/s), the trend of superior power generation by two-bladed turbines, particularly those with the Du06 profile, is consistent with their findings. This highlights the potential of two-bladed configurations in optimizing power output under varying wind conditions, reinforcing the importance of blade count and aerodynamic design in turbine efficiency. Overall, these comparisons underscore that both blade profile and blade number critically influence Savonius turbine performance, with the DU-06-W-200 airfoil and two-bladed configurations showing notable advantages in key operational parameters such as speed-up time and tip speed ratio, while three-bladed designs favor sustained power output.

4. CONCLUSION

This work aimed to investigate Savonius vertical axis wind turbines, with an emphasis on comparing performance parameters across different blade numbers and profiles. Four turbines, ranging from two to three blades and featuring semi-circular and Du06 airfoil profiles, were experimentally analyzed. The comparison focused on key parameters such as speed-up time, final rotational speed, power output, tip speed ratio, and power coefficient. The findings indicated that two-bladed turbines, particularly those with the Du06 profile, outperformed their counterparts in speed-up time and rotational speed, primarily due to their lower weight and aerodynamic advantages of the specific blade profile. However, the three-bladed semi-circular turbine demonstrated the most stable performance and the highest power output, despite the Du06-profile turbines exhibiting slightly lower tip speed ratios. These insights provide a comprehensive understanding of the factors influencing the performance dynamics of Savonius wind turbines, which is essential for optimizing their use in renewable energy applications. In summary, the study yielded several key findings, which are outlined below.

i. Two-bladed turbines, especially those featuring the Du06 profile, exhibited superior performance in speed-up time, tip speed ratio and rotational speed compared to their counterparts.
ii. The semi-circular turbines, particularly the two-bladed design, consistently achieved higher tip speed ratios, reflecting enhanced aerodynamic efficiency in converting wind speed to rotational speed.
iii. The semi-circular three-bladed turbine demonstrated the most stable performance and highest power, even though the Du06-profiled turbines had slightly lower tip speed ratios.
iv. Despite the slightly lower TSRs, the two-bladed Du06 turbine was able to deliver comparable or higher power output than the three-bladed Du06 turbine at elevated wind speeds, demonstrating the aerodynamic advantages of the Du06 profile in power extraction.
v. The lowest speed-up time was recorded for the two-bladed turbine with the Du06 profile, approximately twice and three times faster than both two and three-bladed semi-circular profile turbines
vi. The semi-circular two-bladed turbine achieved the highest final rotational speed at 1433 RPM, outperforming the semi-circular three-bladed turbine by about 33%. In comparison with the two and three-bladed Du06 turbines, this speed was 162% and 219% higher, respectively.
vii. The power coefficient of the semi-circular two-bladed turbine slightly surpasses that of the other turbines, with minimal variances observed among the two-bladed Du06, semi-circular three-bladed, and Du06 three-bladed turbines.

Given the existing theoretical framework, findings, and current limitations, the following research suggestions and directions can be proposed, which provide the groundwork for new research avenues:

1) Investigating and comparing various types of Savonius turbine blades to enhance performance.
2) investigating intermediate designs (e.g., hybrid overlapping blades) and higher blade counts (four or more)
3) Adjusting the overlap ratio of blades to determine the optimal overlapping efficiency in turbines.
4) Conducting experiments at higher inlet speeds.
5) Employing combined Darrieus and Savonius airfoils in vertical-axis turbines.
6) Experimenting and evaluating the performance of turbines implemented in water tunnels to reduce flow velocity and pressure in channels and generate electricity.

STATEMENTS AND DECLARATIONS

Funding The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.

Availability of data and materials The data that support the findings of this study are available on request from the corresponding author.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval We would confirm that this paper has not been published nor submitted for publication elsewhere. We confirm that we have read, understand and agreed to the submission guidelines, policies and submission declaration of the journal.

5. REFERENCES

1. Benbouhenni H. Amelioration effectiveness of torque and rotor flux control applied to the asynchronous generator for dual-rotor wind turbine using neural third-order sliding mode approaches. International Journal of Engineering Transactions C: Aspects. 2022;35(3):517-30. doi: 10.5829/ije.2022.35.03C.04.
2. Adelekan OA, Ilugbusi BS, Adisa O, Obi OC, Awonuga KF, Asuzu OF, Ndubuisi NL. Energy transition policies: a global review of shifts towards renewable sources. Engineering Science & Technology Journal. 2024;5(2):272-87. doi: 10.51594/estj.v5i2.752.
3. Asgari S, Menhaj M, Suratgar AA, Kazemi M. A disturbance observer based fuzzy feedforward proportional integral load frequency control of microgrids. International Journal of Engineering Transactions A: Basics 2021; 34(7): 1694-1702.. doi: 10.5829/ije.2021.34.07a.13.;
4. Al Mubarak F, Rezaee R, Wood DA. Economic, Societal, and Environmental Impacts of Available Energy Sources: A Review. Engineering. 2024;5(3):1232-65. doi: 10.3390/eng5030067.
5. Fard SK, Ahmadi H, Elizei MA. Electricity Supply Model of Conventional Residential Buildings in Tehran with Priority on Renewable Energy Using Adaptive Fuzzy-neural Inference System. ternational Journal of Engineering Transactions A: Basics 2023; 34(10): 1793-1814. doi: 10.5829/ije.2023.36.10a.07.
6. Bozsik N, Szeberényi A, Bozsik N. Impact of Climate Change on the Performance of Household-Scale Photovoltaic Systems. HighTech and Innovation Journal. 2024;5(1):1-15. doi: 10.28991/HIJ-2024-05-01-01.
7. Woodruff TJ. Health effects of fossil fuel–derived endocrine disruptors. New England Journal of Medicine. 2024;390(10):922-33. doi: 10.1056/NEJMra2300476.
8. Shame BH, Tjahjana DD, Aziz M, Mbumba MT. Analysis of Seasonal Wind Energy Potential on Zanzibar Coastal Island. HighTech and Innovation Journal. 2024;5(2):331-48. doi: 10.28991/HIJ-2024-05-02-08.
9. Carfora A, Pansini RV, Scandurra G. Energy dependence, renewable energy generation and import demand: Are EU countries resilient? Renewable Energy. 2022;195:1262-74. doi: 10.1016/j.renene.2022.06.098.
10. Chegini S, Asadbeigi M, Ghafoorian F, Mehrpooya M. An investigation into the self-starting of Darrieus-Savonius hybrid wind turbine and performance enhancement through innovative deflectors: A CFD approach. Ocean Engineering. 2023;287:115910. doi: 10.1016/j.oceaneng.2023.115910.
11. Hannan M, Al-Shetwi AQ, Mollik M, Ker PJ, Mannan M, Mansor M, Al-Masri HM, Mahlia TI. Wind energy conversions, controls, and applications: a review for sustainable technologies and directions. Sustainability. 2023;15(5):3986. doi: 10.3390/su15053986.
12. Foomani AH, Moradlou M, Nazarian P. Utilizing a New Voltage Stability Index in Distribution Power System in Presence of Wind Turbine Units. International Journal of Engineering Transactions A: Basics. 2024;37(4):711-24. doi: 10.5829/ije.2024.37.04a.11
13. Veers P, Bottasso CL, Manuel L, Naughton J, Pao L, Paquette J, Robertson A, Robinson M, Ananthan S, Barlas T, Bianchini A. Grand challenges in the design, manufacture, and operation of future wind turbine systems. Wind Energy Science. 2023;8(7):1071-131. doi: 10.5194/wes-8-1071-2023.
14. Usman FO, Ani EC, Ebirim W, Montero DJP, Olu-lawal KA, Ninduwezuor-Ehiobu N. Integrating renewable energy solutions in the manufacturing industry: challenges and opportunities: a review. Engineering Science & Technology Journal. 2024;5(3):674-703. doi: 10.51594/estj.v5i3.865 .
15. Cruz-Pérez N, Rodríguez Alcántara JS, Koronaiou VL, Jančula A, Rodríguez-Martín J, García Gil A, Fontes JC, Santamarta JC. . SWOT analysis of the benefits of hydropower energy in four archipelagos. Civil Engineering Journal 2024; 10(7). doi: 10.28991/CEJ-2024-010-07-019.
16. Elgendi M, AlMallahi M, Abdelkhalig A, Selim MY. A review of wind turbines in complex terrain. International Journal of Thermofluids. 2023:100289. doi: 10.1016/j.ijft.2023.100289.
17. Al-Rawajfeh MA, Gomaa MR. Comparison between horizontal and vertical axis wind turbine. International Journal of Applied. 2023;12(1):13-23. doi: 10.11591/ijape.v12.i1.pp13-23.
18. Elkodama A, Ismaiel A, Abdellatif A, Shaaban S, Yoshida S, Rushdi MA. Control Methods for Horizontal Axis Wind Turbines (HAWT): State-of-the-Art Review. Energies. 2023;16(17):6394. doi: 10.3390/en16176394.
19. Yusof A, Mohamed M, editors. Vertical Axis Wind Turbines: An Overview. InECCE2019: Proceedings of the 5th International Conference on Electrical, Control & Computer Engineering, Kuantan, Pahang, Malaysia, 29th July 2019; 2020: Springer. doi: 10.1007/978-981-15-2317-5_68.
20. Karimian S, Saham S. Power and Ventilation Performance Study in a Modified Vertical Axis Wind Turbine Based on Semi-analytical Approach. International Journal of Engineering Transactions A: Basics. 2023;36(7):1331-42. doi: 10.5829/ije.2023.36.07a.14
21. Cuevas-Carvajal N, Cortes-Ramirez J, Norato JA, Hernandez C, Montoya-Vallejo M. Effect of geometrical parameters on the performance of conventional Savonius VAWT: A review. Renewable and Sustainable Energy Reviews. 2022;161:112314. doi: 10.1016/j.rser.2022.112314.
22. Karmakar SD, Chattopadhyay H. A review of augmentation methods to enhance the performance of vertical axis wind turbine. Sustainable Energy Technologies and Assessments. 2022;53:102469. doi: 10.1016/j.seta.2022.102469.
23. Maldar NR, Ng CY, Oguz E. A review of the optimization studies for Savonius turbine considering hydrokinetic applications. Energy Conversion and Management. 2020;226:113495. doi: 10.1016/j.enconman.2020.113495.

24. Deda Altan B, Gultekin GS. Investigation of Performance Enhancements of Savonius Wind Turbines through Additional Designs. Processes. 2023;11(5):1473. doi: 10.3390/pr11051473.
25. Wong KH, Chong WT, Sukiman NL, Poh SC, Shiah Y-C, Wang C-T. Performance enhancements on vertical axis wind turbines using flow augmentation systems: A review. Renewable and sustainable Energy reviews. 2017;73:904-21. doi: 10.1016/j.rser.2017.01.160.
26. Layeghmand K, Ghiasi Tabari N, Zarkesh M. Improving efficiency of Savonius wind turbine by means of an airfoil-shaped deflector. Journal of the Brazilian Society of Mechanical Sciences and Engineering. 2020;42(10):528. doi: 10.1007/s40430-020-02598-7.
27. Yuwono T, Sakti G, Aulia FN, Wijaya AC. Improving the performance of Savonius wind turbine by installation of a circular cylinder upstream of returning turbine blade. Alexandria Engineering Journal. 2020;59(6):4923-32. doi: 10.1016/j.aej.2020.09.009.
28. Nimvari ME, Fatahian H, Fatahian E. Performance improvement of a Savonius vertical axis wind turbine using a porous deflector. Energy Conversion and Management. 2020;220:113062. doi: 10.1016/j.enconman.2020.113062.
29. Mohammadi M, Mohammadi R, Ramadan A, Mohamed M. Numerical investigation of performance refinement of a drag wind rotor using flow augmentation and momentum exchange optimization. Energy. 2018;158:592-606. doi: 10.1016/j.energy.2018.06.072.
30. Kalluvila JB, Sreejith B. Numerical and experimental study on a modified Savonius rotor with guide blades. International Journal of Green Energy. 2018;15(12):744-57. doi: 10.1080/15435075.2018.1529574.
31. Amin M, ZoheirSaboohi HHM, Osgouie KG. Investigation of the factors affecting the performance of Savonius wind turbine. Journal of Renewable Energy and New Energy. 2023;10(2). doi: 10.52547/jrenew.10.2.192.
32. Wenehenubun F, Saputra A, Sutanto H. An experimental study on the performance of Savonius wind turbines related with the number of blades. Energy procedia. 2015;68:297-304. doi: 10.1016/j.egypro.2015.03.259.
33. Gumilar L, Kusumawardana A, Habibi MA, Afandi AN, Prihanto D, Aji AF, editors. Performance analysis of vertical wind turbine type Savonius-L based on wind speed, rotation speed, and number of blades. 2019 International Seminar on Application for Technology of Information and Communication (iSemantic); 2019: IEEE. doi: 10.1109/ISEMANTIC.2019.8884302.
34. Prabowoputra DM, Prabowo AR. Effect of Geometry Modification on Turbine Performance: Mini-Review of Savonius Rotor. International Journal of Mechanical Engineering and Robotics Research. 2022;11(10):777-83. doi: 10.18178/ijmerr.11.10.777-783.
35. Al-Shammari SA, Karamallah AH, Aljabair S, editors. Blade shape optimization of savonius wind turbine at low wind energy by artificial neural network. IOP Conference Series: Materials Science and Engineering; 2020: IOP Publishing. doi: 10.1088/1757-899X/881/1/012154.
36. Tartuferi M, D'Alessandro V, Montelpare S, Ricci R. Enhancement of Savonius wind rotor aerodynamic performance: a computational study of new blade shapes and curtain systems. Energy. 2015;79:371-84. doi: 10.1016/j.energy.2014.11.023.
37. Hosseini Imeni SZ, Kaabinejadian A, Ami Ahmadi H, Moghimi M. Optimal design and sensitivity analysis of airfoil-shaped rotor blade for Savonius wind turbine by using response surface methodology. Wind Engineering. 2022;46(4):1203-23. doi: 10.1177/0309524X211072868.
38. Shamsuddin M, Kamaruddin NM. Experimental study on the characterization of the self-starting capability of a single and double-stage Savonius turbine. Results in Engineering. 2023;17:100854. doi: 10.1016/j.rineng.2022.100854.
39. Ikonwa C, Gale W, Ingham D, Ma L, Pourkashanian MJIJoES. Investigation of airfoil profiles for use in self-starting, small-scale, low reynolds number vertical axis wind turbines. The International Journal of Environmental Sustainability 2016;12(3):33-52. doi: 10.18848/2325-1077/CGP/v12i03/33-52.
40. Zamani M, Sangtarash A, Maghrebi MJ. Numerical study of porous media effect on the blade surface of vertical axis wind turbine for enhancement of aerodynamic performance. Energy Conversion and Management. 2021;245:114598. doi: 10.1016/j.enconman.2021.114598.
41. Patel MR. Wind and solar power systems. 2009. doi: 10.1201/9781003042952.
42. Morshed KN. Experimental and numerical investigations on aerodynamic characteristics of Savonius wind turbine with various overlap ratios. 2010. URL: https://digitalcommons.georgiasouthern.edu/etd/773.
43. Bourhis M, Zhang S, Pereira M, Ravelet F, editors. Wind tunnel experimental study of the effects of blade number on the performance and starting behavior of a low tip-speed ratio and micro-scale wind turbine at fixed blade geometry. Turbo Expo: Power for Land, Sea, and Air; 2023: American Society of Mechanical Engineers. doi: 10.3390/fluids9120299.
44. Farzadi R, Bazargan M. 3D numerical simulation of the Darrieus vertical axis wind turbine with J-type and straight blades under various operating conditions including self-starting mode. Energy. 2023;278:128040. doi: 10.1016/j.energy.2023.128040.
45. Du L, Ingram G, Dominy RG. A review of H-Darrieus wind turbine aerodynamic research. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science. 2019;233(23-24):7590-616. doi: 10.1177/0954406219885962.
46. Alomar Z, Khoury G, Rishmany J, Daaboul M. Numerical and Experimental Study on the Effect of Overlap on Savonius Wind Turbine. International Journal of Power and Energy Systems. 2023;43(10):2023.203-0450. doi: 10.2316/J.2023.203-0450.
47. Ali MH. Experimental comparison study for Savonius wind turbine of two & three blades at low wind speed. International Journal of Modern Engineering Research (IJMER). 2013 Sep;3(5):2978-86.
48. Halmy MSM, Didane DH, Afolabi LO, Al-Alimi S. Computational fluid dynamics (cfd) study on the effect of the number of blades on the performance of double-stage savonius rotor. CFD Letters. 2021;13(4):1-10. doi: 10.37934/cfdl.13.4.110.
49. Thiyagaraj J, Rahamathullah I, Anbuchezhiyan G, Barathiraja R, Ponshanmugakumar A. Influence of blade numbers, overlap ratio and modified blades on performance characteristics of the savonius hydro-kinetic turbine. Materials Today: Proceedings. 2021;46:4047-53. doi: 10.1016/j.matpr.2021.02.568.

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

در دهه‌های اخیر، تأثیر زیست‌محیطی تولید انرژی منجر به تلاش‌های فزاینده جهانی برای گذار از سوخت‌های فسیلی آلاینده به منابع انرژی تجدیدپذیر شده و توجه قابل توجهی را از سوی سازمان‌ها و محققان به خود جلب کرده است .این گذار به دلیل افزایش نگرانی‌های زیست‌محیطی و محدودیت دسترسی به سوخت‌های فسیلی ضروری است .همزمان، استقرار توربین‌های بادی در سراسر جهان به طور قابل توجهی گسترش یافته است .هدف این مطالعه ارائه یک طراحی بهینه توربین بادی محور عمودی Savoniusمناسب برای کاربردهای شهری است .دو مدل توربین با تعداد و پروفیل‌های مختلف پره طراحی، ساخته و به صورت تجربی آزمایش شدند .این توربین‌ها شامل پیکربندی‌های دو پره و سه پره بودند که از پروفیل‌های پره نیم‌دایره‌ای و Du06-W-200 استفاده می‌کردند .ارزیابی عملکرد بر پارامترهای کلیدی از جمله ضریب توان، سرعت چرخش نهایی، توان خروجی، نسبت سرعت نوک و زمان رسیدن به حداکثر سرعت چرخش متمرکز بود .نتایج نشان داد که توربین دو پره نیم‌دایره‌ای بالاترین سرعت چرخش نهایی ۱٤۳۳ دور در دقیقه را ارائه می‌دهد که تقریباً ۳۳٪ بیشتر از توربین نیم‌دایره‌ای سه پره و بیش از ۱٦۰٪ بیشتر از توربین‌های با پروفیل Du06 است .همچنین این توربین بالاترین ضریب توان را در بین تمام توربین‌های آزمایش‌شده به دست آورد .توربین دو پره (Du06-W-200) با رسیدن به سرعت نهایی تنها در ۱۱ ثانیه، دو برابر سریع‌تر از توربین نیم‌دایره‌ای دو پره و سه برابر سریع‌تر از پیکربندی نیم‌دایره‌ای سه پره، سریع‌ترین شتاب را نشان داد .در حالی که توربین‌های با پروفیل Du06 نسبت سرعت نوک کمی پایین‌تری نشان دادند، توربین نیم‌دایره‌ای سه پره پایدارترین توان خروجی را نشان داد .این یافته‌ها نشان می‌دهد که توربین‌های ساونیوس با پیکربندی‌های آزمایش‌شده برای محیط‌های شهری مناسب هستند، اگرچه اثربخشی آنها به شدت به شرایط جغرافیایی و آب و هوایی محلی بستگی دارد.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1191-1199

International Journal of Engineering

Journal Homepage: www.ije.ir

Methods of Ensuring the Quality of Assembly of Non-removable Joints from Dissimilar Materials

V. V. Maksarov, M. S. Sinyukov*

Department of Mechanical Engineering, Empress Catherine II Saint Petersburg Mining University, Russia

P A P E R I N F O

Paper history:
Received 24 April 2025
Received in revised form 23 June 2025
Accepted 13 July 2025

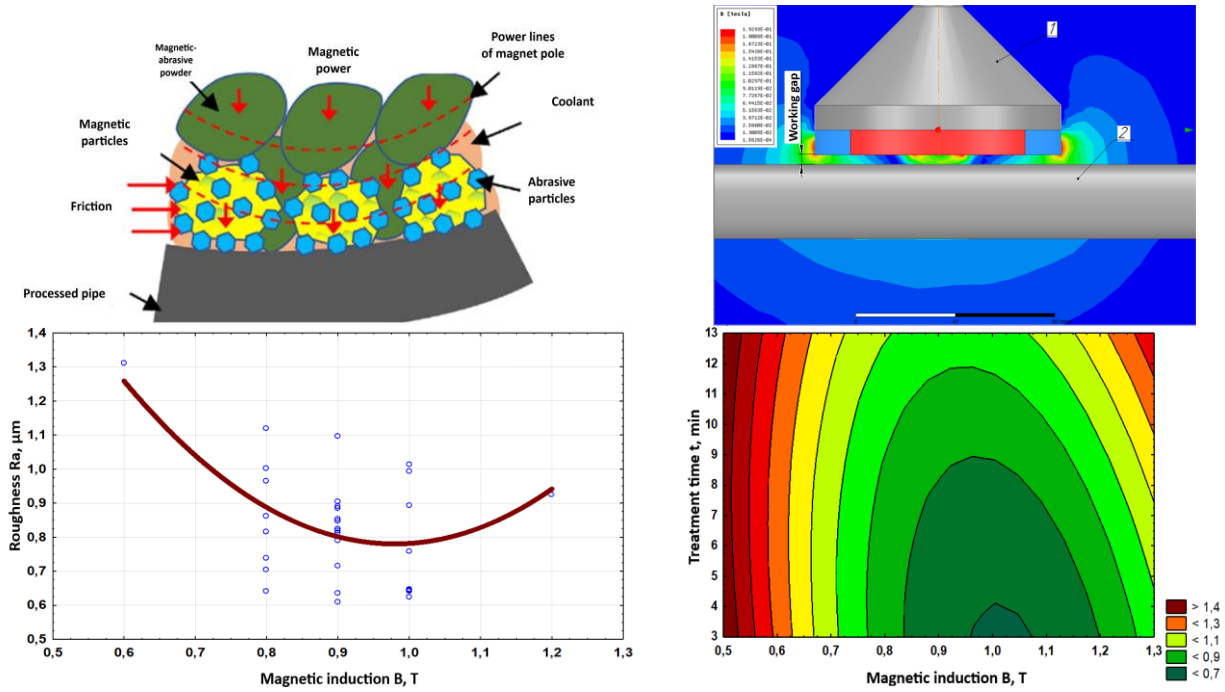
Keywords:
Bonded Abrasive Particles
Electromagnets
Hard-to-process Materials
Magnetic Abrasive Processing
Permanent Magnets
Process Automatization

A B S T R A C T

This article addresses the critical challenge of achieving high-quality surface finishing for dissimilar material joints, which is essential for aerospace, energy, and precision engineering applications. Conventional methods often fail to deliver consistent results due to material property differences. We developed an optimized magnetic-abrasive treatment (MAT) process using NdFe permanent magnets 0.6-1.2 T and composite abrasives Fe/SiC. Experimental results demonstrate that our method achieves superior surface uniformity Ra 0.6-1.0 μm with 2-3 times higher productivity compared to manual grinding. Key findings include the determination of optimal parameters B = 1.0 T, n = 460 min^{-1}. The novelty of this work lies in the systematic optimization of MAT for dissimilar material joints and the first demonstration of industrial scalability for components up to 0.5 m^2. These results expand MAT applications, offering a cost-effective alternative to exist finishing methods.

***doi**: 10.5829/ije.2026.39.05b.12*

Graphical Abstract



*Corresponding Author Email: sms-0305@yandex.ru (M. S. Sinyukov)

Please cite this article as: Maksarov VV, Sinyukov MS. Methods of Ensuring the Quality of Assembly of Non-removable Joints from Dissimilar Materials. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1191-9.

NOMENCLATURE

B	Magnetic induction (T)	t	Treatment time (min)
n	Rotation speed (min^{-1})	S	Longitudinal feed (mm/min)
Ra	Surface roughness (μm)		

1. INTRODUCTION

The magnetic-abrasive treatment (MAT) method is a technology that does not apply the traditional principles of chip removal. Turning, drilling holes, milling, and grinding are traditional chip removal techniques that use turning tools, milling cutters, grinding wheels, and other cutting tools to remove material from the workpiece. Tool wear is inevitable when using these methods, and the tool material must be harder than the material being processed (1). Processing high-strength materials requires significant shear forces, and the processing speed is inversely proportional to the strength of the material (2). At the same time, the release of a significant amount of heat in the cutting area limits the processing speed. Manufacturing of small-size cutting tools is difficult, which, in turn, limits the processing capabilities of small parts. Therefore, it is difficult or impossible to process small parts using traditional methods.

Magnetic-abrasive treatment is one of the modern methods of finishing materials, which is actively used to improve surface quality (3, 4). This method is based on the use of a magnetic field to control abrasive particles, which allows for high precision and uniformity of processing. Permanent magnet devices play a key role in this process, providing a stable magnetic field necessary for effective control of abrasive particles (5, 6).

Magnetic-abrasive treatment plays an important role in ensuring the quality of, for example, soldered joints of flat-oval finned pipes, which are widely used in heat exchangers. Heat exchangers are used in various industries, including the energy sector, the chemical industry, the oil and gas industry, and heating, ventilation, and air conditioning systems (7-9). It is especially important to ensure the quality of heat exchange devices in harsh climatic conditions (10-12). Their main function is to efficiently transfer heat between two or more media, which optimizes heating or cooling processes (13).

Flat-oval finned pipes, due to their design, provide a large heat exchange surface, which helps to increase the efficiency of heat exchangers (14, 15). However, the quality of the soldered joints of such pipes is critically important to ensure the reliability and durability of the heat exchanger. The use of the magnetic abrasive treatment method makes it possible to improve the surface quality in the area of the soldered joint by removing oxides and smoothing out irregularities that negatively affect the strength and tightness of the soldered or welded joint (16).

Soldering of products made of aluminum and its alloys is also complicated by the presence of an oxide film on the surface of the products, which has a melting point several times higher than the metal itself and remains as non-metallic inclusions at the root of the seam. When the aluminum surface that has just been processed on a milling machine interacts with the atmosphere, it is immediately covered with an oxide film, which tends to recover. To create a reliable welded joint, it is necessary to carefully choose the operation of pre-preparation of products before welding (17, 18).

2. LITERATURE REVIEW

Today, with the development of technology, expectations of precision products and their processing methods are growing (19, 20). To meet this need, researchers are developing new surface treatment methods, especially in the mining, aerospace, electronics, and medical industries (21). The treatment of abrasive particles in a magnetic field is a new and developing method that removes an allowance from the surface using abrasive and magnetic particles under the influence of a magnetic field applied to the treatment site (Figure 1) (22-24). This method ensures a high level of surface quality. In the MAT method, shear forces are controlled using a magnetic field, and the abrasive powders that polish the surface have a flexible structure, unlike the one-piece cutting tools used in traditional methods such as grinding and honing (25-27). This avoids the occurrence of stresses on the surface and eliminates microcracks, which can significantly reduce the strength of the material.

As a result of the development of technological requirements, Erden and Cerit drew the attention of engineers to more modern production methods (3). After World War II, significant efforts were made to develop new methods of final surface treatment, and the first non-traditional methods appeared between 1950 and 1970. The development of the MAT method, as an unconventional method, began in the specified period of time. Thanks to advances in materials science, materials with exceptional properties have been created that could not be processed by traditional methods due to their high strength (28). Therefore, there was a need for the MAT method (3, 29).

MAT method was first developed and applied in the Soviet Union, and then it was developed in Japan, which is still happening today (16, 30, 31). A flexible surface treatment tool is created by mixing ferromagnetic and abrasive particles in certain proportions (Figure 2). The ferromagnetic particles in the mixture keep the abrasive particles under pressure, allowing them to penetrate into the processing area and remove a certain layer of metal

(31, 32). Iron powder is usually used as magnetic particles, and materials such as silicon carbide (SiC), aluminum oxide (Al_2O_3), and boron nitride are used as abrasives (32, 33).

Chang et al. (32) added a lubricant that leaves a visible trace to ferromagnetic abrasive powders by mixing them mechanically, and investigated the effect of the prepared mixture during processing of stainless steels by tracking the path of particles. As a result, it was found that the most significant factors affecting the degree of chip removal and surface roughness are the working clearance during processing and the speed of rotation of the workpiece (32). Khairy (25) studied the change in surface roughness during the processing of steels using a mixture of sintered aluminum oxide (Al_2O_3) and iron powders. Kwak (33) investigated the effect of a mixture consisting of silicon carbide abrasives and ferromagnetic particles together with SAE 30 lubricant on the quality of cylindrical outer surfaces after MAT (34).

Lack of automatization in the MAT process is the reason that MAT is not used in industry on a large scale. MAT requires a closed control loop with a feedback system, which will provide better control of output variables by adjusting the operating parameters of the process based on feedback signals (35).

Based on the results of the analysis of scientific literature, it becomes clear that the efficiency of MAT for finishing and the rate of material removal for difficult-to-process materials such as ceramics, superalloys and glass is low. Therefore, it is necessary to consider increasing the efficiency of MAT for these materials by hybridization with other processes, for example, electrochemical and chemical.

The effectiveness of magnetic abrasive treatment depends on many physical and chemical factors. These include, for example, the force acting on individual abrasive grains (Figure 3). One of the important aspects of magnetic abrasive finishing is the creation of a high magnetic flux density in the finishing area, which significantly affects the force acting on individual ferromagnetic abrasive grains and helps to improve the finishing process (33).

3. PROBLEM STATEMENT

The main task in magnetic abrasive processing is to achieve an optimal combination of magnetic field parameters and abrasive particle characteristics to ensure maximum process efficiency.

Preparation for magnetic abrasive treatment includes the choice of the strength of the magnetic field, the speed of movement of the particles of the mixture and the processing time. An important aspect is also the study of the influence of various parameters on the final surface quality, including roughness and structural changes of the material (35).

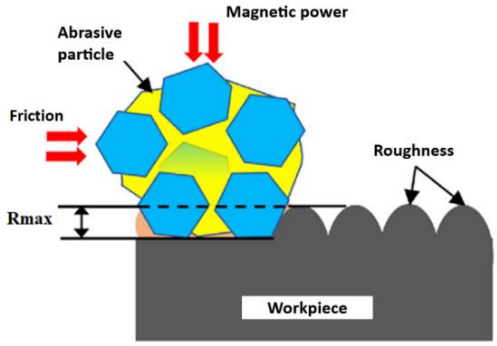


Figure 1. Removal of the material during magnetic abrasive treatment

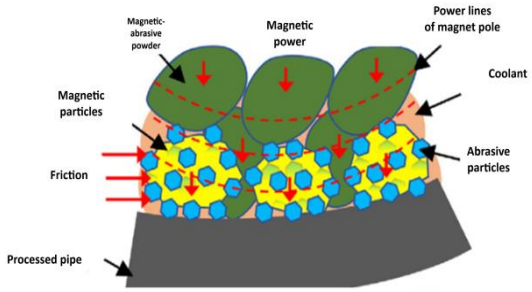


Figure 2. Interaction of grain groups

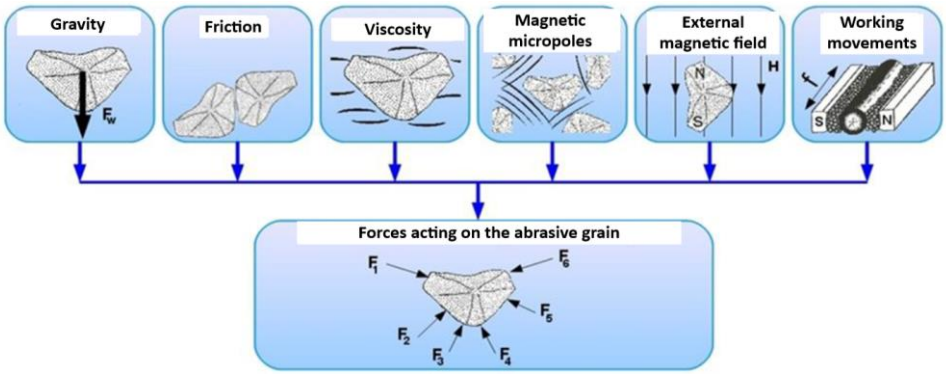


Figure 3. Forces acting on abrasive grain

To assess the effect of the processing mode parameters: magnetic induction B, the rotation speed of the device n, the processing time t and the feed of the workpiece S, the task was set to carry out magnetic abrasive treatment of a flat-oval tube made of aluminum alloy A93003 (AlMn).

A device similar to an end mill with permanent magnets arranged radially is usually positioned relative to the flat surface being processed by a working gap of about 2.5 mm (Figure 4). When the device passes along the processed plane of the flat-oval pipe, the surface is sanded, to which the fin plates (comb) will be soldered later.

Before carrying out magnetic abrasive treatment, it is necessary to simulate the distribution of the magnetic field, since with the help of modeling it is possible to determine the optimal parameters of the magnetic field, such as its intensity and distribution, which contributes to the implementation of processing for a specific selected device.

The uniform distribution of the magnetic field helps to avoid defects such as uneven machining or surface damage, which is especially important for difficult-to-process materials such as aluminum A93003 (AlMn).

Modeling makes it possible to minimize the cost of materials and the time required for experiments by first determining the best processing conditions (2, 29, 36).

The described device and the flat-oval tube were modeled in the KOMPAS 3D software package, then the models were placed in the workspace of the ANSYS Maxwell software package to evaluate the uniformity of the magnetic field distribution in the working gap (Figures 5 and 6). In this case, the following boundary conditions are accepted: the material of the workpiece is aluminum alloy A93003 (AlMn); the material of the pole tips is steel AISI 1017; the material of the magnets is NdFe10 neodymium magnet; the working gap is 2.5 mm.

In this case, the magnetic field distribution was estimated; for these magnets, the induction value B reached only 0.1 T, which is a small value for magnetic abrasive treatment. However, taking into account the rotation of the device on the order of 200-300 min^{-1}, it is assumed that the field will be uniform and processing will be carried out using more powerful NdFe40 and NdFe45 magnets.

The developed MAT approach can benefit from recent methodologies in pipeline monitoring and thermal analysis (37) for enhanced quality control in dissimilar material joints.

4. RESULTS AND DISCUSSION

Experimental studies were carried out, according to the results of which sections of a flat-oval pipe were treated using a device for magnetic abrasive treatment with radially arranged permanent magnets of the NdFe30, NdFe35, NdFe40, NdFe45, NdFe50 grades. Thus, the following ranges of values were selected as controlled processing parameters: magnetic induction B from 0.6 to 1.2 T, processing time t from 4 to 12 minutes, rotation speed of the device n from 115 to 460 min^{-1}, longitudinal feed of the workpiece S from 25 to 225 mm/min. The controlled parameter was the roughness of the treated flat surface Ra. Combinations of operating parameters and roughness results for parameter Ra are shown in Table 1.

Based on the results of experimental studies, graphs of the dependence of surface roughness Ra on magnetic induction B, processing time t, device rotation speed n and longitudinal feed S are drawn (Figure 7). Response surfaces have also been compiled, demonstrating the

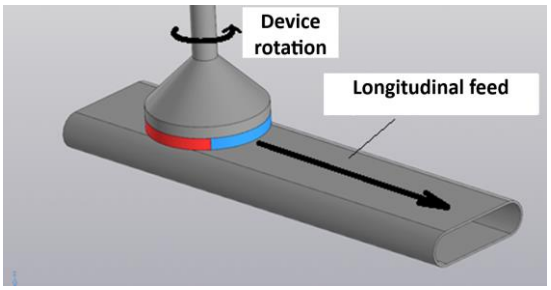


Figure 4. Device operation diagram

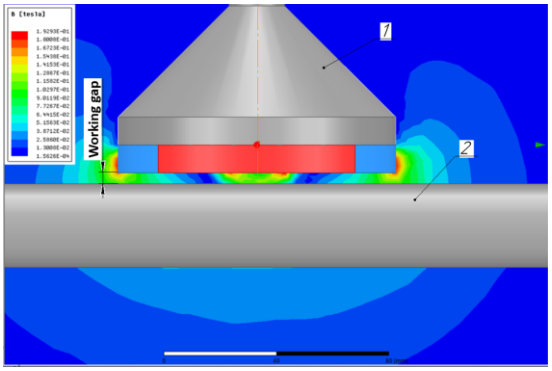


Figure 5. Distribution of the magnetic field in the working gap between the magnets of the device and the flat–oval tube: 1 - device; 2 – workpiece

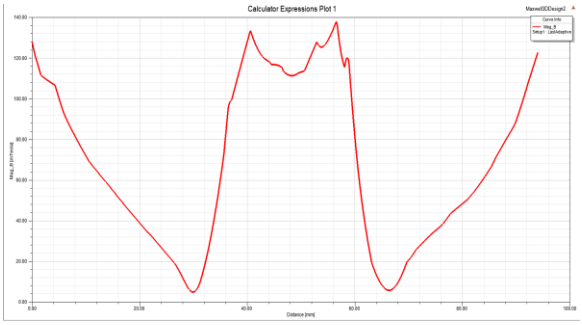


Figure 6. Graph of the magnetic induction distribution along the midline of the working gap

TABLE 1. Matrix of experimental research

Experiment	Magnetic induction B, T	Treatment time t, min	Rotation speed n, min^{-1}	longitudinal feed S, mm/min	Surface roughness Ra, μm
1	0,8	10	460	75	1,002
2	1,0	6	460	75	0,642
3	0,8	6	460	75	0,815
4	1,0	10	190	75	0,994
5	0,8	10	190	75	0,965
6	1,0	6	190	75	1,013
7	0,8	6	190	75	1,12
8	1,2	8	300	125	0,923
9	0,6	8	300	125	1,31
10	0,9	12	300	125	1,095
11	0,9	4	300	125	0,635
12	0,9	8	750	125	0,608
13	0,9	8	115	125	0,715
14	0,9	8	300	225	0,818
15	0,9	8	300	25	0,889
16	0,9	8	300	125	0,884
17	0,9	8	300	125	0,852
18	0,9	8	300	125	0,791
19	0,9	8	300	125	0,903
20	0,9	8	300	125	0,848
21	0,9	8	300	125	0,824
22	0,9	8	300	125	0,814

effect of a combination of these parameters on surface roughness (Figure 8).

A comparative analysis was conducted between MAT and manual grinding for finishing aluminum alloy A93003 surfaces. The results indicate that MAT provides more consistent surface quality Ra 0.6-1.0 μm compared to manual grinding Ra 1.2-2.0 μm, while achieving 3-4 times higher productivity. Unlike manual grinding, MAT eliminates operator-dependent variability and ensures uniform treatment of complex geometries.

Our studies revealed that surface finish consistency varies by geometry complexity. For standard components flat/cylindrical surfaces, Ra variation remained within ±0.1 μm (94% consistency). However, in complex geometries (gears, internal grooves), field gradients caused localized under-treatment Ra up to 1.5 μm in concave areas versus 0.8 μm on convex surfaces. This was mitigated by implementing adaptive magnet arrays and dynamic gap control.

For industrial applications, the MAT system can be adapted in two ways: 1) integration into existing machine tool spindles for processing small-to-medium components (<0.5 m^2), or 2) development of specialized machines with enhanced magnetic systems for large surfaces (>0.5 m^2). While current NdFe magnets effectively handle typical industrial part sizes, scaling beyond 1 m^2 requires either electromagnet arrays or multi-head configurations to maintain uniform abrasive brush behavior across extended surfaces.

The study employed STATISTICA 12 for experimental design (rotatable central composite design) and data analysis, including regression modeling and ANOVA. All mathematical models (Figures 7 and 8, Table 2) were derived and validated using this software.

Most of the available scientific sources and technical literature that address MAT issues are devoted to the finishing of flat, internal and external curved surfaces of cylindrical workpieces. However, the processing of surfaces of arbitrary shape using MAT remains less studied insufficiently. The processing of machine-building industry parts with internal grooves and complex shapes, such as gears and molds, using MAT is

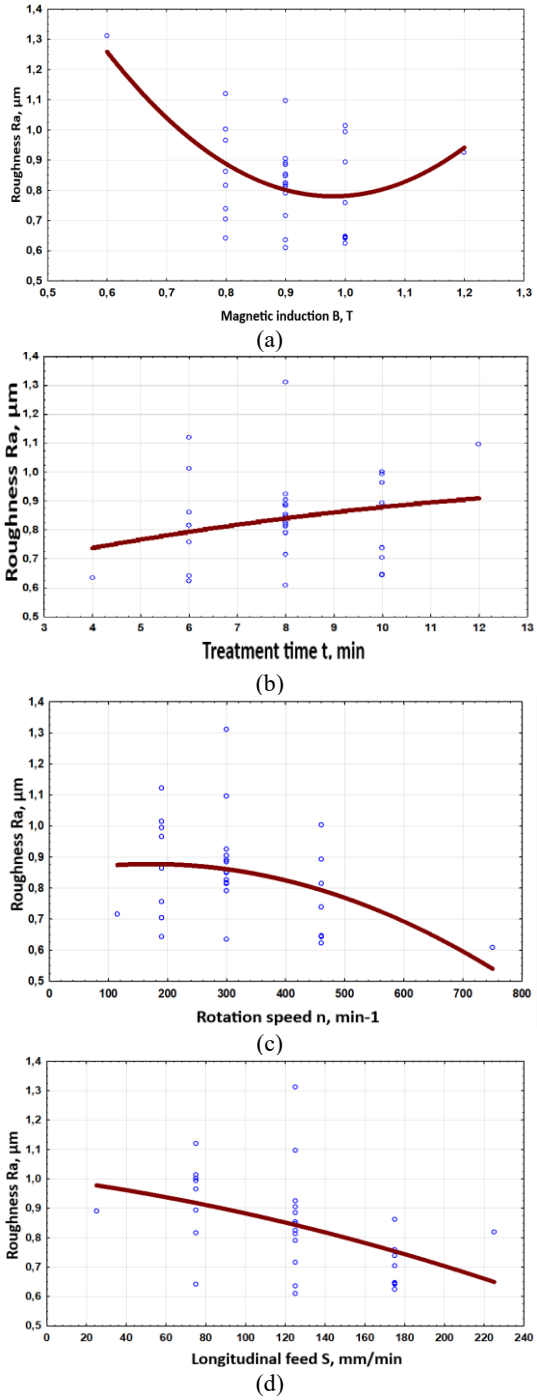


Figure 7. Graphs of the dependence of the roughness of the treated surface on the processing parameters: a) magnetic induction B; b) processing time t; c) rotation speed of the device n; d) longitudinal feed S

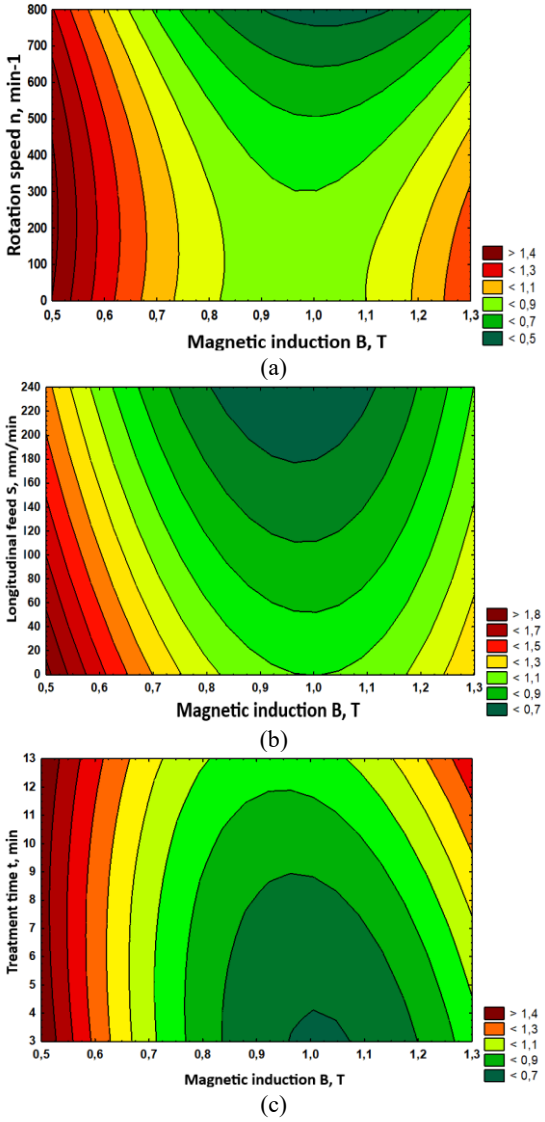


Figure 8. The surface roughness response surface Ra depends on: (a) magnetic induction B and rotational speed n; (b) magnetic induction B and longitudinal feed S; (c) magnetic induction B and processing time

also insufficiently researched. In addition, the processing of surfaces of arbitrary complex shapes while maintaining a constant working gap and the processing of the entire surface using software on numerically controlled machines have been studied even less. It should also be noted that there are practically no algorithms for the development and validation of devices for magnetic abrasive processing.

The results of experimental studies of magnetically abrasive treatment of a flat surface made of A93003 (AlMn) aluminum show that the use of permanent magnets in devices for MAT allows for high stability of the magnetic field, which, in turn, ensures uniform distribution of the brush of magnetically abrasive particles over the treated flat surface. This leads to a significant improvement in the processing quality, a reduction in roughness and an increase in the wear resistance of the material. When processing aluminum surfaces, roughness decreases to values close to mirror ones, which is especially important for parts operating under friction conditions. In addition, the use of

permanent magnets reduces the energy consumption of the process, since there is no need for a constant supply of electric current to create a magnetic field.

For the considered ranges of MAT parameters of a flat surface, patterns of the influence of parameters on surface roughness are obtained, which are presented in Equations 1-7 (Table 2).

TABLE 2. Mathematical dependencies

$Ra = 3{,}3986 - 6{,}5096 \cdot B + 3{,}3225 \cdot B^2$	(1)
$Ra = 0{,}601 + 0{,}0385 \cdot t - 0{,}0011 \cdot t^2$	(2)
$Ra = 0{,}08476 + 0{,}0003 \cdot n$	(3)
$Ra = 1{,}0026 - 0{,}0009 \cdot S$	(4)
$Ra = 3{,}7048 - 5{,}9731 \cdot B + 0{,}0006 \cdot n + 3{,}1121 \cdot B^2 - 0{,}0005 \cdot B \cdot n$	(5)
$Ra = 4{,}339 - 6{,}6955 \cdot B - 0{,}003 \cdot S + 3{,}3546 \cdot B^2 + 0{,}001 \cdot B \cdot S$	(6)
$Ra = 4{,}3395 - 7{,}0348 \cdot B - 0{,}0567 \cdot t + 3{,}3782 \cdot B^2 + 0{,}0531 \cdot B \cdot t + 0{,}0019 \cdot t^2$	(7)

5. CONCLUSION

Magnetic-abrasive treatment using permanent magnet devices is an effective method of finishing flat surfaces, ensuring high quality and precision of processing. The advantages of this method include both the stability of the magnetic field and the uniform distribution of abrasive particles in the working gap, as well as a significant reduction in energy consumption. In the future, we should expect the development of MAT technologies, which will expand the scope of the technology and increase the efficiency of processing various materials, including those that are difficult to process.

The main problems in the processing of compounds from dissimilar materials by the MAO method include:

- Differences in the coefficients of thermal expansion of materials.
- Unequal magnetic permeability of the components.
- The risk of selective processing (one material is processed more intensively than the other).
- Problems with the adhesion of abrasive particles to different materials.

Our method solves these problems through:

- Optimization of the magnetic abrasive composition.
- Magnetic field control.

This study outlines several critical directions for future research: 1) comprehensive comparative analyses between MAT and alternative finishing techniques (laser polishing, chemical-mechanical polishing); 2) extension to complex geometries (gear teeth, spherical surfaces) and internal channels requiring specialized validation algorithms; 3) development of industrial-scale equipment; 4) long-term performance evaluation under operational conditions (wear, thermal cycling, corrosion); and 5) hybrid process development (MAT-ultrasonic/EDM combinations) for superalloys and ceramics. These investigations will build upon the current methodology focused exclusively on external surface treatment.

6. REFERENCES

1. Yungmeister D, Lavrenko S, Yacheikin A, Urazbakhtin RY. Improving the shield machine cutter head for tunneling under the conditions of the Metrostroy Saint Petersburg mines. ARPN Journal of Engineering and Applied Sciences. 2020;15(11):1282-8.
2. Orekhova TN, Sivalneva MN, Frolova MA, Strokova VV, Bondarenko DO. The effect of mechanical and thermal treatment on the characteristics of saponite-containing material. Записки Горного института. 2025(272 (eng)):91-9.
3. CHAWLA G, MITTAL VK, MITTAL S. Experimental investigation of process parameters of al-sic-b4c mmcs finished by a novel magnetic abrasive flow machining setup. Walailak Journal of Science and Technology (WJST). 2021;18(18):9885 (13 pages)- (13 pages). 10.48048/wjst.2021.9885
4. Maksarov V, Maksimov D, Sinyukov M. Quality control of complex contour surfaces in aluminium alloy items during magnetic abrasive finishing. Tsvetnye Metall. 2023;4:96-102. 10.17580/tsm.2023.04.13
5. Jadhav S, Sawant S. Dry sliding behaviour study of novel low-metallic friction materials by using DoE-Taguchi method. International Journal of Engineering. 2024;37(3):452-9. 10.5829/IJE.2024.37.03C.01
6. Safari M, Joudaki J, Rezaei M. Investigation of Laser Cutting of Thin Polymethyl Methacrylate Sheets by Response Surface Methodology. International Journal of Engineering Transactions C: Aspects. 2024;37(3):484-95. 10.5829/ije.2024.37.03c.05
7. Shibanov D, Ivanov S, Yemelyanov A, Pumpur E. Evaluation of working efficiency of open pit shovels in real operating conditions. MIAB Mining Inf Anal Bull. 2020(10):86-94. 10.25018/0236-1493-2020-10-0-86-94
8. Wu J, Zou Y, Sugiyama H. Study on ultra-precision magnetic abrasive finishing process using low frequency alternating magnetic field. Journal of Magnetism and Magnetic Materials. 2015;386:50-9. 10.1016/j.jmmm.2015.03.041
9. Zhukov I, Martyushev N, Zyukin D, Azimov A, Karlina A. Modification of hydraulic hammers used in repair of metallurgical units. Metallurgist. 2023;66(11):1644-52. 10.1007/s11015-023-01480-w
10. Nikolaev AK, Zaripova NA. Substantiation of analytical dependences for hydraulic calculation of high-viscosity oil transportation. Записки Горного института. 2021;252:885-95. 10.31897/PMI.2021.6.10
11. Yushkova EA, Lebedev VA. Enhancement of energy efficiency of the vacuum oil distillation unit using pinch analysis. Записки Горного института. 2023(261 (eng)):415-27.
12. Pryakhin E, Sharapova D. Repair ability of low-alloyed steel strength of K70 (X90) class. Key Engineering Materials. 2020;836:131-5. 10.4028/www.scientific.net/KEM.836.131
13. Zherlygina E, Kuranova M, Gusev V, Odintsov E. Identification of hazardous sites based on studying the development of man-made fractures within the rock mass. Russ Min Ind. 2025;1:162-9. 10.30686/1609-9192-2025-1-162-169
14. Singh P, Singh L, Singh S. Manufacturing and performance analysis of mechanically alloyed magnetic abrasives for magneto

abrasive flow finishing. Journal of manufacturing processes. 2020;50:161-9. https://doi.org/10.1016/j.jmapro.2019.12.033

15. Singh DK, Jain V, Raghuram V. Experimental investigations into forces acting during a magnetic abrasive finishing process. The International Journal of Advanced Manufacturing Technology. 2006;30(7):652-62. 10.1007/s00170-005-0118-6
16. Baron YM. Magnetic abrasive deburring technology for blanks. International Journal of Engineering Research in Africa. 2016;25:1-10.
17. Barman A, Das M. Simulation and experimental investigation of finishing forces in magnetic field assisted finishing process. Materials and Manufacturing Processes. 2018;33(11):1223-32. 10.1080/10426914.2018.1453157
18. Bolobov VI, Popov GG. Methodology for testing pipeline steels for resistance to grooving corrosion. Записки Горного института. 2021;252:854-60. 10.31897/PMI.2021.6.7
19. Ivanov SL, Ivanova PV, Kuvshinkin SY. Promising model range career excavators operating time assessment in real operating conditions. Записки Горного института. 2020;242:228-33. 10.31897/pmi.2020.2.228
20. Litvinenko VS, Petrov EI, Vasilevskaya DV, Yakovenko AV, Naumov IA, Ratnikov MA. Assessment of the role of the state in the management of mineral resources. Записки Горного института. 2023(259 (eng)):95-111. 10.31897/PMI.2022.100
21. Gabov V, Zadkov D, Babyr N, Xie F. Nonimpact rock pressure regulation with energy recovery into the hydraulic system of the longwall powered support. Eurasian Min. 2021;36:55-9. 10.17580/em.2021.02.12
22. Singh DK, Jain V, Raghuram V. Parametric study of magnetic abrasive finishing process. Journal of materials processing technology. 2004;149(1-3):22-9. 10.1007/978-3-030-43312-3_8
23. Salmanian A. PM, Rezaei M. . Investigating the influence of the effective parameters on the tooth geometry of the spherical involute bevel gears through solving Napier's equations. International Journal of Engineering Transactions A: Basics. 2026;39(1):44-58. 10.5829/ije.2026.39.01a.05
24. Singh RK, Gangwar S, Singh D. Experimental investigation on temperature-affected magnetic abrasive finishing of aluminum 6060. Materials and Manufacturing Processes. 2019;34(11):1274-85. 10.1080/10426914.2019.1628263
25. Khairy AB. Aspects of surface and edge finish by magnetoabrasive particles. journal of materials processing technology. 2001;116(1):77-83. 10.1016/S0924-0136(01)00840-8
26. Kumar R, Komma VR. Development and experimental investigation of magnetic abrasive finishing of 17-4PH steel. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science. 2024;238(10):4463-76. 10.1080/10426914.2021.2016822
27. Pandey K, Pandey U, Pandey PM. Statistical modeling and surface texture study of polished silicon wafer Si (100) using chemically assisted double disk magnetic abrasive finishing. Silicon. 2019;11(3):1461-79. 10.1007/s12633-018-9961-6
28. Ko S, Baron YM, Park J. Micro deburring for precision parts using magnetic abrasive finishing method. Journal of Materials Processing Technology. 2007;187:19-25. 10.1016/j.jmatprotec.2006.11.183
29. Girma B, Joshi SS, Raghuram M, Balasubramaniam R. An experimental analysis of magnetic abrasives finishing of plane surfaces. Machining science and Technology. 2006;10(3):323-40. 10.1080/10910340600902140
30. Sharma VK. Modeling and analysis of a novel rotational magnetorheological abrasive flow finishing process. International Journal of Lightweight Materials and Manufacture. 2021;4(3):290-301. 10.1016/j.ijlmm.2021.02.001
31. Vasiliev A. S. GAA. Special strategy of treatment of difficultyprofile conical screw surfaces of single-screw compressors working bodies. Journal of Mining Institute. 2019;235: 60-4. 10.31897/PMI.2019.1.60
32. Chang G-W, Yan B-H, Hsu R-T. Study on cylindrical magnetic abrasive finishing using unbonded magnetic abrasives. International Journal of Machine Tools and Manufacture. 2002;42(5):575-83. 10.1016/S0890-6955(01)00153-5
33. Kwak J-S. Enhanced magnetic abrasive polishing of non-ferrous metals utilizing a permanent magnet. International Journal of Machine Tools and Manufacture. 2009;49(7-8):613-8. 10.1016/j.ijmachtools.2009.01.0
34. Mishra V, Burada DR, Karar V, Manna A, Jha S, Khan GS. Investigations on flexible pad polishing for nano-finishing of freeform optics mold. Journal of Micromanufacturing. 2020;3(2):99-112. 10.1177/2516598420939740
35. Alam Z, Iqbal F, Ganesan S, Jha S. Nanofinishing of 3D surfaces by automated five-axis CNC ball end magnetorheological finishing machine using customized controller. The International Journal of Advanced Manufacturing Technology. 2019;100(5):1031-42. 10.1007/s00170-017-1518-0
36. Chen Y, Zheng Y, editors. Design and Characteristics Analysis of Electromagnets Based on ANSYS Maxwell. 2024 3rd International Symposium on Semiconductor and Electronic Technology (ISSET); 2024: IEEE. 10.1109/ISSET62871.2024.10779827
37. Major Z, Bodnár L, Merczel DB, Szép J, Lublóy É. Analysis of the heating of steel structures during fire load. 2024. 10.28991/ESJ-2024-08-01-01

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

بااین مقاله به چالش حیاتی دستیابی به پرداخت سطح با کیفیت بالا برای اتصالات مواد غیر مشابه می‌پردازد که برای کاربردهای هوافضا، انرژی و مهندسی دقیق ضروری است. روش‌های مرسوم اغلب به دلیل تفاوت در خواص مواد، نتایج ثابتی ارائه نمی‌دهند. ما یک فرآیند عملیات مغناطیسی-سایشی (MAT) بهینه شده با استفاده از آهنرباهای دائمی NdFe 0.6-1.2 T و ساینده‌های کامپوزیتی Fe/SiC توسعه دادیم. نتایج تجربی نشان می‌دهد که روش ما به یکنواختی سطح برتر Ra 0.6-1.0 µm با ۲-۳ برابر بهره‌وری بالاتر در مقایسه با سنگ‌زنی دستی دست می‌یابد. یافته‌های کلیدی شامل تعیین پارامترهای بهینه B = 1.0 T، n = 460 min-1 است. نوآوری این کار در بهینه‌سازی سیستماتیک MAT برای اتصالات مواد غیر مشابه و اولین نمایش مقیاس‌پذیری صنعتی برای قطعات تا ۰.۵ متر مربع نهفته است. این نتایج کاربردهای MAT را گسترش می‌دهد و جایگزینی مقرون به صرفه برای روش‌های پرداخت موجود ارائه می‌دهد.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1200-1210

International Journal of Engineering

Journal Homepage: www.ije.ir

Predicting Optimum Project by Using Structural Equation Modeling

S. T. Y. AL-Qaicy*

Civil Engineering Department, University of Technology, Iraq

PAPER INFO

Paper history:
Received 16 April 2024
Received in revised form 20 May 2025
Accepted 06 July 2025

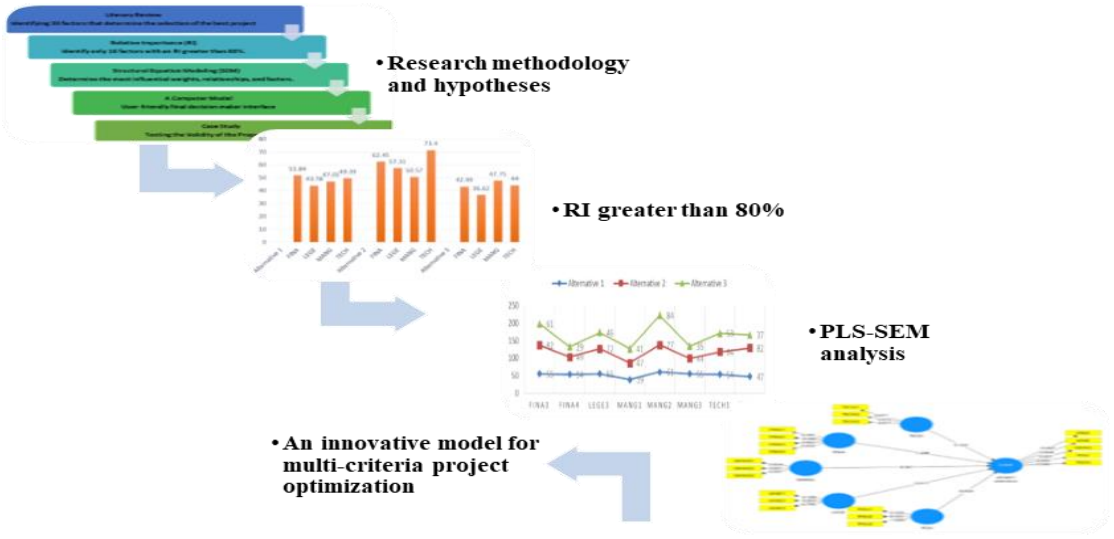
Keywords:
Optimum Project
Structural Equation Modeling
Partial Least Squares

ABSTRACT

Developing a tool to predict the best project according to approved criteria is essential for the success of a project. Traditional assessment models often fall short in capturing the interdependencies among multiple variables. This study fills that gap by offering a predictive model-based framework throw create a simplified and practical interactive model for multi-criteria decision-making. The following steps were taken to overcome the above problems: 1) 33 factors were extracted from open sources and interviews with decision makers. 2) 16 factors achieved a relative importance higher than 80%. 3) Only eight factors were chosen from the Structural Equation Modeling (SEM) to determine weights and relationships having t-values for regression weights were smaller than 1.96, and the R-square was more significant than 0.05, and 4) a mathematical model was built based on the most influential factors, regardless of the lowest bid. To verify the validity of the developed model, the case study was applied to three alternatives (public sector projects under implementation). Four main criteria were consistent with the study's hypothesis, while the POLI criteria were excluded as a major determinant because it had a regression weight greater than 1.96. The results showed the third alternative, was not the best choice, however, the second project is considered the best among them with (medium) degree, while the first and third projects obtained a (weak) degree with a result of 60.44%, 48.01%, and 42.84%, respectively, with goodness of PLS modelwith an R-square of 0.943.

***doi**: 10.5829/ije.2026.39.05b.13*

Geographical Abstract



*Corresponding Author Email: surratalal99@gmail.com (S. T. Y. AL-Qaicy)

Please cite this article as: AL-Qaicy STY. Predicting Optimum Project by Using Structural Equation Modeling. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1200-10.

1. INTRODUCTION

1. 1. Background Choosing a construction project is a crucial decision for investors, as the choice directly impacts the project's success or failure. For selecting the most appropriate project, the AHP, TOPSIS, and VIKOR methods were used, while the ARAS method was rarely used. Construction companies are expected to follow one of the methods of analyzing and selecting projects due to their effectiveness and accurate selection (1). A comprehensive evaluation model was created based on the resulting information, which can be used for an unlimited number of projects, taking into account the specificities of the construction industry in Iraq. The proposed model emphasized the importance of the financial plan and its sound management as one of the most important criteria for project selection. According to the study, the model did not select the least expensive project, but rather the one that fully complied with the basic aspects of the evaluation criteria (2, 3). Applying the Fuzzy Utility Additions (UTASTAR) model for supplier selection is an integrated model for decision-making in multi-criteria projects. The results of the UTASTAR model have been compared and proven effective in selecting suppliers to ensure sustainable development (4). Selecting the best project according to the multi-criteria decision-making technique is a global approach that clarifies the criteria and influencing factors according to the region and surrounding conditions, ensuring the appropriate decision is reached (5-8). Decision-making is a recurring process in our daily lives. However, with development, more criteria and controls govern this process, creating decision-making one of the most challenging topics. However, decision-makers agree that evaluating and ranking alternatives is a fundamental feature of different decision-making methods (9-11). This study aims to create a simplified and practical interactive model to fill the gap for multi-criteria decision-making, by offering a predictive model-based SEM- PLS. Consisting of several primary and secondary criteria for selecting the best project.

1. 2. Study Questions and Hypothesis The study questions revolved around:

- Are there specific criteria governing the selection of a construction project?
- How can these criteria be determined?

Present study assumes that the following criteria have the greatest impact on selecting the best construction project: technical (TECH), financial (FINA), managerial (MANA), legislative (LEGE), and political (POLI). The PLS-SEM technique was used to identify the influential criteria through standardized regression weights for each criterion.

2. MATERIALS AND LITERATURE REVIEW

2. 1. Factors Affecting Project Selection and Decision-Making Determining the best project is a process that requires several factors to ensure the project is executed to the highest quality within each project's parameters. A comprehensive database of similar projects must be available, a thorough analysis of available data must be conducted. The potentials must be evaluated, alternatives must be researched and compared. Then, the modern methods and techniques must be employed (1). Two combined approaches were used to reach the best project selection: the Fuzzy Delphi Technique (FDT) and the Analytical Network Process (ANP) to select the influencing factors and determine the priorities and bilateral relationships between them. This approach yields the best results, providing a comprehensive view of the efficiency of the project components: structural/organizational, technical, and financial (2). Supply chains are so important that they can be considered a key criterion for project selection in terms of chain integration, time and cost savings, use of modern technologies, speed and reliability of operations, and support for sustainable development goals (4). Due to the complexity of digital data representation, Zadeh (12) introduced fuzzy sets (fuzzy or imprecise data). Many studies have described various fuzzy sets, such as hesitant FS, circular intuitionistic FS, spherical FS, linguistic intuitionistic FS, and others, which aim to improve decision-making in the face of numerous challenges, complexities, and uncertainties. This has led to the development of new cognitive structures that support decision-making (8, 13, 14). The factors influencing the decision to choose the best multi-criteria project are summarized as follows (2):

- Integration of project information and documents.
- A common platform for proper communication
- An appropriate organizational structure
- Crystallize the final idea
- Using modern technologies
- Effective financial risk management
- Detailed financial plans

Uncertainty, incomplete data, and gaps between criteria relationships often challenge the Iraqi construction sector. Fuzzy sets, on the other hand, provide a valuable framework for dealing with ambiguity and imprecision, making them particularly suitable for MCDM analysis.

2. 2. Structural Equation Modeling Multiple regression, logistic regression, and analysis of variance are considered first-generation data analysis and classification techniques, relying on fundamental methods for testing relationships between variables and

significantly contributing to the development of subsequent techniques. Many areas of scientific research routinely deal with theoretical concepts, such as perceptions, attitudes, and intentions, so the limitations of first-generation techniques are fundamental (15-17). The model is evaluated through the following steps:

- Examining the structural model to identify potential collinearity issues.
- Estimating coefficients using ordinary least squares (OLS) regressions for internal consistency.
- Evaluating high levels of collinearity between structural models.
- Evaluating the significance and reliability of path coefficient relationships.
- Examining the explanatory power and predictive ability of the model.
- Calculating and comparing alternative models generated from different theories or contexts to facilitate PLS-SEM analysis.

Structural equation modeling (SEM) is a sophisticated statistical technique for identifying and analyzing relationships between variables in different structures. This logic combines the principles of factor analysis and multiple regression analysis. Variables take two forms in structural equations: 1) endogenous variables (outcomes) and 2) exogenous variables (causes). The advantages of PLS-SEM are (16, 18, 19):

- The method does not assume a normal distribution of data, therefore, it can handle heterogeneous data.
- It can handle multi-criteria relationships.
- It reduces outliers and multicollinearity.
- It provides high accuracy and precision for research hypotheses.

2. 3. Multi-Criteria Project Decision Making The multi-criteria decision model effectively contributes to sound decision-making. It is a structured reasoning process for determining project acceptability indicators by identifying, evaluating, and analyzing criteria. The proposed models strongly support senior management and decision-makers in determining project selection (20-23). The criteria related to internal and external processes in the construction industry in India and tested for applicability (22). Project evaluation models often succeed because they consider system dynamics and analyze them according to primary and secondary criteria. When deciding on a project, the optimal response to all emerging changes is considered, and the success of the proposed model is measured by its ability to adapt to internal and external factors (24, 25). Several criteria are involved in evaluating or selecting the best project, including technical, financial, administrative, political, legislative, and other criteria. Each includes sub-criteria. The proposed models seek to balance the importance of the requirements and the degree of their application within the project. Decision-making models may address the gap between planning and implementation by determining the importance of each criterion relative to the other criteria and their direct and indirect relationships (26-31). Decision-making methods vary from one project to another according to several imposed or established determinants. The experience of decision-makers is considered the primary determinant of evaluation, depending on the method used, including AHP, ANP, AMOS, FUZZY AHP, PLS-SEM, TOPSIS Techniques (32, 33).

3. RESEARCH METHODOLOGY

To achieve the research goal, three approaches are used with the following details:

1. Examining the most important factors affecting project selection according to their relative importance (1, 6).
2. Finding the weights of the criteria using Structural equation modeling (SEM) (15-17).
3. Create a computer model using Excel to choose the best project (2). Figure 1 shows the research methodology followed in this study.

3. 1. First: Relative Importance to Identify and Test the Factors Affecting Project Selection After a comprehensive literature review of academic articles and interviews with decision makers, including project managers and construction industry professionals, in Table 1. In fact, 33 primary and secondary factors influencing the selection of the best project were identified.

3. 1. 1. Data Collecting and Sampling Data for this study were collected using a structured questionnaire

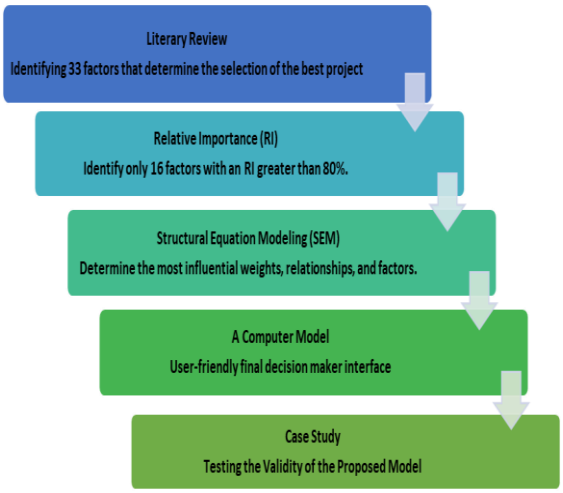


Figure 1. Research methodology

TABLE 1. Main and sub-criteria

No	Factors	Ref.
	Technical	
1	Detailed feasibility studies (environmental, economic, social, etc.)	(1-3)
2	Using modern technologies in planning and monitoring work progress	(1-3)
3	Alternative plans based on potential risks	(1-3)
4	Using modern tools and equipment in the work	(4)
5	Developing a workflow plan that clarifies the timeframe for each activity	(5, 6)
6	Continuous monitoring, oversight, and feedback	(5, 7)
7	Clarity of project sections and identification of those responsible for each section	(2, 3)
	Financial	
1	Detailed financial feasibility study	(4, 5, 8)
2	Effective financial management that takes into account the experiences of similar projects	(8)
3	Continuous financial oversight throughout the project with transparent access to reports	(4, 5, 8)
4	A documented financial plan that clarifies funding sources	(3, 8, 9)
5	A financial plan for expected risks	(5, 12)
6	Identifying project funding sources	(1, 2, 4)
7	A timetable for disbursing project payments	(8, 16)
8	Integration of project financial components	(4, 5)
9	Adopting a safety and environmental approach	(7, 11)
	Managerial	(15, 17)
1	Pre-qualification of technical and administrative teams	(11, 14)
2	Continuous coordination among all stakeholders	(8)
3	Integration of administrative components	(2, 5, 15)
4	Educating senior management	(11, 14)
5	Identifying challenges and ways to address them	(4, 8)
6	Documenting all project documents and ensuring their access to relevant parties for review	(2, 6)
	Legislative	
1	Compliance with the country's governing laws and regulations	(2, 3)
2	Proper approvals for foreign labor:	
3	Obtaining proper approvals before commencing work	(4, 5, 8)
4	A legal team reviews all types of contracts	(8)
5	Adhering to laws and regulations regarding imports, equipment, and labor	(4, 5, 8)
	Political	
1	An alternative plan for political fluctuations	(5, 12)
2	The project must not conflict with existing policies in the country	(7, 11)
3	Adapting to political obstacles	(15, 17)
4	The project must comply with the country's policies and orientations according to the phase.	(11, 14)
5	Defining project objectives and their alignment with the country's general policy	(8, 12, 13)
6	Priority is given to employing citizens and attracting experienced and qualified personnel	(2, 5, 6)

designed to capture the necessary variables relevant to the research objectives. A total of 240 questionnaires were distributed among targeted respondents. These respondents were selected using purposive sampling, focusing on individuals with relevant experience and expertise to ensure the data's validity and relevance, given the research topic's specialized nature.

Only 205 were returned, yielding a response rate of approximately 85.4%. After initial screening, 196 questionnaires were found to be complete and valid for analysis. The selection of participants was based on predefined inclusion criteria, such as professional background, years of experience, and involvement in project management, to ensure alignment with the study's objectives.

The data were analyzed using IBM SPSS Statistics Version 20. To identify key factors, the analysis in Table 2 focused on calculating the Relative Importance (RI) index for each item. Items with an RI greater than 80% were considered significant and retained for interpretation. This threshold was chosen based on prior literature and methodological standards in similar studies to highlight the most impactful factors.

TABLE 2. Criteria having RI greater than 80%

No	Factors	Coding
	Technical	
1	Detailed feasibility studies (environmental, economic, social, etc.)	TECH1
2	Using modern technologies in planning and monitoring work progress	TECH2
3	Continuous monitoring, oversight, and feedback plans	TECH3
	Financial	
1	Detailed financial feasibility study	FINA1
2	Effective financial management that takes into account the experiences of similar projects	FINA2
3	Continuous financial oversight throughout the project, with transparent access to reports	FINA3
4	A financial plan for expected risks	FINA4
	Managerial	
1	Pre-qualification of technical and administrative teams	MANG1
2	Continuous coordination among all stakeholders	MANG2
3	Documenting all project documents and ensuring their access to relevant parties for review	MANG3
	Legislative	
1	Compliance with the country's governing laws and regulations	LEGE1
2	A legal team reviews all types of contracts	LEGE2
3	Adhering to laws and regulations regarding imports, equipment, and labor	LEGE3
	Political:	
1	The project must not conflict with existing policies in the country	POLI1
2	The project must comply with the country's policies and orientations according to the phase.	POLI2
3	Priority is given to employing citizens and attracting experienced and qualified personnel.	POLI3

The questionnaire form consists of two parts: the first includes information about the respondents, and the second provides questions to determine the concepts and indicators that have been imposed. The questionnaire form is designed to be answered as follows: 1 strongly disagree, 2 Disagree, 3 Unsure, 4 Agree, 5 strongly agree. The questionnaire included 66 civil engineers (34%), 38 architects (19%), 23 MEP engineers (12%), 21 electrical engineers (11%), 32 surveyors (16%), and 16 others (technicians and administrators) (8%). The responses showed that 83 respondents (42%) had more than 20 years of experience, 69 respondents (35%) had 10-20 years of experience, and 44 respondents (23%) had less than 10 years of experience. The data were analyzed by using SPSS-20 to allocate the factors with RI greater than 80%. The analysis of the experiences indicates that the respondents have extensive experience, reflecting the careful selection of the research sample to suit the importance of the study. To assess the internal consistency of the questionnaire, Cronbach's alpha was calculated. The overall Cronbach's alpha is 0.93.

3. 2. Second: Weights of Criteria Using Structural Equation Modeling SEM According to the conceptual framework, the structural equation model was used to examine the data and choose the best project. The following programs were used:

- IBM- SPSS Statistics 20 to enter, analyze, and examine the data.
- Partial Least Squares Technique- Smart PLS 3.0) to test the relationships between the factors. Measurement models reflect a direct relationship between the construct and the indicators, and the indicators are treated as error-prone coefficients of the underlying construct. Equation 1 illustrates the relationship between a latent variable and its observed indicators (34):

$$x = l.y + e \quad (1)$$

Where:

X is the observed indicator variable, Y is the latent variable, l is the regression coefficient quantifying the strength of the relationship between X and Y, and e represents the random measurement error. The presence of structural error in the causal indicators can be equivalent to the conceptual variable. If the indicators x1, x2, and x3 represent causal indicators, then the error term Y1, z1, Equation 2 will describe these other "causes." This model can be expressed as (34):

$$Y = \sum_{K=1}^{K} Wk.Xk + Z \quad (2)$$

Where wk represents the contribution of xk (k = 1,..., K) to Y, and z is an error term associated with Y. The measurement model using composite in Equation 3 indicators takes the following form, where Y is a linear combination of indicators xk (k = 1,..., K), each weighted by a specific weight as stated below (34):

$$Y = \sum_{K=1}^{1} Wk.Xk \quad (3)$$

PLS path modeling was chosen instead of AMOS because it explains the variance between the internal variables to reach the maximum value. Structural models in PLS path modeling are evaluated by finding t-values for regression weights greater than 1.96, where independent factors with regression weights less than the specified value of t-value are removed and considered not to support the proposed model. Figure 3 shows the exclusion of the main criterion (POLI) that did not meet the conditions. R-squared value for predictive power is determined to judge the factors and evaluate the predictive model comprehensively, which should not exceed 0.05 (17, 34). For the direct model (shown in Figure 2), the regression weight of the assumed structure and the predictive power of the model (R-square) are shown in Table 2. Figure 2 shows the adopted criteria.

3. 3. Third: Integration Model for Best Project Selection To facilitate the selection of the best project, a simplified model was created in Excel with a simple final user interface that includes the criteria weights, with expert opinion required to determine the final value of each criterion based on its specific application for each project. Table 3 and Equation 4 illustrates the formula used in programming the user interface.

$$Best\ Project = MAX[\sum_{i=1}^{m}(RI\ sub\ factors * score\ of\ each\ criteria)/m] \quad (4)$$

where m = number of sub-factors for each main criterion.

The proposed model considers the lack of familiarity of the end user with the software or techniques used to access the factors of interest, as the technique is one of the methods of decision support.

4. CASE STUDY MODEL

To verify the validity and effectiveness of the proposed model, a case study was applied to evaluate three government projects implemented by the public sector. These projects were selected for the following reasons:

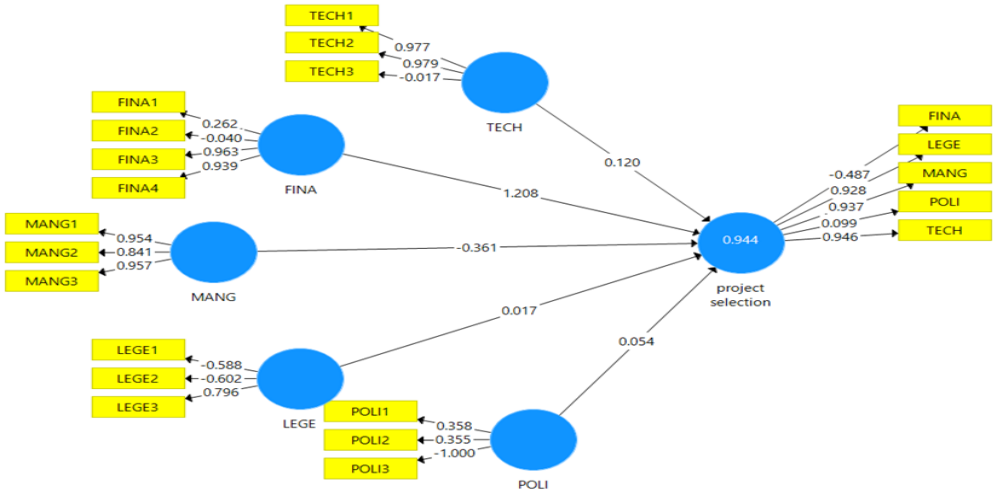


Figure 2. Assessment of the direct structural model

TABLE 3. The regression weight of the proposed structure

CRITERIA	FINA	LEGE	MANG	TECH
FINA3	0.963			
FINA4	0.939			
LEGE3		0.796		
MANG1			0.954	
MANG2			0.841	
MANG3			0.957	
TECH1				0.977
TECH2				0.979
R- square	0.943			

1) still under construction and funded from the 2023 operating budget, 2) recently selected for implementation, and 3) There is an apparent delay in some projects, 4) implemented by the Iraqi public sector, which reflects a model of the precision of project selection in Iraq with delay rates and estimated costs and 5) The public sector's reliance on the lowest bids. Table 4 summarizes the information for each alternative that was later used in the case study.

Three projects were selected and coded as Alternatives 1, 2, and 3. The three alternatives were re-evaluated by specialists, with each project assigned a score from 100 degrees (as a standardized scale) for each sub-factor. The feasibility of each alternative was re-

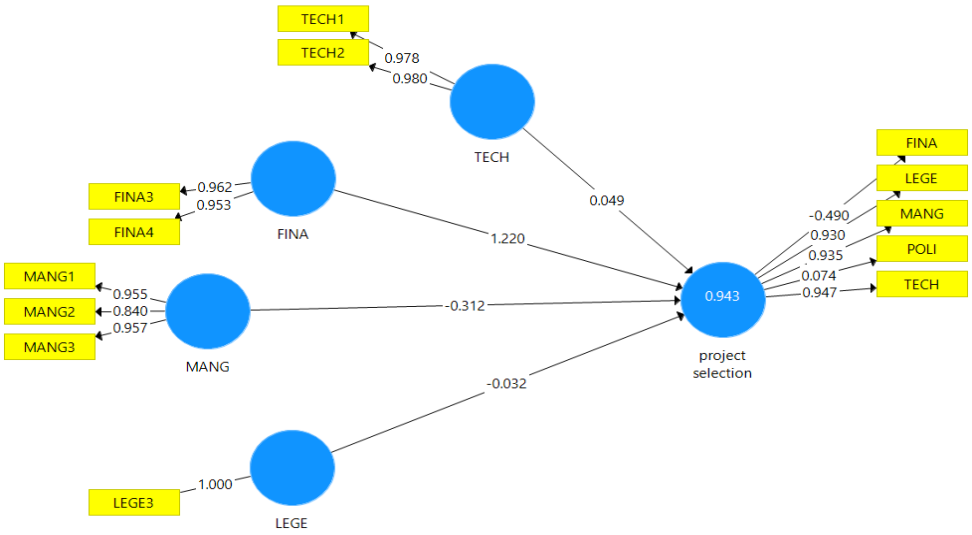


Figure 3. Assessment of the direct structural model after correction

TABLE 4. Details of case study (alternatives)

No.	Project Name	Provence	Deviation rate	Cost (IQD)
Alt.1	Ministry of Planning, Buildings Complex	Baghdad	17%	54,116,864,345
Alt.2	Rehabilitation of Jalawla General Hospital	Diyala	2%	9,286,000,000
Alt.3	Zurbatiyah Border Crossing	Wasit	37%	75,918,166,660

examined, taking into account aspects of integration and assigning it the appropriate score. Table 5 shows the results of analyzing the three alternatives.

5. RESULTS AND DISCUSSION

5. 1. Influential and Excluded Criteria Selecting and implementing construction projects based on scientific and practical principles is essential for mitigating challenges. 33 sub-criteria were identified and divided into five main criteria: technical (TECH), financial (FINA), managerial (MANA), legislation (LEGE), and political (POLI). Only 16 sub-criteria with an RI more significant than 80% were selected; using PLS, eight influential criteria were identified to create the optimal project selection model. In the PLS application, the POLI criterion was excluded from the requirements assumed in the research. Experts supported the exclusion of the POLI criterion due to its weak impact on locally implemented projects, specifically in the public sector. Other sub-criteria that did not meet the minimum thresholds were excluded, and only eight criteria were retained with an R-square of 0.943. The TECH criterion achieved significant superiority over the other criteria, with a weight of 0.947, as it included 1) Detailed feasibility studies (environmental, economic, social, etc.); 2) using modern technologies in planning and monitoring work progress; 3) Continuous monitoring, oversight, and feedback plans, followed by the MANG criterion, with a weight of 0.935, which included: 1) Pre-qualification of technical and administrative teams; 2) Continuous coordination among all stakeholders; 3) Documenting all project documents and ensuring their access to relevant parties for review, following this was the LEGE criterion, with a weight of 0.93, which emphasized the enactment of fair laws for awarding bids, particularly selecting the lowest-cost bid.

5. 2. The Case Study Application The proposed model was applied to three public sector projects currently under implementation. The results of the application varied between the three alternatives. The following standard classification was adopted to determine the project's importance: 0-20 (very weak), 21-40 (weak), 41-60 (medium), 61-80 (good), 81-100 (very good). After excluding POLI and its sub factors, eight factors were retained and are shown in Figure 4 with their standard deviation weights.

Figure 5 represents the results of applying the mathematical model to the secondary criteria.

Based on the results of the secondary factors, Figure 6 shows the results of the main factors, which demonstrate the superiority of the second alternative as the best choice among the three.

5. 2. 1. Main Criteria for Each Project The TECH criterion had a very high impact in the second alternative, less than in the first alternative, and the least impact in the third one, 71.4% (good), 49.39% (medium), and 44% (medium), respectively. In the same order, the results of the FINA criterion analysis were in favor of the second alternative, the first, and then the third: 62.45% (good), 51.84% (medium), and 42.99% (medium), respectively. Regarding the third main criterion, MANG, the second alternative obtained 50.57% (medium), the first alternative with 47.05% (medium), and the third one

TABLE 5. Results of the case study for the best project selection

Results of the case study for the best project selection				
Alternative 1				
CRITERIA	**Regression weight**	**criteria score**	**Sub-criteria result**	**Main criteria result**
FINA3	0.963	55	52.965	51.836
FINA4	0.939	54	50.706	
LEGE3	0.796	55	43.78	43.780
MANG1	0.954	39	37.206	47.047
MANG2	0.841	61	51.301	
MANG3	0.957	55	52.635	
TECH1	0.977	54	52.758	49.386
TECH2	0.979	47	46.013	
The final score of Alternative 1				**48.012**
Alternative 2				
CRITERIA	**Regression weight**	**criteria score**	**Sub-criteria result**	**Main criteria result**
FINA3	0.963	82	78.966	62.489
FINA4	0.939	49	46.011	
LEGE3	0.796	72	57.312	57.312
MANG1	0.954	47	44.838	50.568
MANG2	0.841	77	64.757	
MANG3	0.957	44	42.108	
TECH1	0.977	64	62.528	71.403
TECH2	0.979	82	80.278	
The final score of Alternative 2				**60.443**
Alternative 3				
CRITERIA	**Regression weight**	**criteria score**	**Sub-criteria result**	**Main criteria result**
FINA3	0.963	61	58.743	42.987
FINA4	0.939	29	27.231	
LEGE3	0.796	46	36.616	36.616
MANG1	0.954	41	39.114	47.751
MANG2	0.841	84	70.644	
MANG3	0.957	35	33.495	
TECH1	0.977	53	51.781	44.002
TECH2	0.979	37	36.223	
The final score of Alternative 3				**42.839**

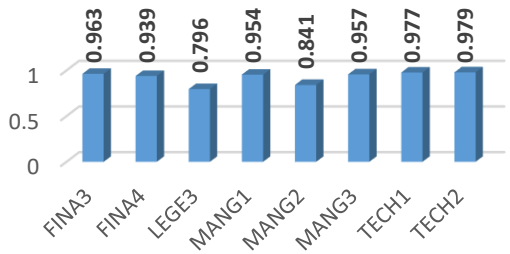


Figure 4. Weight of sub-factors

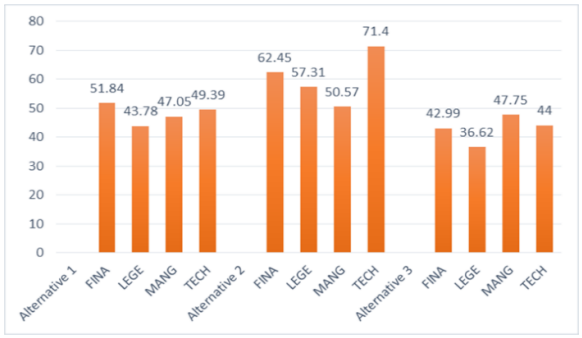


Figure 7. Comparison between alternatives

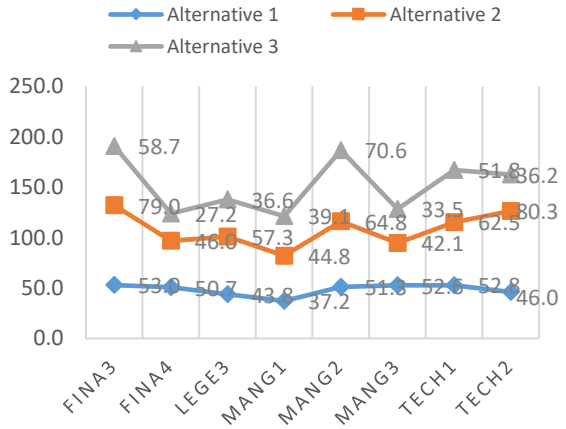


Figure 5. Sub-criteria results

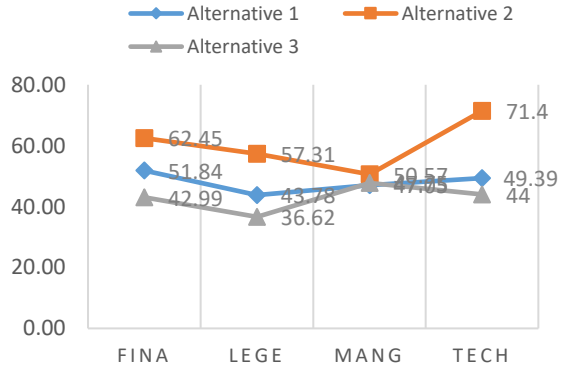


Figure 6. Main criteria results

47.75% (medium). The last main criterion was LEGE, and the second alternative was in the lead, followed by the first and then the third at 57.31% (medium), 43.78% (medium), and 36.62% (poor), respectively, Figure 7. Shows the comparison between alternatives.

5. 2. 2. The Alternatives The three alternatives are under construction and were previously selected as the best projects among the submitted alternatives. The three projects suffer from delays and delays in completion rates as specified in Table 4, according to the statistics of the Iraqi Ministry of Housing, Construction, and Public Works, 2025 data. According to the PLS analysis, the second alternative received the highest evaluation according to the deviation weights. However, its importance as the best alternative is still medium, with a rate of 60.44%. As for the first and third alternatives, their level of selection as the best project was medium, with a rate of 48.01% and 42.84%, respectively. The above results explain and analyze the reasons for the delays of the projects: rehabilitation of Jalawla General Hospital with a delay of 2% only, while the Ministry of Planning Buildings Complex project and the Zurbatiyah Border Crossing project with delays reached 17% and 37%, respectively. Although the second alternative received the highest scores, it remains weak as a preferred alternative. The criteria mentioned in the research should be taken into account when preparing feasibility studies and selecting projects to avoid wasted time and cost. The innovative structural model of PLS analysis provides an easy-to-use tool to overcome major challenges that hinder the successful delivery of construction projects in Iraq.

6. CONCLUSION

Selecting the best project based on pre-established criteria that take into account the country's changing factors and circumstances has proven effective in project evaluation. Selecting the best project from among several may not be a simple task, as it involves numerous challenges, such as maintaining costs and project delivery time, as well as quality and optimal project management. Identifying the most influential criteria highlights the focus for multi-criteria decision-making. It is recommended:

- The work team must be highly experienced to assess projects accurately.

- Learn from the experiences of neighboring countries with similar circumstances.
- Adopting a structured decision support system that includes the approved SEM model to prioritize decisions according to the criteria defined in this study.
- Engaging stakeholders:
- Applying the Fuzzy Risk Management Model (MCDM) to project bidding and evaluation.
- Investing in project planning, cost estimation, and conflict resolution programs to enhance project implementation outcomes.

This study examines the critical challenge of accurately predicting the optimum characteristics of successful projects by employing Structural Equation Modeling (SEM). The research introduces a statistically robust approach to identifying, validating, and quantifying the key factors that contribute to project success. For future studies, it is recommended to use the proposed model for more specialized projects and to compare alternatives or specific departments, in addition to applying the model to projects under construction as a tool for evaluating projects and determining delay rates.

7. REFERENCES

1. Yikici O, Ozcelik TO. Investment project selection with multi criteria decision making techniques in FMCG industry. International Journal of Engineering and Management Research. 2023;13(3):116-33. https://doi.org/10.31033/ijemr.13.3.15
2. AL-Qaicy S, Abdullah K. Integration Management Decision Support Model for Best Project Selection. International Journal of Engineering, Transactions B: Applications. 2024;37(11):2333-43. https://doi.org/10.5829/ije.2024.37.11b.18
3. Ansari R, Banihashemi S, Taherkhani R, Moradi S. Decision support system for analyzing key performance indicators in construction projects management. International Journal of Engineering. 2022;35(5):865-74. https://doi.org/10.5829/ije.2022.35.05b.03
4. Hakimi M, Daneshvar A, Ehsanifar M, Nouri I. Applying Fuzzy Utilities Additives STAR for Sustainable Supplier Selection in Iran's Construction Industry. International Journal of Engineering. 2024;37(10):2008-20. https://doi.org/10.5829/ije.2024.37.10a.12
5. Mousavi S, Gitinavard H, Vahdani B. Evaluating construction projects by a new group decision-making model based on intuitionistic fuzzy logic concepts. 2015. https://doi.org/10.5829/idosi.ije.2015.28.09c.08
6. Davoudabadi R, Mousavi SM, Šaparauskas J, Gitinavard H. Solving construction project selection problem by a new uncertain weighting and ranking based on compromise solution with linear assignment approach. Journal of Civil Engineering and Management. 2019;25(3):241-51. https://doi.org/10.3846/jcem.2019.8656
7. Ali J, Ali W, Alqahtani H, Syam MI. Enhanced EDAS methodology for multiple-criteria group decision analysis utilizing linguistic q-rung orthopair fuzzy hamacher aggregation operators. Complex & Intelligent Systems. 2024;10(6):8403-32. https://doi.org/10.1007/s40747-024-01586-x
8. Ali G, Afzal A, Sheikh U, Nabeel M. Multi-criteria group decision-making based on the combination of dual hesitant fuzzy sets with soft expert sets for the prediction of a local election scenario. Granular computing. 2023;8(6):2039-66. https://doi.org/10.1007/s41066-023-00414-w
9. Ali J, Bashir Z, Rashid T. A cubic q-rung orthopair fuzzy TODIM method based on Minkowski-type distance measures and entropy weight: J. Ali et al. Soft Computing. 2023;27(20):15199-223. https://doi.org/10.1007/s00500-023-08552-8
10. Eslamipoor R. A fuzzy multi-objective model for supplier selection to mitigate the impact of vehicle transportation gases and delivery time. Journal of Data, Information and Management. 2022;4(3):231-41. https://doi.org/10.1007/s42488-022-00076-4
11. Ali J. Norm-based distance measure of q-rung orthopair fuzzy sets and its application in decision-making. Computational and Applied Mathematics. 2023;42(4):184. https://doi.org/10.1007/s40314-023-02313-x
12. Zadeh LA. Fuzzy sets. Information and control. 1965;8(3):338-53. https://doi.org/10.1016/S0019-9958(65)90241-X
13. Jana C, Dobrodolac M, Simic V, Pal M, Sarkar B, Stević Ž. Evaluation of sustainable strategies for urban parcel delivery: Linguistic q-rung orthopair fuzzy Choquet integral approach. Engineering Applications of Artificial Intelligence. 2023;126:106811. https://doi.org/10.1016/j.engappai.2023.106811
14. Liu P, Naz S, Akram M, Muzammal M. Group decision-making analysis based on linguistic q-rung orthopair fuzzy generalized point weighted aggregation operators. International Journal of Machine Learning and Cybernetics. 2022;13(4):883-906. https://doi.org/10.1007/s13042-022-01560-5
15. Becker J-M, Ringle CM, Sarstedt M. Estimating Moterating effects in PLS-SEM andPLSc-SEM: interaction term gerneration* data treatment. Journal of Applied Structural Equation Modeling. 2018(2):1-21. https://doi.org/10.47201/jasem.2018.02.02.01
16. Lee S-Y. Structural equation modeling: A Bayesian approach: John Wiley & Sons; 2007.
17. Henseler J, Fassott G. Testing moderating effects in PLS path models: An illustration of available procedures. Handbook of partial least squares: Concepts, methods and applications: Springer; 2009. p. 713-35.
18. Schreiber JB. Core reporting practices in structural equation modeling. Research in social and administrative pharmacy. 2008;4(2):83-97. https://doi.org/10.1016/j.sapharm.2007.10.002
19. Vermunt JK, Magidson J. Structural equation models: Mixture models. Encyclopedia of statistics in behavioral science: Wiley; 2005. p. 1922-7.
20. Stoianova A, Trofimets V, Stoianova O, Matrokhina K. Structural model of decision support system for sustainable development of oil and gas companies. International Journal of Engineering, Transactions A: Basics. 2025;38:701-9. https://doi.org/10.5829/ije.2025.38.04a.03
21. Hakimi M, Daneshvar A, Ehsanifar M, Nouri I. Applying Fuzzy Utilities Additives STAR for Sustainable Supplier Selection in Iran's Construction Industry. International Journal of Engineering, Transactions A: Basics. 2024;37(10):2008-20. https://doi.org/10.5829/ije.2024.37.10a.12
22. Mousavi S, Gitinavard H, Vahdani B. Evaluating construction projects by a new group decision-making model based on intuitionistic fuzzy logic concepts. International Journal of Engineering, Transactions C: Aspects. 2015;28 (9):1312-9. https://doi.org/10.5829/idosi.ije.2015.28.09c.08
23. Dorfeshan Y, Mousavi SM, Vahdani B, Mohagheghi V. Solving critical path problem in project network by a new enhanced multiobjective optimization of simple ratio analysis approach

with interval type-2 fuzzy sets. International Journal of Engineering, Transactions C: Aspects. 2017;30(9):1352–61. https://doi.org/10.5829/ije.2017.30.09c.03

24. Goodarzi H, Ehsanifar M, Mirhosseini S, Mazaheri H. A Combined Modeling Approach for Delay Management in Construction Projects Based on a Two-objective Mathematical Model and System Dynamics. International Journal of Engineering, Transactions A: Basics. 2025;38(4):945–63. https://doi.org/10.5829/ije.2025.38.04a.22
25. Selcuk O, Turkoglu H, Polat G, Hajdu M. An integrative literature review on the causes of delays in construction projects: evidence from developing countries. International Journal of Construction Management. 2024;24(6):610-22. https://doi.org/10.1080/15623599.2022.2135939
26. Ajayi BO, Chinda T. Dynamics of pertinent project delay variables in the Thai construction sector: mathematical analysis. Engineering Management in Production and Services. 2022;14(2). https://doi.org/10.2478/emj-2022-0014
27. Monirabbasi A, Ramezani Khansari A, Majidi L. Simulation of delay factors in sewage projects with the dynamic system approach. Industrial Engineering and Strategic Management. 2021;1(1):15-30. https://doi.org/10.22115/IESM.2020.232300.1006
28. Shamshiri F, Shahnazari-Shahrezaei P, Fallah M, Kazemipour H. Fuzzy Dynamic Modeling for Export Consortia in Small and Medium-Sized Enterprises. International Journal of Engineering Transactions A: Basics. 2024;37(07):1422. https://doi.org/10.5829/ije.2024.37.07a.19
29. Shi L, Peng, B. . Optimization of Operating Parameters of a Scroll Expander Based on Response Surface and NSGA II. International Journal of Engineering Transactions A: Basics. 2024;37(1):136–50. https://doi.org/10.5829/ije.2024.37.01a.13
30. Ansari R, Banihashemi S, Taherkhani R, Moradi S. Decision support system for analyzing key performance indicators in construction projects management. International Journal of Engineering, Transactions B: Applications. 2022;35(5):865-74. https://doi.org/10.5829/ije.2022.35.05b.03
31. Awasthi A, Chauhan SS. A hybrid approach integrating Affinity Diagram, AHP and fuzzy TOPSIS for sustainable city logistics planning. Applied Mathematical Modelling. 2012;36(2):573-84. https://doi.org/10.1016/j.apm.2011.07.033
32. Aragonés-Beltrán P, Chaparro-González F, Pastor-Ferrando J-P, Pla-Rubio A. An AHP (Analytic Hierarchy Process)/ANP (Analytic Network Process)-based multi-criteria decision approach for the selection of solar-thermal power plant investment projects. Energy. 2014;66:222-38. https://doi.org/10.1016/j.energy.2013.12.016
33. Sablatzky T. The delphi method. Hypothesis: Research Journal for Health Information Professionals. 2022;34(1). https://doi.org/10.18060/26224
34. Sarstedt M, Ringle CM, Hair JF. Partial least squares structural equation modeling. Handbook of market research: Springer; 2021. p. 587-632.

COPYRIGHTS

©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

توسعه ابزاری برای پیش‌بینی بهترین پروژه بر اساس معیارهای تأیید شده برای موفقیت یک پروژه ضروری است. مدل‌های ارزیابی سنتی اغلب در ثبت وابستگی‌های متقابل بین متغیرهای متعدد، کوتاهی می‌کنند. این مطالعه با ارائه یک چارچوب مبتنی بر مدل پیش‌بینی، این شکاف را پر می‌کند و یک مدل تعاملی ساده و کاربردی برای تصمیم‌گیری چند معیاره ایجاد می‌کند. مراحل زیر برای غلبه بر مشکلات فوق انجام شد: ۱) ۳۳ عامل از منابع آزاد و مصاحبه با تصمیم‌گیرندگان استخراج شد. ۲) ۱۶ عامل اهمیت نسبی بالاتر از ۸۰٪ را به دست آوردند. ۳) تنها هشت عامل از مدل‌سازی معادلات ساختاری (SEM)برای تعیین وزن‌ها و روابط با مقادیر t برای وزن‌های رگرسیون کوچکتر از ۱.۹۶ و R-square معنادارتر از ۰.۰۵ انتخاب شدند و ٤) یک مدل ریاضی بر اساس تأثیرگذارترین عوامل، صرف نظر از کمترین پیشنهاد، ساخته شد. برای تأیید اعتبار مدل توسعه‌یافته، مطالعه موردی بر روی سه گزینه (پروژه‌های بخش دولتی در حال اجرا) اعمال شد. چهار معیار اصلی با فرضیه مطالعه سازگار بودند، در حالی که معیارهای POLI به عنوان یک عامل تعیین‌کننده اصلی حذف شدند زیرا وزن رگرسیونی آنها بیشتر از ۱.۹٦ بود. نتایج نشان داد که گزینه سوم بهترین انتخاب نبود، با این حال، پروژه دوم با درجه (متوسط) بهترین در بین آنها در نظر گرفته می‌شود، در حالی که پروژه‌های اول و سوم به ترتیب با نتیجه ٦۰.٤٤٪، ٤۸.۰۱٪ و ٤۲.۸٤٪ درجه (ضعیف) کسب کردند و مدل PLS با ضریب تعیین ۰.۹٤۳ برازش داشت.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1211-1225

International Journal of Engineering

Journal Homepage: www.ije.ir

An Improved Control Chart Pattern Classification using a Transfer Learning-Based VGG-16 Network

S. Nazari*[a], F. Sogandi[b]

[a] Faculty of Electrical Engineering, Shahid Beheshti University, Tehran, Iran
[b] Department of Industrial Engineering, University of Torbat Heydarieh, Torbat Heydarieh, Iran

PAPER INFO

Paper history:
Received 20 February 2025
Received in revised form 06 July 2025
Accepted 07 July 2025

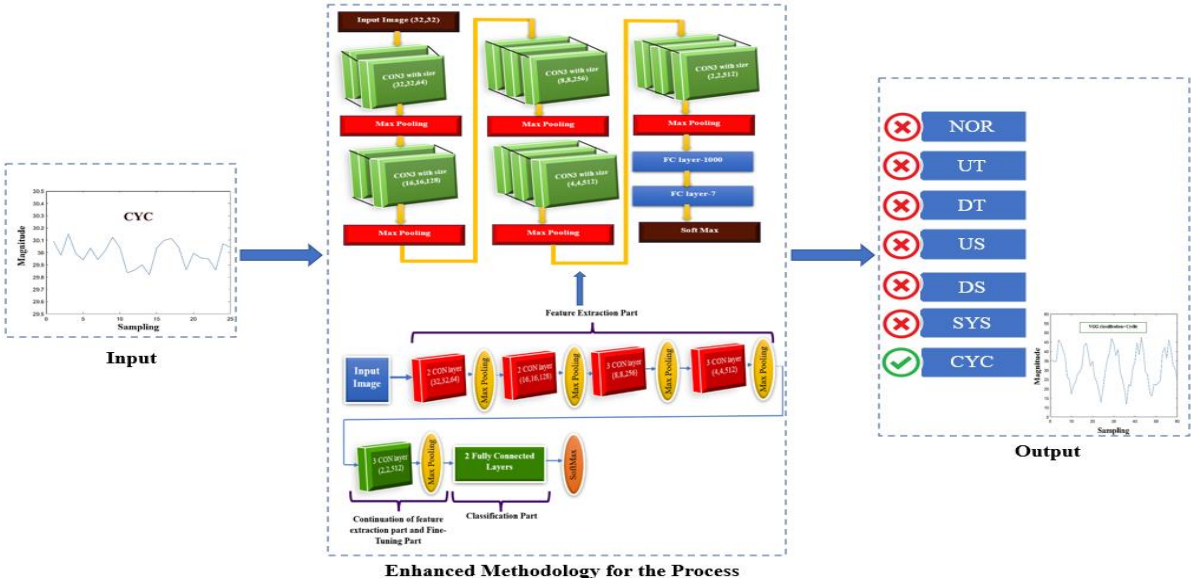
Keywords:
Pattern Classification
VGG-16 Network
Deep Transfer Learning
Control Chart

ABSTRACT

Control chart pattern recognition is a crucial statistical process control tool used to determine whether a process operates within intended parameters or exhibits abnormal behavior. Accurate and automated recognition is essential for manufacturing companies to maintain high-quality production. Recent studies have combined machine learning with control chart pattern recognition, yet these methods often assume identical feature spaces and statistical distributions between training and test data—a condition rarely met in practice. Retraining models from scratch with new data is time-consuming. Transfer learning offers a more efficient alternative by leveraging knowledge from pre-trained models. This study employs the pre-trained VGG-16 deep convolutional neural network to classify control chart patterns, significantly improving recognition performance without extensive retraining. Through Monte Carlo simulations and a real-world case study, the proposed method achieved a recognition accuracy of 99.3%, outperforming conventional MLP, 1D-CNN, and CNN models. Sensitivity analysis demonstrated that the use of ReLU activation, batch normalization, and dropout techniques substantially enhanced accuracy and model robustness. The results confirm the potential of the proposed transfer learning-based recognizer to improve intelligent quality control in manufacturing processes by effectively handling limited training data and complex feature extraction.

***doi**: 10.5829/ije.2026.39.05b.14*

Graphical Abstract



*Corresponding Author Email: soheilanazari21@yahoo.com (S. Nazari)

Please cite this article as: Nazari S, Sogandi F. An Improved Control Chart Pattern Classification using a Transfer Learning-Based VGG-16 Network. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1211-25.

1. INTRODUCTION

Statistical Process Control (SPC) are some techniques that involve monitoring and evaluating different stages of the manufacturing process to maintain a stable and acceptable level, ensuring that products and services meet predetermined standards. Control charts are a crucial tool in SPC. When a control chart exhibits a natural pattern, it shows that the process is in-control state. Conversely, an unnatural pattern signals an out-of-control state due to at least an assignable cause. There are different types of non-random patterns that can appear in-control charts, and various analytical tools have been created to help identify these patterns and their underlying assignable causes. However, the accuracy of identifying unnatural patterns using run rules is dependent on the practitioner's experience and skill, and the identification rate is not satisfactory. Moreover, although these methods can identify abnormalities indicating out-of-control conditions, they do not provide information about the specific non-random pattern that occurred .

Traditional control charts in manufacturing processes have a significant limitation: they only utilize the information from the most recent plotted point, disregarding potentially valuable data about disturbances contained in previous samples. Incorporating pattern recognition into the control chart can address this challenge. By knowing the type of Control Chart Pattern (CCP) and the corresponding process factors that could cause it, the diagnosis search can be expedited. Hence, timely recognition of CCPs is crucial for identifying potential assignable causes (1). There are several common examples of assignable causes that can lead to the appearance of abnormal, non-random patterns in manufacturing control charts such as operator fatigue. Figure 1 illustrates seven common CCPs. A normal pattern is one where the control chart only includes common causes, showing that the process is in-control state.

Machine learning techniques, especially neural networks and deep learning models, offer powerful tools for automated control chart pattern recognition. They enable accurate classification of a wide range of patterns, including subtle and complex ones, and can adapt effectively to variable-length and multivariate data sequences. Additionally, these methods are robust in handling autocorrelation and dynamic changes in process behavior. As a result, they significantly enhance online monitoring capabilities, facilitating timely detection of assignable causes and improving overall process control. These advancements make machine learning-based control chart pattern recognition highly valuable in modern manufacturing and quality control, where automation and precision are essential.

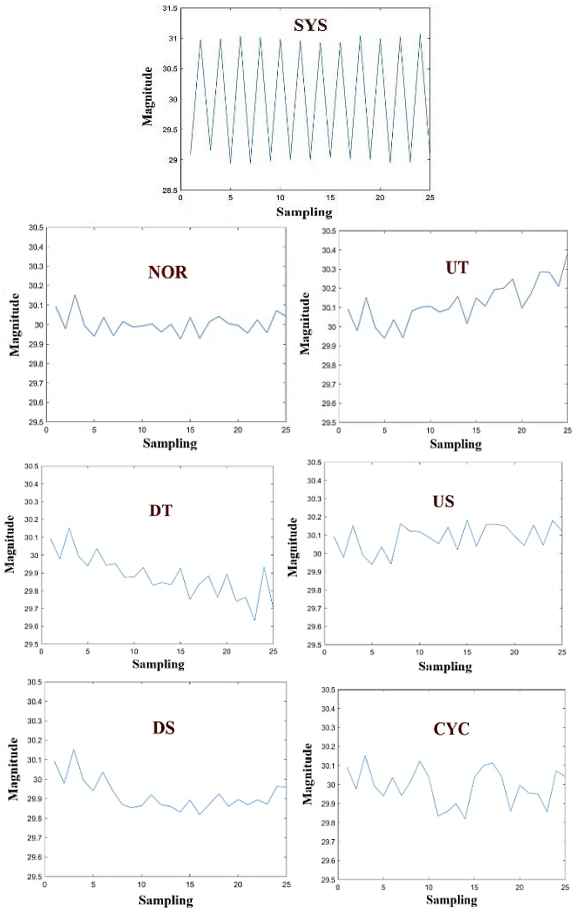


Figure 1. A sample for the well-known and typical CCPs

The remainder of the paper is structured as follows: Section 3 outlines the proposed methodology for the proposed intelligence classifier of control chart patterns in detail. Section 4 presents the comprehensive simulation results obtained. Section 5 applies the proposed intelligent CCP recognizer in a real-world data. The next section presents some managerial insights on this topic. Section 7 concludes the paper and offers suggestions for future research.

2. LITERATURE REVIEW

Due to the limitations of supplementary rules and the growing demands of intelligent manufacturing, there is an increasing interest in developing intelligent CCP recognition algorithms. As a result, the CCP recognition challenge is being addressed as a pattern recognition issue and addressed using machine learning methods (2). In this respect, a systematic literature review by Hachicha and Ghorbel (3) covered 120 research indexed papers until 2010. Moreover, Garcia et al. (4) presented a comprehensive taxonomic classification of studies

focused on identifying concurrent patterns, with most papers relying on ANNs until 2021. There are various machine learning-based approaches to detect different types of unnatural CCPs, considering the autoregressive nature of the process observations.

As mentioned, literature reviews shown that most studies have initially focused on ANNs. As a pioneer, Cheng (5) proposed a neural network-based CCP recognizer in SPC scope. After that, several works have conducted on this subjectin recent years too. For example, Kadakadiyavar et al. (6) presented an approach to classify CCPs by developing an adaptive radial basis function neural network that uses a stochastic gradient method to map pattern features to their respective categories. Recently, Zaidi et al. (7) presented a pattern recognition tool using multilayer feed-forward neural networks to accurately recognize Compositional data patterns.

Recent advancements in CCP recognition have involved the adoption of more complex frameworks to enhance classification accuracy. In this regard, SVMs have been widely used, achieving good results. However, SVM models can be complex due to the need to choose the most suitable kernel function and hyperparameter settings. In this regard, optimization algorithms can be practical, including cuckoo optimization (8), Genetic (9), and particle swarm optimization algorithms (10). In another viewpoint, to enhance the recognition performance, Lee et al. (11) combined SVM with adaptive neuro-fuzzy inference and spectral clustering techniques, respectively. Additionally, some researchers have employed various techniques, including expert systems (12), fuzzy- methods (13) and decision trees (14), to recognize CCPs.

Convolutional Neural Networks (CNNs) have gained prominence in various domains due to their ability to train directly from raw inputs to classifier outputs, reducing the need for manual feature engineering. This shift towards CNNs is particularly relevant in CCP recognition, where they can learn discriminative features directly from data, making them well-suited for these applications. For example, a CCP classifier is introduced by Zan et al. (15) leveraging a 1D-CNN. This approach eliminates the need for manual feature extraction, as the 1D-CNN learns to identify the most relevant features directly from the raw data. This automated feature learning process enables the method to efficiently identify CCP. Additionally, a new method that leverages a multi-label CNN was presented by Cheng et al. (16) to show concurrent CCPs. They extended a multi-channel deep CNN model to involve 1D and 2D representations of control chart information for pattern recognition in multivariate processes. However, a major challenge with CNNs is that they require significant expertise to choose the proper parameters, such as the number of kernels, kernel sizes, and learning rates. These parameters are interdependent, making it difficult to fine-tune them. To address this issue, Golilarz et al. (17) employed the Harris Hawks optimization algorithm to optimize CNN parameters. Also, CNN and long short-term memory were integrated based on Genetic algorithm as a classifier for CCPs by Yu and Zhang (18). Cheng et al. (19) proposed control chart mixture pattern classification using multi-label CNN. Wu et al. (20) have developed a CNN-based technique to show CCPs in autocorrelated processes, which is a recent contribution to the field. Recently, Xue et al. (21) designed a method of recognition based on multi-feature fusion by using CNN for the automatic recognition of imbalanced CCPs.

2. 1. Research Gap The literature demonstrates how well CNNs perform complicated classification tasks and can extract features from unprocessed data. Nevertheless, in real-world circumstances, it might not be the case that training and test data have the same properties and distribution, as most research assume. Because of this limitation, models must be rebuilt from the ground up. We used a modified version of VGG-16 to address this CCP recognition difficulty (22). It can be applied to the process of transfer learning to extract characteristics. We explored the value of transfer learning in cross-domain fault detection and diagnosis tasks, similar to the approaches used by Go et al. (23). The pretrained VGG-16 is a potent model whose pre-trained layers may be adjusted to obtain high accuracy for particular tasks. It can be used as a base for various image categorization applications like CCP recognition. Consequently, this work proposes a transfer learning-based control chart recognition approach. To enhance the network model's capacity for generalization throughout the feature extraction process, the convolutional layer is simultaneously fine-tuned using the fine-tune approach. Subsequently, the classifier is trained based on the number of control chart categorization classes. To classify the CCP, the classifier uses the output of the feature extraction layer as its input. We evaluated the model's performance using Monte Carlo simulations and real-world case, comparing it to traditional classifiers. This approach has the potential to contribute to developing smart SPC systems.

3. METHODOLOGY

3. 1. Data Generation Despite the significant advancements in automatic production capacity, leading to a substantial accumulation of data. Garcia et al. (4) found that most studies (41 out of 44) still rely on simulated data to assess the performance of CCP recognition methods due to the limited availability of publicly accessible, thoroughly documented databases. Only three studies used real data. In this research, a

significant number of CCP samples are generated using the Monte Carlo method, which is commonly employed for training and testing CCPs. The literature frequently referred to seven fundamental CCPs: NOR, UT, DT, US, DS, CYC and SYS. Based on Zan et al. (15), the process means and two noise components are utilized for the different patterns as follows:

$$y(t) = \mu + x(t) + d(t), \quad (1)$$

in which $y(t)$ shows the value of a sample gathered during sampling *t*. The *t* is the discrete-time in which the monitored process variable is gathered. $x(t)$ shows random noise at t^{th} time, and it follows a normal distribution with a mean of zero and a variance of σ^2. Note that μ and σ^2denote the mean and variance when process is in-control state, respectively. In addition, $d(t)$ demonstrates particular pattern due to assignable causes manufacturing process at t^{th} time.

Based on Equation 1, the simulation approach for CCPs is performed. For NOR pattern, we consider $d(t) = 0$, and for $d(t) = \pm v \times s$ in which v is responsible for the shift position such that it is equal to zero before the change and to one after the change. Moreover, *s* represents magnitude of shift, considering its direction. The UT and DT patterns are shown as Equation 2.

$$d(t) = \pm v \times d \times t, \quad (2)$$

in which v is responsible for the shift position similar to the concept of upward-and downward-shift patterns. Moreover, d shows the slope of a trend in which signs "+" and "-" are applied for the UT and DT patterns, respectively, and CYC pattern is shown by Equation 3.

$$d(t) = v \times a \times sin\left(\frac{2\pi t}{\omega}\right), \quad (3)$$

in which ω and a denote the period and amplitude of a cycle, respectively. On the other hand, when a systematic pattern has been observed in a control chart based on $d(t) = b \times (-1)^t$, it is commonly referred to as a SYS pattern. This type of pattern indicates that there is an underlying, consistent cause influencing the process, which requires further investigation and corrective action to maintain the desired level of quality and consistency.

3. 2. The Deep Neural Network for CCP Classification CNNs can simplify the difficult feature extraction in the conventional recognition because they are feed-forward, weight-sharing neural networks. A common CNN consists of the following layers: input, convolutional, downsampling, fully connected, and output. The computation depending on input can be advantageous for the feature map. Equation 4 represents the output of channel *j* of the convolutional layer *l* as x_j^l, where $f(.)$ is the activation function.

When the output characteristic x_j^{l-1} of the previous layer is weighted and biased, it produces the net activation of channel j in the convolutional layer l, represented as u_j^l in Equation 5. The feature map input set denoted by M_j is utilized to obtain the convolution kernel matrix k_{ij}^l, which in turn yields u_j^l. The bias of the feature map is b_j^l.

$$x_j^l = f(u_j^l). \quad (4)$$

$$u_j^l = \sum_{i \in M_j} x_j^{l-1} * k_{ij}^l + b_j^l. \quad (5)$$

Computing the feature map's input is one method of figuring out the downsampling layer's output. In the down-sampling layer *l*, the net activation of channel *j* is obtained by biasing and weighting the output characteristic x_j^{l-1} of the previous layer represented by u_j^l in Equation 6. β represents the bias weight coefficient. The function for subsampling is down (.).

$$u_j^l = \beta_j^l down\left(x_j^{l-1}\right) + b_j^l. \quad (6)$$

The fully connected layer's output is calculated using the activation function and the weighted sum of the input. The completely linked *l* layer's net activation in Equation 7 obtained by biasing and weighting the output characteristic x^{l-1} of the layer that came before it.

$$u^l = w^l x^{l-1} + b^l. \quad (7)$$

The back propagation method computes the error based on the actual network output and expectation. Equation 8 shows the total error of training, where y_n is the sample training prediction value for the n^{th} sample and t_n is the true value of the label for the nth sample.

$$E(w, \beta, k, b) = \frac{1}{2}\sum_{n=1}^{N} \|t_n - y_n\|^2. \quad (8)$$

The primary applications of CNNs include semantic segmentation of scenes, object detection, face recognition and verification, and image categorization. Within this field, one of the main areas of research is image categorization. Picture classification can be significantly aided by using the ImageNet (24) dataset, which comprises 3200000 images, 5246 synonym sets, and 12 subtrees. Using a CNN, Krizhevsky et al. (25) developed the AlexNet model, and he took first place in the image categorization competition. The VGG16 model from Oxford University (22), GoogleNet from Google (26), ResNet from Microsoft (27), and so on are examples of common CNN models. The Visual Geometry group's VGG16 network model is currently regarded as a top-notch image categorization model. Thirteen convolutional layers, five downsampling layers, and three fully connected layers comprise the VGG16 network structure. The parameters in the 16-level network total 1380000000.

3. 2. 1. The Modified VGG 16 Architecture The suggested model (Figure 2) is a modified version of the VGG net. Because of the VGG's straightforward architecture, which makes result analysis easier, and the 32x32 input would not disappear via it, we decided to utilize it as the base net. Moreover, one of the strongest deep CNNs is the VGG net. The results in Section 4 will demonstrate how much more powerful our VGG-based deep model is in classifying the CCP dataset when compared to other types of convolutional networks. The modified VGG16 contains 13 convolution layers and two fully connected layers. The proposed model has five groups of convolution layers and 1 group of fully connected layers. Each convolution filter has a pooling zone of 2 by 2 without overlap and a kernel size of 3 by 3 with a stride of 1. It should be noted that we minimize the number of parameters greatly by shrinking the two 4096-dimension fully connected layers to one 1000-dimension fully connected layer.

3. 2. 1. 1. Batch Normalization Small modifications in earlier levels of very deep networks will amplify layer by layer until they become an issue. The term "internal covariate shift" refers to the problem that arises when the input distribution of a layer changes since the parameter must iteratively adjust to the new distribution (28). To address this issue, Ioffe et al. (28), a researcher at Google, presented Batch Normalization. This method normalizes the inputs of each layer, which accelerates network convergence, lowers error rates, and somewhat reduces overfitting. With adding Batch Normalization layers to GoogleNet, Ioffe et al. (28) beat human performance records in ImageNet classification tasks. Batch normalization, in our opinion, can also optimize extremely complex models like VGG. Before the VGG16 architecture and dense layers, we add the Batch Normalization layer. Our implementation differs slightly from the conventional one, though. We apply the same z-score-like function to the Batch Normalization layer's inputs.

$$\hat{x}^{(k)} = \frac{x^{(k)} - E(x^{(k)})}{\sqrt{var(x^{(k)})}}. \quad (9)$$

After that, perform the linear operation on $\hat{x}^{(k)}$:

$$y^{(k)} = \gamma^{(k)}\hat{x}^{(k)} + \beta^{(k)}. \quad (10)$$

All operations are channelwise since k indicates the channel. Instead of using the fixed mean and variance, we utilize exponentially-weighted moving averages, where the subscript n denotes the number of iterations.

$$Mean_n = \begin{cases} BatchMean_1 & n = 1 \\ \alpha BatchMean_{(n)} + (1-\alpha)Mean_{n-1} & n > 1 \end{cases} \quad (11)$$

$$Mean_n = \begin{cases} BatchVar_1 & n = 1 \\ \alpha BatchVar_{(n)} + (1-\alpha)Var_{n-1} & n > 1 \end{cases} \quad (12)$$

The outcomes demonstrate that batch normalization greatly raises the accuracy of recognition. With the aid of dropout, which lowers the error rate, it can awaken the hidden power of extremely deep models.

3. 2. 1. 2. Dropout Setting The first thing we need to consider when applying a deep network to a limited dataset is how to lessen overfitting. Dropout is the most widely used technique to lessen overfitting. Dropout is typically used by setting it at the fully connected layers,

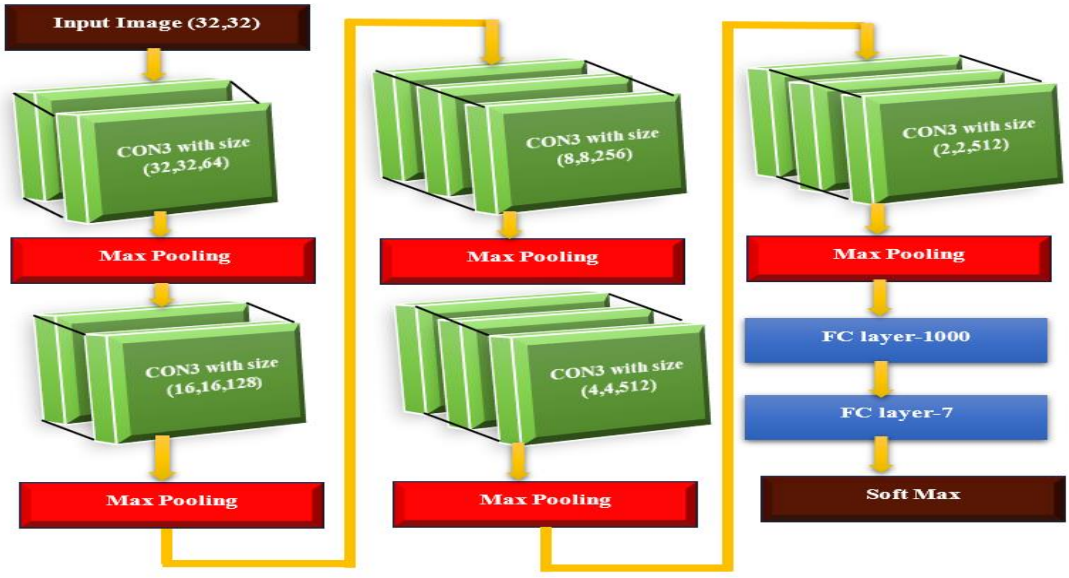


Figure 2. The VGG 16 network's architecture after modification

which contain most of the parameters. In this case, the fully connected layers' dropout rate was adjusted to 0.5 because the model is quite deep and the dataset is small. The overfitting of the deep model is considerably controlled by the dropout value, according to the results. The deep model's hidden power was unleashed in conjunction with the application of batch normalization, lowering the error rate.

3. 2. 1. 3. Training Acceleration While deep models can yield powerful representations, training a deep model is computationally demanding; that is, both the forward and backward passes are slower than with shallow models. For classification learning to be reliable, there must be an adequate number of training examples to guarantee the classification model's correctness, and the samples must be independent and uniformly distributed. Nonetheless, there exist specific issues such as a restricted quantity of training samples and unsatisfactory independence and distribution of samples. The researchers proposed a new machine learning technique called transfer learning (29). Transfer learning synchronous satisfies two fundamental difficulties of traditional machine learning, allowing for the application of current knowledge to a limited amount of labelled sample data in the target area without learning issues. At the moment, transfer learning is divided into three components: instance-based transfer learning in isomorphic space, transfer learning using features in isomorphic space, and transfer learning in heterogeneous space. This work uses the isomorphic space characteristic for transfer learning. There are four steps in the generic CNN method that are based on transfer learning.

The random starting parameters of the model are saved and trained using large data sets from related disciplines, which improves the network's generalization capacity. Second, features are extracted from the trained network model's convolutional and downsampling layers. Thirdly, samples from the target domain are used to train the pattern classifier. The target application of pattern classification is finished when the feature extraction layer's output is fed into a pattern classifier.

The processes of extracting VGG16 features, training classification models, and fine-tuning them are all part of the transfer learning-based control chart recognition principle. As illustrated in Figure 3, the whole network structure for control chart identification consists of the feature extraction part, fine-tuning section, and classification components of VGG16. Convolutional and downsampling layers are used in the VGG16 feature extraction step of the model, and the fully connected layer is removed from the modified VGG16 model. The feature extraction portion's weight parameters are equal to the VGG 16 weights when trained in ImageNet, which enhances the network's capacity to generalize and yield the best classification results, particularly in the absence of a large amount of training data. In the last stage of feature extraction, the fine-tuning portion includes three convolutional layers with size (2,2,512) and one downsampling layer. We selectively freeze the layers in the fine-tuning portion in addition to the feature extraction part's layers.

Freezing layers serves to prevent changing parameter values and maintain the network's capacity for generalization. None-freezing layers are used to increase the network's capacity for adaptive learning to the target samples. There are six different types of anomalous patterns that are the defined output results of the two completely connected layers that make up the categorization section.

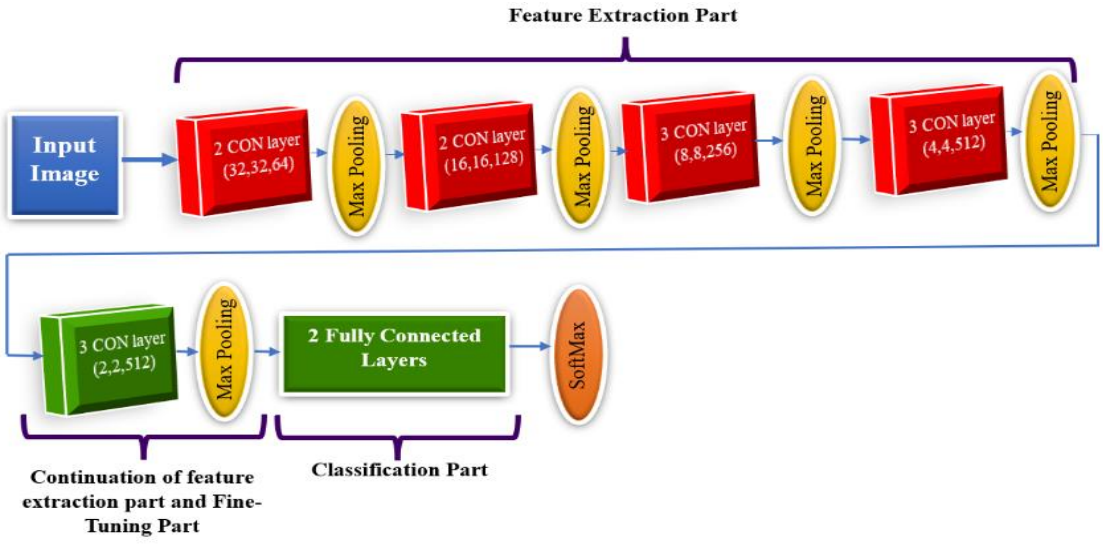


Figure 3. The principle of transfer learning in network architecture

4. SIMULATION AND VERIFICATION

4. 1. CCP Parameter Settings In this subsection, seven common CCPs were analyzed using the Monte-Carlo simulation method, which is described in Section 3, to generate datasets for the modified pretrained VGG16. In this respect, Table 1 presents the equations and parameter settings used in the simulation process.

In the real applications, the changes in quality characteristic are complex (30). To accurately simulate this, CCP parameters were generated from a uniform distribution. The μ and σ values are set to 30 and 0.05, respectively. Next, the slope d is randomly selected within the range of 0.1σ to0.3σ. The s value follows a uniform distribution between 1.5σ and 3σ. Furthermore, a uniformly distributed between $1.5\sigma$and 4σ. The ω value is respectively set to 4, 5, 6, 7, and 8. The responsible position of CCPs is in range of 4 to 9. Also, the data length for each pattern is 25. In contrast to VGG, comparing recognition rates across various sample sizes revealed that beyond 2000 samples, the recognition rate improvement becomes negligible for other models. Conversely, the recognition accuracy of VGG network is enough high with 750 samples per CCP, resulting in a total of 5250 unique samples. The dataset was categorized into two separate parts in a random manner. 4200 samples were used to train the VGG, while the remaining samples were reserved for testing.

4. 2. Expremental Setup To finish the experiment, Keras based on TensorFlow, serves as the experimental setting in this research. The input consists of 32*32 jpeg files produced by Keras' matplotlib module based on sample numerical data. The output vector has seven lengths and an element value of either 0 or 1. The output vector is represented by the expression $y = [y_1, y_2, y_3, y_4, y_5, y_6, y_7]$; $y_i = 1$ denotes if the output is the input pattern, while $y_i = 0$ denotes that it is not. As an illustration, [1, 0, 0, 0, 0, 0, 0] denotes a normal pattern. The data set is randomly divided into two subsections of training and test data, which are completely separate from each other. Typically, 80% samples are used for training and 20% samples are used for testing. In this work, there are 50 training iterations and a 4:1 ratio between training and test data.

TABLE 1. Parameters for generating the six CCPs

CCP	Formula	Parameters values
NOR	$y(t) = \mu + x(t)$	$\mu = 30, \sigma = 0.05$
UT	$y(t) = \mu + x(t) + v \times d \times t$	$d \sim U(0.1\sigma, 0.3\sigma)$
DT	$y(t) = \mu + x(t) - v \times d \times t$	$d \sim U(0.1\sigma, 0.3\sigma)$
US	$y(t) = \mu + x(t) + v \times s$	$d \sim U(1.5\sigma, 3\sigma)$
DS	$y(t) = \mu + x(t) - v \times s$	$d \sim U(1.5\sigma, 3\sigma)$
CYC	$y(t) = \mu + x(t) + v \times a \times sin\left(\frac{2\pi t}{\omega}\right)$	$d \sim U(1.5\sigma, 4\sigma)$, $\omega \in \{4,5,6,7,8\}$
SYS	$y(t) = \mu + x(t) + b \times (-1)^t$	$b = 1$

4. 3. Analysis of Classification Results

4. 3. 1. Optimizing the Feature Extraction Layers The sample size for each pattern in this paper is 25 points, and the total number of dataset samples is 5250, of which 600*7 is the training sample and 150*7 is the test sample. The experiment's goal is to optimize the pretrained, modified VGG16 network by freezing various layers and identifying the optimal layer to freeze based on the control chart's recognition accuracy. Table 2 displays the findings of the experiment. The modified pretrained VGG16 model's last three convolutional layers and one downsampling layer have been fine-tuned, according to the experimental data, and the control chart's overall classification accuracy is higher than 99%. When the modified pretrained VGG16 model has 13 frozen layers instead of 12, the overall recognition rate increases. On the control chart, the modified pretrained VGG16 network performs best for recognition with 14 frozen layers. Simultaneously, 99.3% of normal and aberrant patterns are recognized. To achieve the best recognition result, we therefore set the modified pretrained VGG16 model's number of frozen layers to 14 in the subsequent trials.

4. 3. 2. Improve the Performance of Training Batch normalization-also referred to as batch norm-using re-centering and rescaling the layer inputs allows artificial neural networks to be trained more quickly and steadily. Additionally, "dropout" in machine learning describes the technique of randomly ignoring some nodes in a layer when training. A regularization strategy called a dropout makes sure that no units are codependent on

TABLE 2. The network model's recognition outcomes for the various frozen layers

Frozen layer of the modified pretrained VGG16	Test Accuracy (%)
12	97.1
13	99.3
14	99.3
15	98.9
16	98.6
17	98.4
18	98.4

another, which inhibits overfitting. As a result, we create four models for the CCP dataset's classification to conduct comparison studies.

1. Baseline: The modified pretrained VGG16 network was used as a baseline. The training process was performed without modification.
2. The modified pretrained VGG16 for CCP classification based on BN: This model consists of Batch Normalization layer before the VGG16 architecture and dense layers in the structure is shown in Section 3.2.1.3.
3. The modified pretrained VGG16 for CCP classification based on DROPOUT: This model includes dropout on the structure is shown in Section 3.2.1.3.
4. The modified pretrained VGG16 for CCP classification based on DROPOUT and BN: This model use Batch Normalization + dropout on the structure is shown in Section 3.2.1.3.

Our models were trained using the CCP dataset. Our experiment's results are displayed in Table 3. We discovered that the accuracy gained by the very deep model was not particularly high when we used only Batch Normalization or the dropout option. However, compared to the other models, the modified pretrained VGG16 model with dropout and batch normalization gains greater accuracy. This table demonstrates that the extremely deep model achieves high accuracy when both the BN and dropout are used. We believe that depth provides a high accuracy opportunity, BN keeps the training process from being stuck in subpar local optima, and dropout greatly reduce overfitting. This demonstrates how the power of depth can even be used with tiny datasets.

4. 3. 3. Different Nonlinearity To determine the best activation function, a number of simulations was done, the results of which are displayed in Table 4. Scheme 1 used the Sigmoid function to activate the unfrozen layers of the modified pretrained VGG16 architecture; Scheme 2 used the Softmax function to activate the output layer and the Leaky ReLU function to activate the unfrozen layers. The Softmax function for the output layer and the ReLU functions for the unfrozen layers were Scheme 3's activation functions.

TABLE 3. The comparative studies comparing for effects of batch normalization and dropout on CCP categorization

Network architecture	Accuracy
Baseline	88.2
The modified pretrained VGG16 with batch normalization	90.1
The modified pretrained VGG16 with Dropout	86.7
The modified pretrained VGG16 + Dropout + Batch normalization	99.3

The network using the ReLU and Softmax activation functions outperformed the other activation function, as shown in Table 4. Results employing Leaky Relu and Relu nonlinearity indicate no discernible difference. It's possible that we didn't apply the ideal negative slope setting for Leaky ReLU. However, given they are all non-saturate nonlinearities, we think that there isn't much of a distinction between these nonlinearities.

4. 4. Comparison of CCP Recognition A comparative study of 1D-CNN, CNN, MLP, and a modified pretrained VGG16 network model was carried out to assess the efficiency of the suggested approach in more detail. We utilized the MLP in this work's comparison since it is a machine learning technique with exceptional performance that has been applied extensively in the CCP recognition and has shown positive outcomes (30). An ordinary three-layer neural network, or MLP, was employed in the experiment. It had 25 neurons that received the raw CCP data and nine neurons that received the feature set. Six neurons in the output layer and fifteen in the hidden layer represented the six common CCPs. The Sigmoid function served as each layer's activation function. Two convolution and pooling layers with a fully connected layer make up the 1D-CNN and CNN structure. The fully connected layer performs classification, while the convolution and pooling layers handle feature extraction. The input was split into four categories: raw data for the 1D-CNN, control chart image data (60 × 60 pixels), feature set or raw data for the MLP, and 32*32 pixels for the modified pre-trained VGG-16. The MLP feature set included the statistical and shape characteristics that have been demonstrated to be very successful in the CCP recognition introduced by Garcia et al. (4). You may get a thorough explanation of the characteristics defined by Zan et al. (15).

In this experiment, the modified pre-trained VGG16 network (MPVGG16) and other network models (MLP, 1DCNN, CNN) are compared for their ability to recognize a control chart using the same samples. The experiment has a sample size of 25 for each pattern, the training and test data are randomly selected from the dataset and their ratio is 4:1, and 14 frozen layer was considered in the modified pre-trained VGG16 network. Table 5 displays the findings of the experiment. The experimental results show that, firstly, both the

TABLE 4. Modified pretrained VGG16 accuracy comparison using various activation functions

Accuracy	Activation Function of the unfrozen layers
99.3	Relu + Soft max
99.1	Leaky Relu + Soft max
98.1	Sigmoid

MPVGG16 network model and other network models constantly improve their control chart identification rate with an increase in training data volume. Second, with low training data, the MPVGG16 network model achieves above 98% recognition accuracy. Thirdly, under varying sample data quantities, the MPVGG16 network model performs better overall than other network models in the classification accuracy of the control charts. This shows that, two-dimensional training data in the form of control chart image is a good input candidate for the modified pre-trained VGG16 model. Additionally, it shows that a pre-trained VGG16 network with low design and tuning complexity can classify the control chart more accurately, which has significance for research as well as practical applications.

Table 5 presents a comparison of the five deep network examples' performance based on varying numbers of samples per pattern. The number of training samples for each CCP was considered 200, 600, or 2,000, as Table V illustrates. It can be concluded from Table V that the suggested approach outperformed the conventional MLP network and the CNN method in terms of recognition accuracy.

The confusion matrix was considered to assess the classifier's performance in a more intuitive manner. The percentage of correctly identified patterns was indicated by the values on the diagonal of the confusion matrix. The percentage of misclassifications is represented by the other values in the matrix. The average value of every member on the matrix diagonal equaled the control chart recognition rate. Tables 6-10 display the confusion matrix, and it is evident that the modified pretrained VGG16 performed better than the other approaches. In particular, the modified pretrained VGG16 network model produced a lower mistake rate and improved recognition accuracy.

The suggested approach was contrasted with the techniques documented in the literature review to assess the efficacy of the modified pretrained VGG16 model even more. Several classifiers were utilized for the CCP

TABLE 5. Comparison findings for various sample of training and test patterns

Train sample	Test sample	MLP/raw data (15)	MLP/ feature (15)	1DCNN (15)	CNN (15)	This work
200 sample/pattern	50 sample/pattern	82.61	92	92.4	93.2	98
600 sample/pattern	150 sample/pattern	-	-	-	-	99.3
2000 sample/pattern	500 sample/pattern	85.93	93.27	98.33	96.33	99.5

TABLE 6. The CCP recognition confusion matrix for the modified pretrained VGG16, total accuracy with VGG=99.3%

	NOR	UT	DT	USS	DS	CYC	SYS
NOR	99.9	0	0	0	0	0.1	0
UT	0.1	99.2	0	0.7	0	0	0
DT	0	0	99.1	0	0.9	0	0
US	0	0.9	0	99	0	0.1	0
DS	0	0	1.1	0	98.9	0	0
CYC	1	0	0	0	0	99	0
SYS	0	0	0	0	0	0	100

TABLE 7. The CCP recognition confusion matrix the MLP and raw data reported by Zan et al. (15), total accuracy with VGG=85.93%

	NOR	UT	DT	US	DS	CYC
NOR	73.6	2.4	1.6	5.6	6	10.8
UT	4.8	92	0	2.8	0	0.4
DT	2.4	0	91.6	0.4	5.2	0.4
US	6.4	4.4	0	86..4	0	2.8
DS	3.2	0	5.2	0	91.6	0
CYC	18.8	0	0	0.4	0.4	80.4

TABLE 8. The CCP recognition confusion matrix for the MLP and feature set of data reported by Zan et al. (15), Total accuracy=93.27%

	NOR	UT	DT	USS	DS	CYC
NOR	97.6	0	0.4	0	0.4	1.6
UT	1.2	89.6	0	9.2	0	0
DT	0	0	96	0	4	0
US	0	8.8	0	91.2	0	0
DS	0.4	0	10.8	0	88.8	0
CYC	3.6	0	0	0	0	96.4

TABLE 9. The CCP recognition confusion matrix the 1D-CNN and raw data reported by Zan et al. (15), total accuracy=98.33%

	NOR	UT	DT	USS	DS	CYC
NOR	100	0	0	0	0	0
UT	0	97.6	0	2.4	0	0
DT	0	0	98	0	2	0
US	0	1.2	0	98.4	0	0.4
DS	0	0	3.2	0	96.8	0
CYC	0	0	0.4	0	0.4	99.2

TABLE 10. The CCP recognition confusion matrix for the CNN and image data reported by Zan et al. (15), total accuracy=96.33%

	NOR	UT	DT	USS	DS	CYC
NOR	99.2	0	0	0	0.4	0.4
UT	0.8	92.4	0	6.8	0	0
DT	0	0	97.2	0	2.8	0
US	0	2.8	0	97.2	0	0
DS	0	0	6.4	0	93.6	0
CYC	1.6	0	0	0	0	98.4

recognition, as indicated in Table VI, including the Multi-Layer perceptron (MLP), PNN, Fuzzy ARTMAP, SVM, and RBF. Either a feature set or raw data was used as the input. According to the results of Table VI, in neural networks other than convolution networks, the use of pre-extracted feature vectors in the input has a better performance than applying raw data as input. The performance was especially noticeable when statistical and shape features were combined. However, in convolution networks, due to the presence of convolution and pooling layers as feature extractors, raw data input is a suitable option. However, new learning transfer techniques can address the issue of convolution networks having too many parameters and requiring a large amount of training data for the training phase. Table 11 demonstrates the high accuracy of the suggested pre-trained network (99.3% classification accuracy) when compared to alternative approaches, despite using less training data.

Most of the literature used control chart samples with 60 or more sampling points each, which may guarantee a more significant difference between the various CCP kinds. Moreover, over time, the trend patterns might become more pronounced, aiding in achieving high classification accuracy. To identify irregularities in the manufacturing process more quickly and lower the loss of manufacturing companies, there were 25 sampling points per control chart sample in this work. But the patterns might grow remarkably similar, particularly the shift and trend patterns, making identification even more challenging.

5. A REAL APPLICATION

A control chart diagnostic system has been created and used to monitor an actual dataset from the production environment to better demonstrate the suggested technique. An instrument that measures three coordinates can be used to measure a part's diameter, which is thought to be the primary quality attribute. There were sixty sample points in all. As demonstrated in Figure 4 and Table 12, the modified pretrained VGG model can identify all five unnatural CCPs.

The Nelson rules are applied to control charts to identify patterns that deviate from a random, stable process, allowing manufacturers to investigate and address the underlying causes. In this regard, Figure 5 demonstrates Nelson rules for the first six sample in the real case. The numbers displayed within the circular markers in individual charts indicate the specific rules, based on Nelson's rules, that were violated and consequently triggered a warning signal.

Figure 5 shows that specific patterns that may indicate process instability or other issues for these samples based on Nelson rules method. However, it's essential to consider the rules in combination and the context of the specific process to draw accurate conclusions about potential causes of unnatural patterns. While, the proposed diagnostic method, which can

TABLE 11. Comparing the effectiveness of various techniques

Model	Number of patterns	Input representation	Approach	Test accuracy
Guh and Tannock (31)	4	Raw data	MLP	94.38
Hassan et al. (32)	6	Statistical features	MLP	96.80
Assaleh and Al-Assaf (33)	4	Frequency features	MLP	97.22
Cheng and Ma (34)	6	Raw data	PNN	95.58
Zan et al. (35)	6	Autoregressive (AR) spectrum	Fuzzy ARTMAP	95.00
Ranaee and Ebrahimzadeh (36)	6	Shape and statistical features	MLP	99.15
Kao et al. (37)	7	Independent component analysis (ICA)	SVM	98.94
Zhou et al. (9)	6	Shape and statistical features	FSVM	99.28
Addeh et al. (38)	8	Shape and statistical features	Bees-RBF	99.63
Zan et al. (15)	6	Raw data	1D-CNN	98.33
This work	7	Raw data	The modified pretrained VGG16	99.3

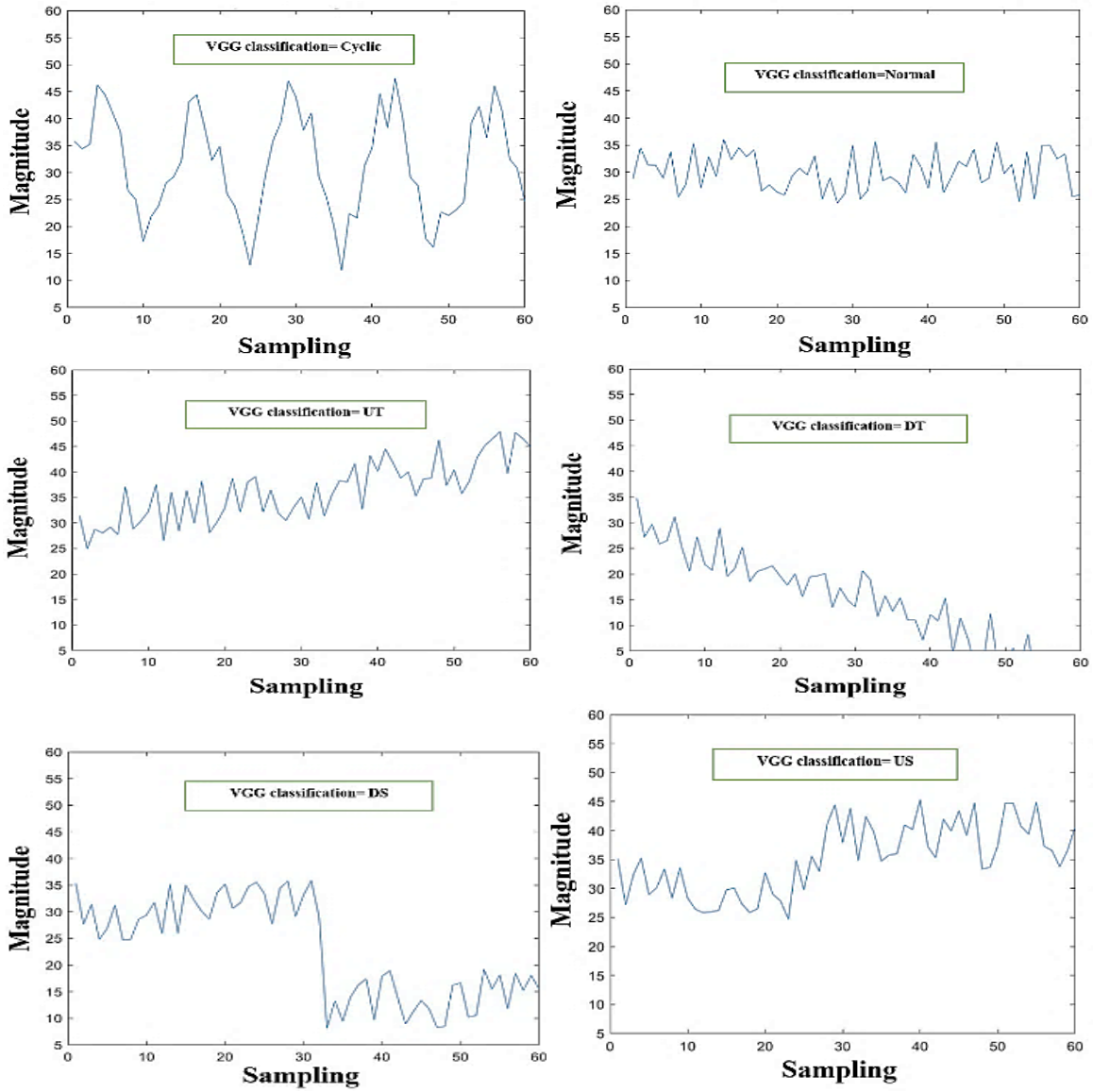


Figure 4. The suggested model's detection of normal and abnormal CCPs in the real data

TABLE 12. Accuracy of real data CCP recognition by the modified pretrained VGG16 (test sample: 100 samples/pattern)

	Real Data						Overall accuracy
Class	**NOR**	**UT**	**DT**	**US**	**DS**	**CYC**	
Accuracy	100	99.4	99.2	97	97.1	918	98.45

accurately predict the specific pattern, has the potential to significantly improve manufacturers' capacity to identify the root causes of process variations. By providing a more precise pattern classification, this approach enables manufacturers to more effectively pinpoint the reasons for unnatural patterns and take targeted corrective actions. As a result, the implementation of the proposed pattern recognition technique can lead to substantial savings in both time and costs associated with quality control and process improvement efforts. Hence, it was discovered via correspondence with enterprise engineers that the exchange of equipment and tool status could cause these abnormal CCPs to appear during processing. This indicates the model trained using a simulated dataset was still able to identify real datasets from the real-world

setting with good accuracy. Furthermore, the recognition results of the suggested model and Nelson rules complement each other, and utilizing both can improve the recognition system.

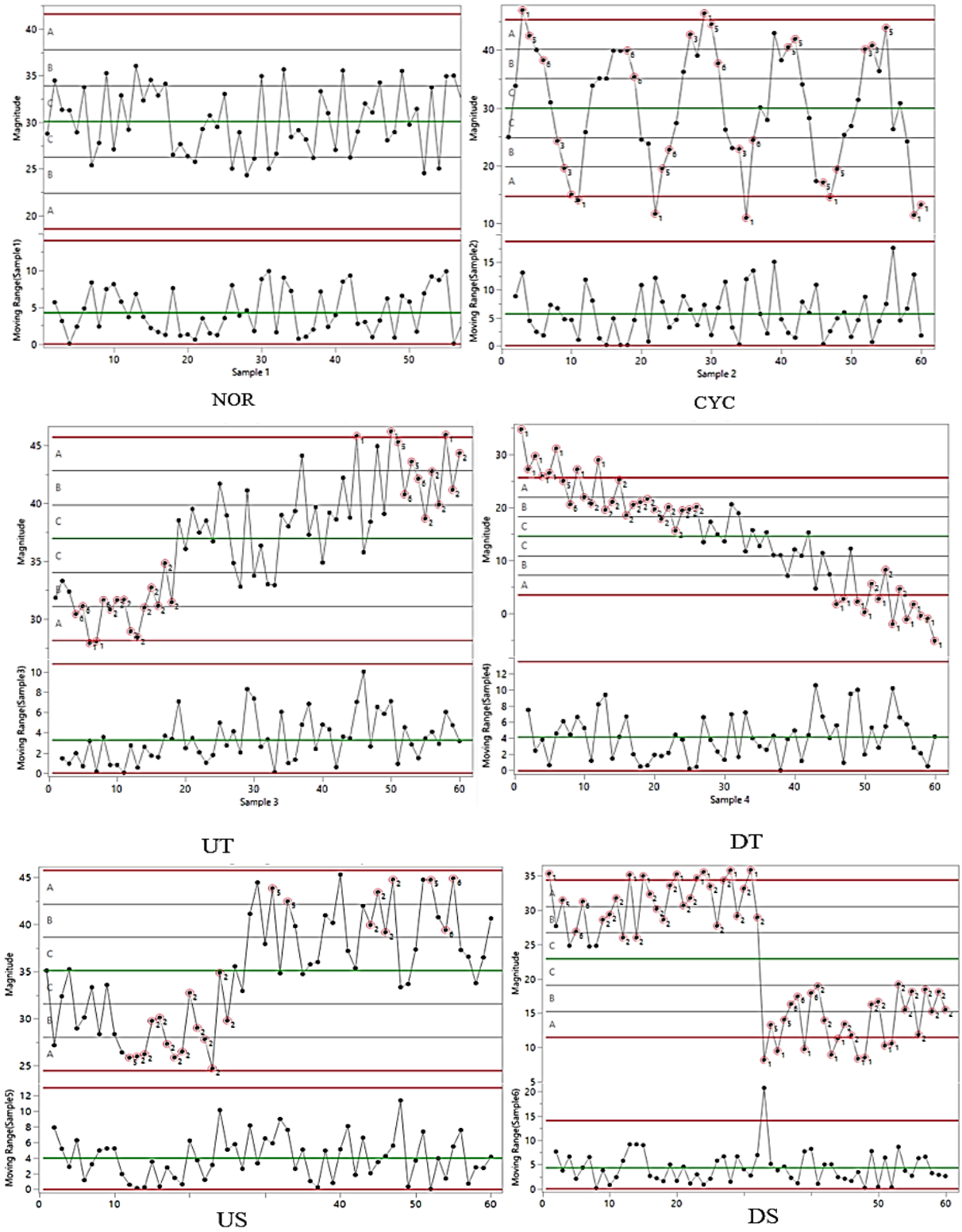


Figure 5. Nelson rules in Individual & Moving Range (IMR) chart in the real data

6. MANAGERIAL INSIGHTS

The findings of this study provide several important managerial implications for manufacturing and quality control professionals:

- Enhanced Quality Control through Advanced AI Techniques: By adopting transfer learning with pre-trained deep networks such as VGG-16, managers can significantly improve the accuracy and reliability of control chart pattern recognition. This leads to earlier and more precise detection of process abnormalities, minimizing defects and reducing waste (39).
- Cost and Time Efficiency in Model Development: The use of transfer learning reduces the need for extensive retraining and large volumes of labeled data. Managers can thus deploy intelligent quality control systems more rapidly and cost-effectively, even when historical data are limited or when processes change frequently.
- Improved Decision-Making with Robust Recognition: The integration of techniques like ReLU activation, batch normalization, and dropout enhances model robustness and consistency. This reliability supports confident, data-driven decision-making in production environments, helping managers maintain process stability and product quality.
- Scalability and Adaptability: The proposed approach's ability to generalize well across different control chart patterns and conditions means that manufacturing firms can scale their quality monitoring systems across multiple production lines or facilities without the need for rebuilding models from scratch.
- Strategic Investment in AI-Driven Quality Systems: Managers should consider investing in AI-powered quality control solutions as a strategic asset that not only improves operational efficiency but also strengthens competitive advantage by ensuring superior product quality and compliance with industry standards.

Overall, the study highlights how leveraging state-of-the-art machine learning techniques can transform traditional quality control practices into intelligent, adaptive systems that support proactive management and continuous process improvement.

7. CONCLUSION AND FUTURE RESEARCH

The accurate and automated classification of CCPs is crucial for manufacturing processes to maintain high-quality production processes. Recently, there has been extensive research and exploration into integrating CCP recognition with machine learning methods. This study proposes a modified pretrained VGG16 model for end-to-end CCP recognition realization and feature learning. The modified pre-trained VGG16 model was trained in this paper's unfrozen layers in the isomorphic space using the image of the CCP. The training set of control chart images is really used for fine-tuning which can enhance and satisfy the need for control chart recognition in complicated circumstances.

The results collected allow for the drawing of the following conclusions. 1) Despite the large number of network structure parameters in VGG16, this problem was resolved and the maximum recognition accuracy compared earlier deep networks was attained by applying the transfer learning technique. 2) The ReLU activation function significantly increased the recognition accuracy and was better suited for deep learning. 3) The accuracy of control chart recognition can be increased by using the batch normalization and dropout technique. 4) Using a pre-trained deep network improves the accuracy of CCPs' correct detection even with a low number of training sample. 5) The feature set recovered manually was of lower quality than the one that was acquired by the suggested modified pretrained VGG16 network. The suggested network identified the patterns properly with a 99.3% recognition accuracy when the image of control chart patterns was utilized as an input. This significantly improved identification accuracy over the conventional MLP model, 1DCNN, and CNN. In conclusion, the suggested model helps to increase the intelligence level of quality control in businesses by avoiding the challenge of extracting various complicated features. This makes it more suited to actual quality control.

Limitations: While the proposed model demonstrates excellent recognition accuracy under the tested conditions, it has certain limitations. First, the study primarily focused on individual control chart patterns under the assumption of normality and independence, potentially limiting its applicability to more complex, non-normal, or correlated processes. Second, the generalizability of the model to different types of control charts and industrial settings requires further validation. Third, the computational resources required for training and fine-tuning deep learning models like VGG16 may be a barrier to adoption for some smaller enterprises.

Recommendations for Future Research: Based on these limitations, future research can proceed in several directions. First, extending CCP recognition systems for non-normal data distributions, such as those belonging to the exponential family, is essential. Second, expanding the proposed method to recognize concurrent patterns would address more complex real-world scenarios. Third, developing this intelligence system for multivariate processes and exploring its applicability to different control chart types (e.g., EWMA, CUSUM) could broaden its practical utility. Finally, investigating techniques to reduce the computational demands of the model, such as model compression or efficient fine-tuning strategies, would make it more accessible to a wider range of users.

8. REFERENCES

1. Sogandi F, Vakilian F. Isotonic change point estimation in the AR (1) autocorrelated simple linear profiles. International Journal of Engineering Transactions A:Basics. 2015;28(7):1059-67. 10.5829/idosi.ije.2015.28.07a.12.
2. Owlia M, Fallahnezhad M. A new statistical approach for recognizing and classifying patterns of control charts (research note). International Journal of Engineering-Transactions A: Basics. 2015;28(7):1040. 10.5829/idosi.ije.2015.28.07a.10.
3. Hachicha W, Ghorbel A. A survey of control-chart pattern-recognition literature (1991–2010) based on a new conceptual classification scheme. Computers & Industrial Engineering. 2012;63(1):204-22. 10.1016/j.cie.2012.03.002.
4. Garcia E, Penabaena-Niebles R, Jubiz-Diaz M, Perez-Tafur A. Concurrent control chart pattern recognition: A systematic review. Mathematics. 2022;10(6):934. https://doi.org/10.3390/math10060934.
5. Cheng C-S. A neural network approach for the analysis of control chart patterns. International Journal of Production Research. 1997;35(3):667-97. https://doi.org/10.1080/002075497195650.
6. Kadakadiyavar S, Ramrao N, Singh MK. Efficient mixture control chart pattern recognition using adaptive RBF neural network. International Journal of Information Technology. 2020;12(4):1271-80. https://doi.org/10.1007/s41870-019-00381-z.
7. Zaidi FS, Dai H-L, Imran M, Tran KP. Analyzing abnormal pattern of hotelling T2 control chart for compositional data using artificial neural networks. Computers & Industrial Engineering. 2023;180:109254. https://doi.org/10.1016/j.cie.2023.109254.
8. Khormali A, Addeh J. A novel approach for recognition of control chart patterns: Type-2 fuzzy clustering optimized support vector machine. ISA transactions. 2016;63:256-64. https://doi.org/10.1016/j.isatra.2016.03.004.
9. Zhou X, Jiang P, Wang X. Recognition of control chart patterns using fuzzy SVM with a hybrid kernel function. Journal of Intelligent Manufacturing. 2018;29:51-67. https://doi.org/10.1007/s10845-015-1089-6.
10. Ranaee V, Ebrahimzadeh A, Ghaderi R. Application of the PSO–SVM model for recognition of control chart patterns. ISA transactions. 2010;49(4):577-86. https://doi.org/10.1016/j.isatra.2010.06.005
11. Lee P-H, Torng C-C, Lin C-H, Chou C-Y. Control chart pattern recognition using spectral clustering technique and support vector machine under gamma distribution. Computers & Industrial Engineering. 2022;171:108437. https://doi.org/10.1016/j.cie.2022.108437.
12. Swift JA, Mize JH. Out-of-control pattern recognition and analysis for quality control charts using lisp-based systems. Computers & Industrial Engineering. 1995;28(1):81-91. https://doi.org/10.1016/0360-8352(94)00028-L.
13. Zaman M, Hassan A. Improved statistical features-based control chart patterns recognition using ANFIS with fuzzy clustering. Neural Computing and Applications. 2019;31(10):5935-49. https://doi.org/10.1007/s00521-018-3388-2.
14. Wang C-H, Guo R-S, Chiang M-H, Wong J-Y. Decision tree based control chart pattern recognition. International Journal of Production Research. 2008;46(17):4889-901. https://doi.org/10.1080/00207540701294619.
15. Zan T, Liu Z, Wang H, Wang M, Gao X. Control chart pattern recognition using the convolutional neural network. Journal of Intelligent Manufacturing. 2020;31:703-16. https://doi.org/10.1007/s10845-019-01473-0.
16. Cheng C-S, Chen P-W, Hsieh Y-C, Wu Y-T. Multivariate Process Control Chart Pattern Classification Using Multi-Channel Deep Convolutional Neural Networks. Mathematics. 2023;11(15):3291. https://doi.org/10.3390/math11153291.
17. Golilarz NA, Addeh A, Gao H, Ali L, Roshandeh AM, Munir HM, Khan RU. A new automatic method for control chart patterns recognition based on ConvNet and harris hawks meta heuristic optimization algorithm. Ieee Access. 2019 Oct 4;7:149398-405. DOI: 10.1109/ACCESS.2019.2945596.
18. Yu Y, Zhang M. Control chart recognition based on the parallel model of CNN and LSTM with GA optimization. Expert Systems with Applications. 2021;185:115689. https://doi.org/10.1016/j.eswa.2021.115689.
19. Cheng C-S, Chen P-W, Ho Y. Control Chart Concurrent Pattern Classification Using Multi-Label Convolutional Neural Networks. Applied Sciences. 2022;12(2):787. https://doi.org/10.3390/app12020787.
20. Wu C, Hou H, Lei C, Zhang P, Du Y. A Novel Scheme of Control Chart Patterns Recognition in Autocorrelated Processes. Mathematics. 2023;11(16):3589. https://doi.org/10.3390/math11163589.
21. Xue L, Wu H, Zheng H, He Z. Control chart pattern recognition for imbalanced data based on multi-feature fusion using convolutional neural network. Computers & Industrial Engineering. 2023;182:109410. https://doi.org/10.1016/j.cie.2023.109410.
22. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014. https://doi.org/10.48550/arXiv.1409.1556.
23. Gao Z, Chen N, Yang Y, Li L. An innovative multisource multibranch metric ensemble deep transfer learning algorithm for tool wear monitoring. Advanced Engineering Informatics. 2024;62:102659. https://doi.org/10.1016/j.aei.2024.102659.
24. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L, editors. Imagenet: A large-scale hierarchical image database. 2009 IEEE conference on computer vision and pattern recognition; 2009: Ieee. 10.1109/CVPR.2009.5206848.
25. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems. 2012;25. https://doi.org/10.1145/3065386.
26. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. InProceedings of the IEEE conference on computer vision and pattern recognition 2015 (pp. 1-9).10.1109/CVPR.2015.7298594.
27. He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. https://doi.org/10.48550/arXiv.1512.03385.
28. Ioffe S, Szegedy C, editors. Batch normalization: Accelerating deep network training by reducing internal covariate shift. International conference on machine learning; 2015: pmlr. https://doi.org/10.48550/arXiv.1502.03167.
29. Li J, She Q, Fang F, Chen Y, Zhang Y. Multi-Source Transfer Learning via Optimal Transport Feature Ranking for EEG Classification. Neurocomputing. 2024:127944. https://doi.org/10.1016/j.neucom.2024.127944.
30. Chen J, Li J, Peng Y, Wang Z, Sun T, Zhu Y. In-situ manifestation of lapping mechanisms by rapid intelligent pattern recognition analysis (RIPRA) of acoustic emission via a point density fuzzy C-means (PD-FCM) method. Journal of Manufacturing Processes. 2022;79:924-33. https://doi.org/10.1016/j.jmapro.2022.05.035.
31. Guh R-S, Tannock J. A neural network approach to characterize pattern parameters in process control charts. Journal of Intelligent

Manufacturing. 1999;10:449-62. https://doi.org/10.1023/A:1008975131304.

32. Hassan A, Baksh MSN, Shaharoun AM, Jamaluddin H. Improved SPC chart pattern recognition using statistical features. International Journal of Production Research. 2003;41(7):1587-603. https://doi.org/10.1080/0020754021000049844.
33. Assaleh K, Al-Assaf Y. Features extraction and analysis for classifying causable patterns in control charts. Computers & industrial engineering. 2005;49(1):168-81. https://doi.org/10.1016/j.cie.2005.01.021.
34. Cheng Z, Ma Y, editors. A research about pattern recognition of control chart using probability neural network. 2008 ISECS International Colloquium on Computing, Communication, Control, and Management; 2008: IEEE. 10.1109/CCCM.2008.168.
35. Zan T, Wang M, Fei RY. Pattern recognition for control charts using AR spectrum and fuzzy ARTMAP neural network. Advanced Materials Research. 2010;97:3696-702. https://doi.org/10.4028/www.scientific.net/AMR.97-101.3696.
36. Ranaee V, Ebrahimzadeh A. Control chart pattern recognition using neural networks and efficient features: A comparative study. Pattern Analysis and Applications. 2013;16:321-32. https://doi.org/10.1007/s10044-011-0246-6.
37. Kao L-J, Lee T-S, Lu C-J. A multi-stage control chart pattern recognition scheme based on independent component analysis and support vector machine. Journal of Intelligent Manufacturing. 2016;27:653-64. https://doi.org/10.1007/s10845-014-0903-x.
38. Addeh A, Khormali A, Golilarz NA. Control chart pattern recognition using RBF neural network with new training algorithm and practical features. ISA transactions. 2018;79:202-16. https://doi.org/10.1016/j.isatra.2018.04.020.
39. Maleki MR, Sahraeian R. Online monitoring and fault diagnosis of multivariate-attribute process mean using neural networks and discriminant analysis technique. International Journal of Engineering-Transactions B: Applications. 2015;28(11):1634-42. 10.5829/idosi.ije.2015.28.11b.11.

COPYRIGHTS

©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

تشخیص الگوی نمودار کنترل، یک ابزار کنترل فرآیند آماری حیاتی است که برای تعیین اینکه آیا یک فرآیند در محدوده پارامترهای مورد نظر عمل می‌کند یا رفتار غیرعادی از خود نشان می‌دهد، استفاده می‌شود. تشخیص دقیق و خودکار برای شرکت‌های تولیدی جهت حفظ تولید با کیفیت بالا ضروری است. مطالعات اخیر، یادگیری ماشین را با تشخیص الگوی نمودار کنترل ترکیب کرده‌اند، با این حال این روش‌ها اغلب فضاهای ویژگی و توزیع‌های آماری یکسانی را بین داده‌های آموزشی و آزمایشی فرض می‌کنند – شرایطی که به ندرت در عمل برآورده می‌شود. آموزش مجدد مدل‌ها از ابتدا با داده‌های جدید زمان‌بر است. یادگیری انتقالی با بهره‌گیری از دانش مدل‌های از پیش آموزش‌دیده، جایگزین کارآمدتری ارائه می‌دهد. این مطالعه از شبکه عصبی کانولوشن عمیق VGG-16 از پیش آموزش‌دیده برای طبقه‌بندی الگوهای نمودار کنترل استفاده می‌کند و عملکرد تشخیص را بدون آموزش مجدد گسترده به طور قابل توجهی بهبود می‌بخشد. از طریق شبیه‌سازی‌های مونت کارلو و یک مطالعه موردی در دنیای واقعی، روش پیشنهادی به دقت تشخیص ۹۹.۳٪ دست یافت که از مدل‌های MLP معمولی، 1DCNN و CNN بهتر عمل می‌کند. تجزیه و تحلیل حساسیت نشان داد که استفاده از تکنیک‌های فعال‌سازی ReLU، نرمال‌سازی دسته‌ای و dropout، دقت و استحکام مدل را به طور قابل توجهی افزایش می‌دهد. نتایج، پتانسیل تشخیص‌دهنده مبتنی بر یادگیری انتقالی پیشنهادی را برای بهبود کنترل کیفیت هوشمند در فرآیندهای تولید با مدیریت مؤثر داده‌های آموزشی محدود و استخراج ویژگی‌های پیچیده تأیید می‌کند.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1226-1237

International Journal of Engineering

Journal Homepage: www.ije.ir

Ensuring Vibration Strength of Mainline Pumping Unit Systems Using Digital Prototypes

I. N. Budilov[a], S. N. Tyulyandin*[b]

[a] Department of Mechanics, Empress Catherine II Saint Petersburg Mining University, Saint Petersburg, Russia
[b] Department of Transport and Storage Oil and Gas, Empress Catherine II Saint Petersburg Mining University, Saint Petersburg, Russia

PAPER INFO

Paper history:
Received 05 May 2025
Received in revised form 21 June 2025
Accepted 14 July 2025

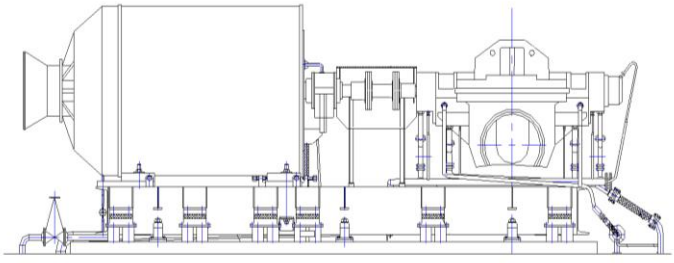
Keywords:
Computational Modeling
Digital Prototype
Finite Element Method
Mainline Pumping Unit
Multiphysics Models
Vibration-compensation System

ABSTRACT

The relevance of this research stems from the necessity to enhance existing vibration-compensation systems for reducing vibrational loads on power units of mainline pumping units (MPUs). Algorithms and computational results of the spatial vibration state of the vibration-compensation system (VCS) for MPUs under various operational regimes are presented. Mathematical and spatial finite-element models, as well as versatile digital prototypes of the VCS, have been developed. The novelty of the findings lies in the developed digital prototype of the MPU, which offers broad capabilities for accounting for design-technological factors, accommodates diverse operational regimes, and exhibits high precision and computational efficiency. The practical significance of the work is that the developed simulation methodology based on digital prototypes of the MPU VCS enables the analysis of vibrational impacts from electric motors, pumps, and load-bearing frames. Comprehensive computational studies of the stress-strain state of the VCS frame were performed, and the system's natural frequencies during operation were evaluated. The objective of this work is to study the behavior of the pumping unit when replacing elastomeric supports in the compensation system with hydrofilm dampers. An indisputable advantage of the developed digital model is its foundation for a universal methodology to assess the effectiveness of vibration compensators in the MPU VCS. A unique digital prototype of the MPU was developed considering the engineered VCS. The feasibility of using a finite-element-based digital prototype for evaluating vibrational loading on MPUs is substantiated. Numerical experiments using the MPU's digital twin established that implementing hydrofilm dampers in the VCS reduces peak dynamic stresses by 18%. Values of axial and resultant displacements of the MPU using hydrofilm and elastomeric dampers in the VCS were obtained, measuring 11.1 mm and 10.0 mm, respectively.

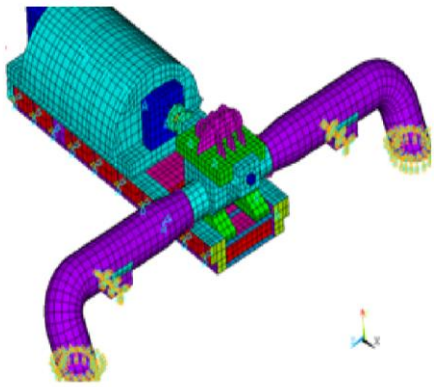
***doi**: 10.5829/ije.2026.39.05b.15*

Graphical Abstract



Resonant frequencies of the rigidly mounted MPU

Frequency, Hz						
f_1	f_2	f_3	f_4	f_5	f_6	f_7
38,2	52,2	55,3	59,1	62,4	65,7	66,8



* Corresponding Author Email: s235105@stud.spmi.ru (S. N. Tyulyandin)

Please cite this article as: Budilov IN, Tyulyandin SN. Ensuring Vibration Strength of Mainline Pumping Unit Systems Using Digital Prototypes. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1226-37.

1. INTRODUCTION

Mainline pumping units (MPUs) are widely used in oil and gas transportation systems to maintain pressure during fluid transfer and are integral components of oil pumping stations (OPS).

VCSs for MPUs often include bellows compensators to reduce static and dynamic loads transmitted from pipelines to the pump.

However, computational and experimental analyses (1) demonstrate limited compensation efficiency due to the pressure-dependent stiffness of bellows compensators, which becomes comparable to the stiffness of pipeline bends during MPU operation.

The generally accepted global practice in the design and manufacturing of mainline pumping units (MPUs) involves the expert evaluation of technical solutions based on digital prototypes (2, 3). In the field of improving the design of vibration compensation systems (VCS), notable works are known (4, 5). By utilizing digital prototypes and relying on the fundamental laws of continuum mechanics (CM) embedded within them, as well as through the application of modern computational methods (6) it becomes possible to track trends in load increases when changing the operating parameters of VCS.

Digital prototypes also facilitate process optimization by isolating random factors inherent in real-world production and operation.

MPUs must maintain operational integrity under standard loads from connected pipelines, as per regulatory code NP-068-95[1]

Natural frequency calculations are mandatory for vibration strength verification (7-9), complying with SNiP 2.05.06-85 (10), GOST 32388 (11), and other standards (12-15), ensuring pipeline system reliability and safety (16-18) while accounting for gas distribution failure statistics (19, 20). MPU design per PNAE G-7-002-86 requires static strength, brittle fracture resistance, and vibration strength analyses (21). Studies (22-24) address pipeline crack propagation and fracture toughness assessment.

The implementation of novel approaches, particularly during initial design stages—specifically through the use of fuzzy Bayesian belief networks (FBBNs) (25)- enables significant risk reduction in addressing systemic challenges across project lifecycle phases. Machine learning further ensures high operational efficiency during preliminary prototype development (26).

This work validates the feasibility of equipping MPUs with vibration-compensation systems and selects optimal damper types. Stress-strain states and vibration levels were calculated for the NM 7000–210/STDP–6300–2BUKhL4/10000 MPU under static (weight, temperature, pressure) and dynamic (rotor unbalance) loads.

Four foundation mounting configurations were analyzed: rigid supports vs. dampers (AMP-1700, A-2000, AGP-2.1).

VCS Design Description. The external view of the NM 7000–210 MPU is shown in Figure 1.

The primary load-bearing components of the VCS include: the load-bearing frame, elastomeric dampers, a flexible coupling and a horizontal centrifugal pump with a double-suction impeller. The MPU is driven by an asynchronous STDP-6300 electric motor with a power output of up to 10 kW (25).

Vibrations in the MPU system arise from dynamic interactions between the pump-motor assembly and the load-bearing frame, which are subsequently transmitted to the foundation (26, 27). The pump and electric motor form a coupled dynamic system, generating significant vibrational loads (28, 29).

Since the MPU incorporates structural components with heterogeneous stiffness, these assemblies experience uneven dynamic loading. Consequently, the service life of the MPU is primarily determined by operational modes, pump inlet pressure, and the structural integrity of its flow path components (30, 31). The unit is also subjected to electromagnetic and hydrodynamic forces.

Frequent transitions between operational regimes—from nominal operation to start/stop cycles—coupled with prolonged transient-state operation, induce significant dynamic stress in load-bearing assemblies, adversely affecting the MPU's overall operational reliability.

Notably, the annual number of MPU starts can exceed 300, resulting in substantial dynamic loading intensity that further compromises operational durability (32).

A distinctive feature of the VCS is its incorporation of reactive supports equipped with elastomeric dampers. The primary objective in designing the VCS is to mitigate

Figure 1. The external view of the NM 7000–210 MPU/ STDP- 6300

[1] NP-068-95, Federal norms and rules in the field of atomic energy use «Pipeline fittings for nuclear power plants. General technical requirements».

dynamic loads acting on the MPU's load-bearing components.

2. 1. Problem Statement Despite the high operational efficiency of VCS in industrial applications, their drawback lies in the use of angular rubber-cord or straight-through compensators[1]. Current practice widely adopts hydrofilm dampers without compensators. Based on numerical modeling results of vibration load fields for both damped and rigidly mounted MPUs, recommendations are provided for installing MPUs on hydrofilm dampers.

Consider a VCS comprising: inlet pipes: 820×12 mm, weight per linear meter (oil-filled): 666.67 kg; outlet bend: Ø820 mm, mass (oil-filled): 678 kg.

The adjustable vibration-damping pipeline support consists of a steel plate with connecting and reinforcing elements, fitted with three elastomeric dampers (APM-1700) as shown in Figure 2.

2. 2. Spectral Analysis Equations In oil and gas field mechanics, the finite element method (FEM) (33), implemented in modern computational platforms, has gained widespread adoption[2].

TABLE 1. Characteristics of APM-1700 Dampers

Rated load, kgf	P_Y	**2000,0**
	K_{xi}	410,0
Static stiffness, kN/m	K_{zi}	330,0
	K_{yi}	5400,0
	K_{xi}	637,0
Dynamic stiffness, kN/m	K_{zi}	588,0
	K_{yi}	6020,0
Allowable vertical deformation, mm		32,0
Allowable horizontal deformation, mm		18,2

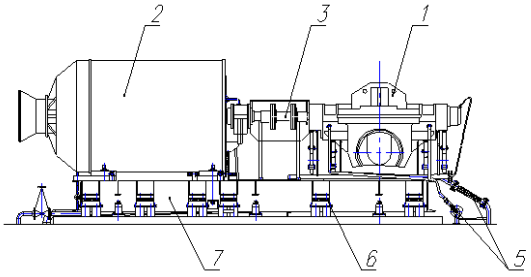


Figure 2. MPU NM 7000–210 MPU/ STDP-6300-2
1. Pump, 2. Electric motor, 3. UKM coupling, 4. MKS compensator-vibration dampers, 5. Flexible connections of auxiliary pipelines, 6. Elastomeric damping supports, 7. Reinforced unit frame; 8. Reactive supports for pipelines

The governing equation for evaluating undamped oscillations using spectral analysis can be expressed as follows (34-36):

$$[M]\{u''\}+[C]\{u'\}+[K]\{u\}=\{F\} \quad (1)$$

where [M]; [C], [K] - mass, damping, and stiffness matrices;

{u''},{u'},{u} - nodal acceleration, velocity, and displacement vectors.

The time-dependent force vector {F} is expressed as:

$$\{F\}=\{F^{nd}\}+s\{F^{s}\} \quad (2)$$

where: $\{F^{nd}\},\{F^{s}\}$ - load vectors; s - scaling coefficient.

The load vector $\{F^{s}\}$ is computed during modal analysis.

Introducing the substitution,

$$\{u\}=\sum_{i=1}^{n}\{\varphi_i\}y_i \quad (3)$$

into the system's equation of motion yields the governing equation:

$$[M]\sum_{i=1}^{n}\{\varphi_i\}y_i''+[C]\sum_{i=1}^{n}\{\varphi_i\}y_i'+[K]\sum_{i=1}^{n}\{\varphi_i\}y_i=\{F\} \quad (4)$$

where: $\{\varphi_i\}$ - i-th mode shape; n - number of modes; y_i;- modal coordinates.

Multiplying by $\{\varphi_j\}^T$, we derive the final equation:

$$\{\varphi_j\}^T[M]\sum_{i=1}^{n}\{\varphi_i\}y_i''+\{\varphi_j\}^T[C]\sum_{i=1}^{n}\{\varphi_i\}y_i'+ +\{\varphi_j\}^T[K]\sum_{i=1}^{n}\{\varphi_i\}y_i=\{\varphi_j\}^T\{F\} \quad (5)$$

or equivalently:

$$\{\varphi_j\}^T[C]\{\varphi_j\}=2\xi_j\sqrt{K_jM_j}\left(\frac{1}{\sqrt{M_j}}\right)^2=2\xi_j\omega_j \quad (6)$$

Introducing the notation

$$f_j=\{\phi_j\}^T\{F\} \quad (7)$$

The equation of motion becomes:

$$y_j''+2\omega_j\xi_jy_j'+\omega_j^2y_j=f_j \quad (8)$$

[1] Gumerov A.G., Gumerov R.S, Akberdin A.M. Diagnostics of oil pumping station equipment. M.: OOO «Nedra-Biznestsentr», 2003, 347 p.

[2] Finite element method in problems of oil and gas field mechanics. N. Alikin, I.E. Litvin, S.M. Shcherbakov, V.P. Borodavkin. M.: Nedra, 1992, 288 p.

2. 3. Description of Finite Element Model The digital prototype of the MPU, shown in Figure 3, includes the following components: the unit frame, the STDP-6300-2BUKhL4/10000 electric motor, the NM 7000–210 pump, the shaft assembly with a coupling, the inlet pipeline section, and the outlet pipeline section.

To assess the influence of the frame damper stiffness characteristics on the stress-strain state (SSS) of the MPU, three damped mounting configurations were analyzed: mounting on elastomeric dampers (APM-1700), mounting on elastomeric dampers (A-2000), and mounting on hydrofilm dampers (AGP-2.1)[1].

Simplifications in the model involve replacing mounting dampers with rod elements represented as spring-type elastic elements (COMBINE14) in three directions, accounting for the stiffness of actual dampers specified in Table 3. The model employed rigid contacts between connecting nodes of structural elements using "surface-to-surface" and "point-to-surface" types, with slippage of contact elements disregarded.

The finite element model (FEM) of the MPU incorporating boundary conditions is illustrated in Figure 4.

FEM Constraints: A distinctive feature of the FEM is the assignment of reduced density for various structural components. The mass of the inlet/outlet pipelines was calculated considering the contained fluid product, the motor mass corresponded to its simplified model representation, and the same approach was applied to the pump. The element size was selected based on the smallest constructible mesh dimension. The final FEM comprised 800,000 hexahedral elements, balancing analytical precision and computational efficiency.

Thermal stresses and subsequent total stresses in the pipes were derived via the superposition principle. Incremental calculations summed mechanical and thermal deformations using the formula:

$$\Delta\varepsilon^0 = \varepsilon^e + \varepsilon^T \quad (9)$$

where ε^e - elastic deformation increment, and $\Delta\varepsilon^T=\alpha\times\Delta T$ - increment of components accounting for thermal deformations, where α is the coefficient of linear thermal expansion.

2. 4. Calculation of Stress-Strain State of Nm 7000–210/STDP–6300–2 Unit The objective of this computational experiment is to perform a computational assessment of the stress-strain state (SSS) of the NM 7000–210/STDP–6300 MPU. Two installation configurations of the unit frame on the PS foundation were analyzed: mounting on rigid supports or on dampers, with pipelines resting on an adjustable vibration-damping support. This support consists of a steel plate with connecting and reinforcing elements, equipped with three elastomeric APM–1700 dampers.

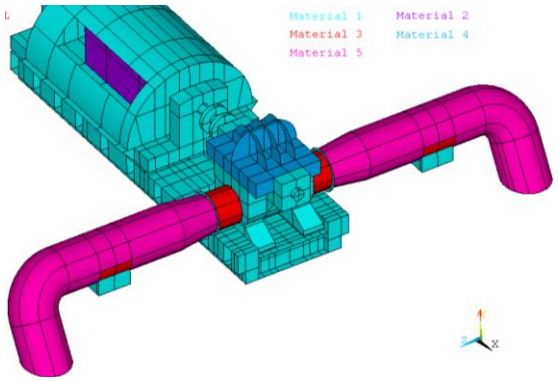


Figure 3. Solid-state model of MPU

The SSS evaluation of the MPU was conducted under the following loads: gravitational force (37), starting torque M_{st}=45410 N/m, temperature differential $\Delta t=20^0$C, inlet pipeline pressure p_1=2,8 MPa, and outlet pipeline pressure p_2=4,8 MPa. The SSS and vibration analysis for both installation configurations were performed using the finite element method (FEM) based on a digital platform (38).

2. 5. Description of Digital Prototype The FEM mathematical model employed five material types, whose properties are listed in Table 2. All materials corresponded to Grade 20 structural steel, with variations only in reduced density assignments to match the weight of individual components. The table also specifies their applications.

Material Group 1 corresponds to standard steel properties. For Groups 2–5, adjusted density values were assigned to ensure equivalence between the mathematical model's mass and the actual structural mass. Material Group 1 exhibits standard steel properties. Groups 2–5 feature reduced density values determined by equating

TABLE 2. Material Characteristics

№	E, Pa	ν	ρ, kg/m³	α·10⁵	Application
1	2,06·10¹¹	0,3	7850	1,1	Unit frame, shaft assembly
2	2,06·10¹¹	0,3	35500	1,1	Electric motor
3	2,06·10¹¹	0,3	15700	1,1	Pump, pipeline
4	2,06·10¹¹	0,3	12200	1,1	Pump
5	2,06·10¹¹	0,3	21700	1,1	Pipeline

[1] EIMA.304242.015 TU. Technical conditions. Metal plate shock absorbers of AMP type. Tekhnicheskie usloviya, Severodvinsk, PO «SevmaSH», 1996, 60 p.

the model's mass with the actual structural mass. Pipeline material corresponded to Grade 20 steel ([σ]=147,0 MPa, σ_B=410,0 MPa; σ_T=234,0 MPa), while the support frame material matched Grade 09G2S steel ([σ]=196,0 MPa, σ_B=500,0 MPa; σ_T=350,0 MPa).

The general configuration of the finite element model (digital prototype) of the MPU with boundary conditions is shown in Figure 4.

The finite element model includes components such as the unit frame, electric motor, shaft drive, coupling, pump, and pipeline. Shell-type finite elements were used for mesh generation, with displacement constraints applied as boundary conditions. A key constraint involves interactions between volumetric Solid 185 and shell-type SHELL63 elements with incompatible degrees of freedom. This limitation is resolved using next-generation SolSH 190 volumetric elements that maintain compatibility with shell elements (39).

The characteristics of the COMBINE14 elastic elements in terms of axial stiffness are provided in Tables 3 and 4.

3. RESULTS OF CALCULATIONS

The stress-strain state (SSS) of the rigidly mounted mainline pumping unit (MPU) was calculated under the following loads: gravitational force (g = 9,807 m/s^2 gravitational acceleration); starting torque M_{st} = 45,41 kN·m; temperature differential Δt = 20^{0}C; inlet pipeline pressure p_1 = 2,8 MPa; and outlet pipeline pressure p_2 = 4,8 MPa.

The SSS assessment under gravitational load was performed using static analysis in ANSYS/WORKBENCH 2019 R3. Calculations were performed on the workstation (processor: Intel i7-10920, memory: DDR4 64 GB, graphics processing unit (GPU): GeForce RTX 3070 Ti 12 GB, storage: Samsung SSD 2 TB). The total computation time was 4 hours. The general

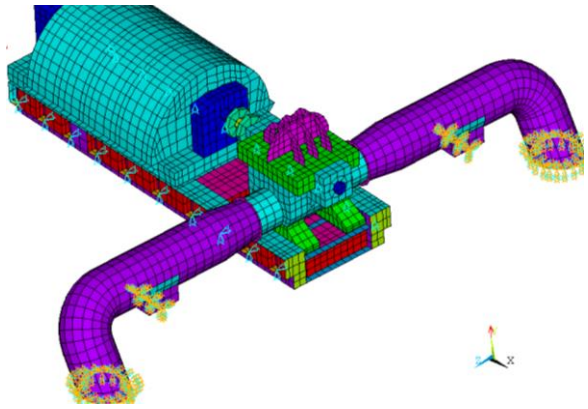


Figure 4. Mathematical model of the MPU accounting for boundary conditions

TABLE 3. Static characteristics of elastic elements

Designation	Static stiffness, N/m		
	AMP-1700	A–2000	AGP-2.1
K_X	0,41·10^6	1,47·10^6	0,91·10^6
K_Z	0,33·10^6	0,44·10^6	0,56·10^6
K_Y	5,40·10^6	2,16·10^6	1,98·10^6

TABLE 4. Dynamic characteristics of elastic elements

Designation	Dynamic stiffness, N/m		
	AMP-1700	A–2000	AGP-2.1
$\boldsymbol{K_X}$	0,64·10^6	1,96·10^6	1,35·10^6
$\boldsymbol{K_Z}$	0,59·10^6	0,84·10^6	0,95·10^6
$\boldsymbol{K_Y}$	6,02·10^6	2,84·10^6	2,10·10^6

view of the MPU mathematical model with boundary conditions and applied loads is shown in Figure 4. The SSS results are graphically presented as contour plots in Figures 5 to 8.

The von Mises criterion (40-43) was adopted as the basis for strength assessment. Stress distributions according to the von Mises criterion σ_{von} in the inlet pipeline and outlet pipelines are presented as contour bands and shown in Figure 8.

The maximum stress, σ_{max}=25,2 MPa, was observed at the central rigid support locations.

The SSS under combined operational loads was analyzed using an engineering analysis system (49, 50). The boundary conditions are illustrated in Figure 4.

Calculations in Figures 5a and 5b depict the loading scheme for the starting torque and pressure distribution in the pipeline and pump, respectively.

Figures 5a and 5b show the deformation pattern of the MPU with contour lines of resultant (Figure 5a) and vertical (Figure 5b) displacements. Dashed lines represent the undeformed state. Maximum displacements were calculated as follows: resultant displacement W_{max} = 0,0765·10^{-3} m; vertical displacement UY_{min} = – 0,0761·10^{-3} m.

Results in Figures 6 to 8 display contour plots of resultant, longitudinal, vertical, and lateral displacements. Extreme displacement values are: resultant: W_{max} = 1,72·10^{-3} m (Figure 5a); longitudinal: U_{Xmin} = – 1,437·10^{-3} м, U_{Xmax} = 1,422·10^{-3} m (Figure 5b); vertical: U_{Ymin} = – 0,250·10^{-3} m, U_{Ymax}=1,285·10^{-3} m (Figure 5b); lateral: U_{Zmin} = – 1,214·10^{-3} m, U_{Zmax} = 1,093·10^{-3} m.

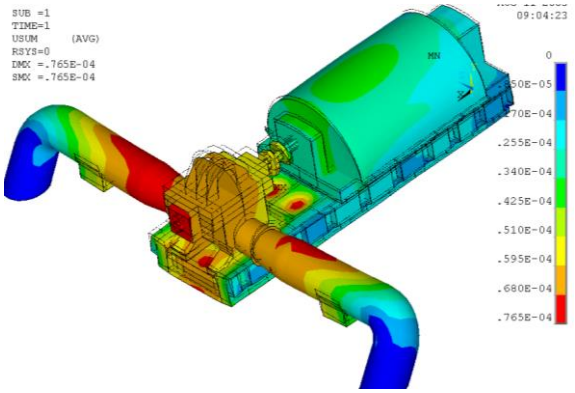
Structural strength was evaluated using the Mises criterion. Figure 8 shows the Mises stress distribution across the MPU and its components. Maximum stresses were: outlet pipeline: σ_{out} = 285 MPa; Inlet pipeline:

σ_{inl} = 234,0 MPa; pump casing: σ_{pump} = 101,0 MPa; MPU frame: σ_{frame} = 25,5 MPa.

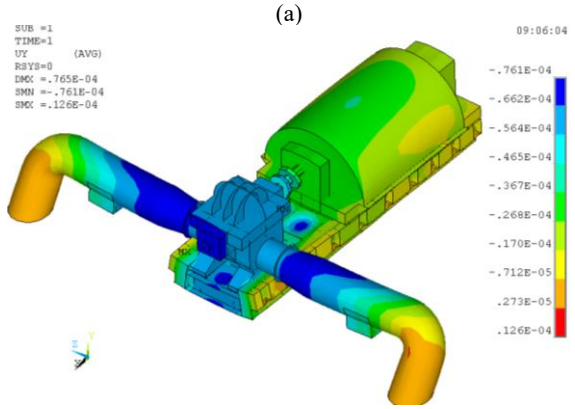
Thus, it has been established that the maximum stresses in the MPU occur at the outlet pipeline bend. This high-stress zone exhibits a localized character.

Forced vibration results for the rigidly mounted MPU are shown in Figures 9 to 14, including: -Frequency response of bearing and pipeline vibration displacements (Figures 9 and 10);

Vibration displacements at the motor bearing near the coupling and at the field-side motor bearing (Figure 10);

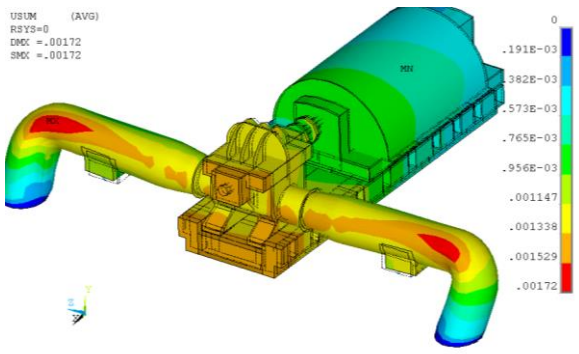


(a)

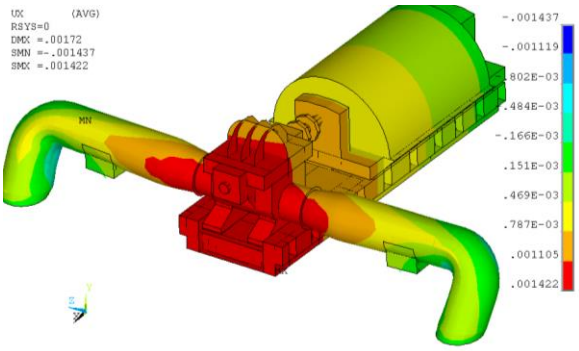


(b)

Figure 5. Resultant displacements (a) (m) and vertical displacements (b) (m) of the MPU

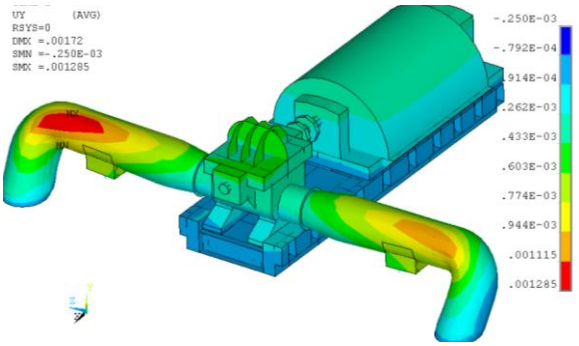


(a)

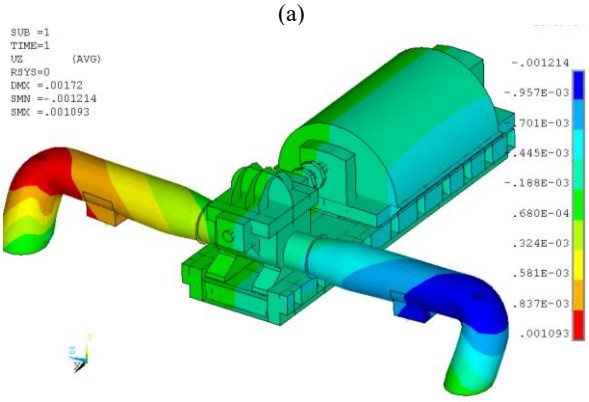


(b)

Figure 6. Total and longitudinal displacements of MPU (m)

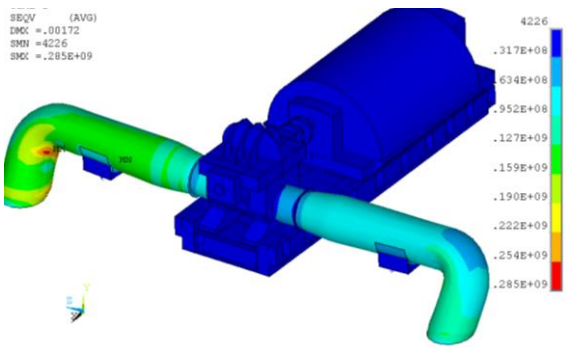


(a)



(b)

Figure 7. Vertical (a) and lateral (b) displacements of the MPU (m)



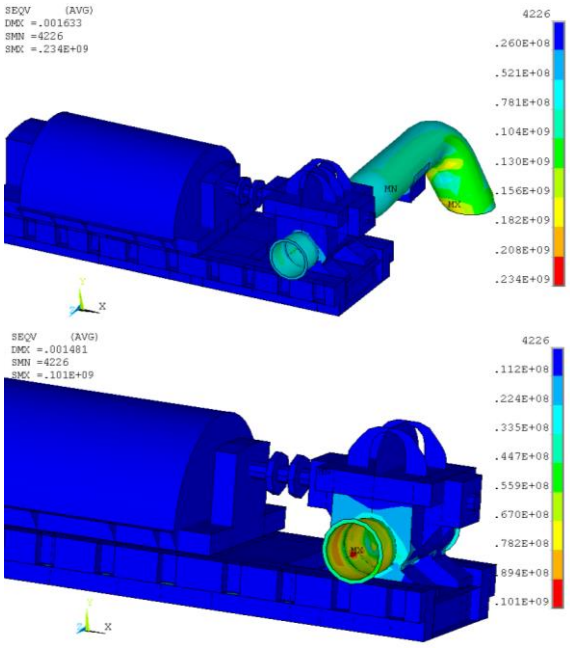
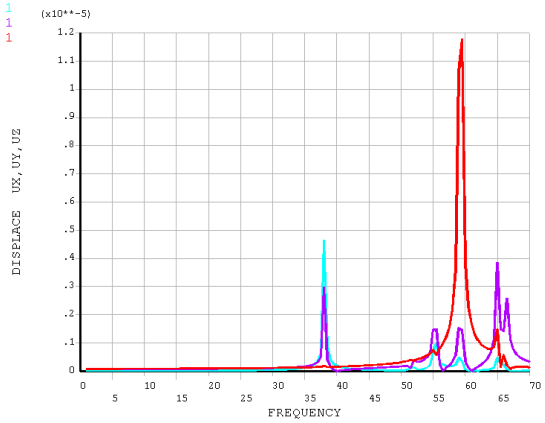
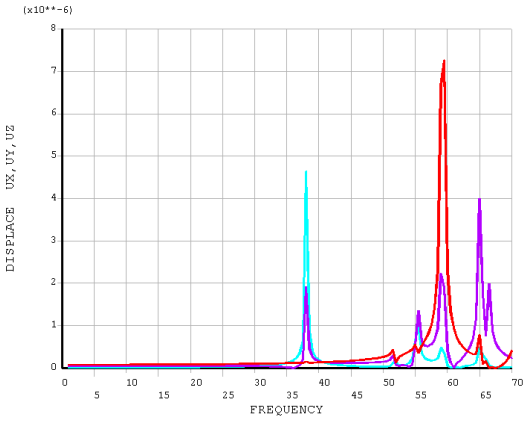


Figure 8. Equivalent stresses in the MPU inlet pipeline (Pa)

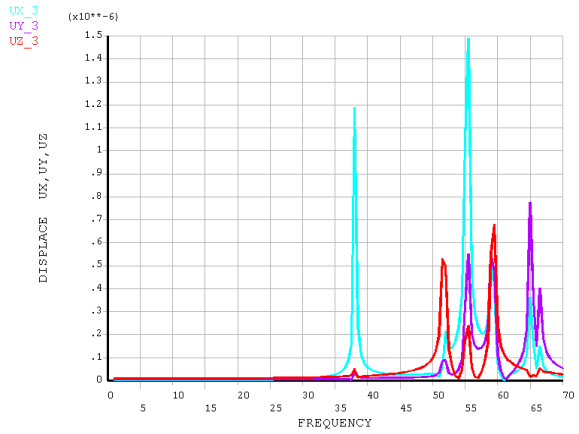


(a)

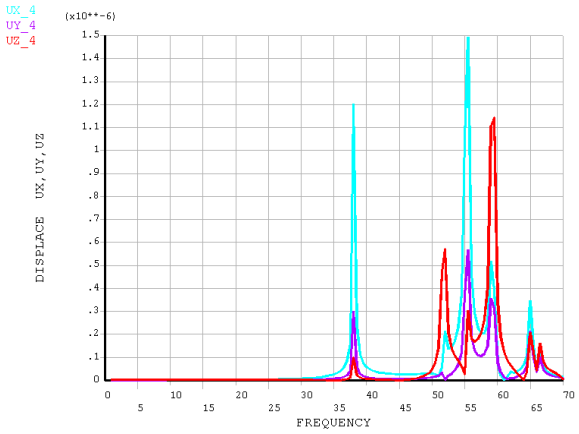


(b)

Figure 9. Vibration displacements of the non-damped MPU at the field-side pump bearing (a) and at the pump bearing near the coupling (b)



(a)



(b)

Figure 10. Vibration displacements at the motor bearing near the coupling (a) and at the field-side motor bearing (b) of the non-damped MPU

Vibration displacements at the outlet and inlet pipelines (Figure 11) and vibration displacements under the field-side motor bearing and motor bearing near the coupling (Figure 12);

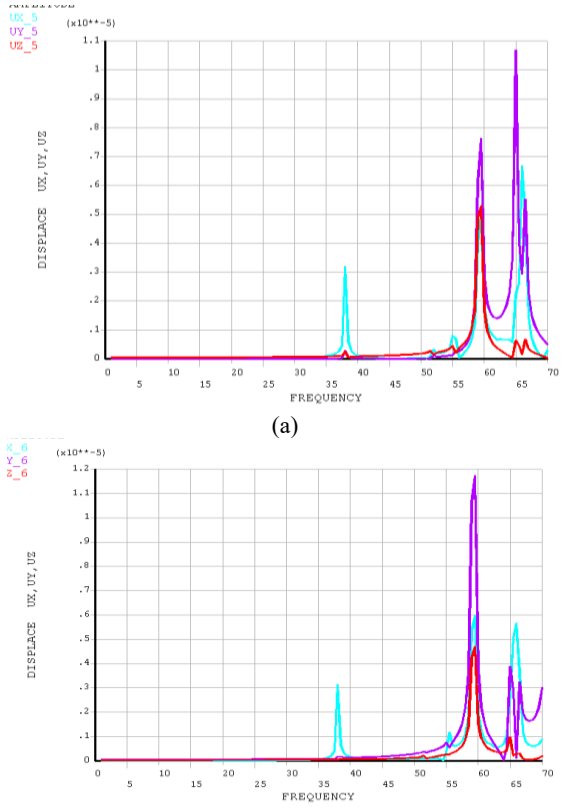
The vibration responses of the non-damped MPU at points beneath the pump bearing are shown in Figure 13., and forced vibration mode shapes at resonant frequencies (Figure 14).

The calculated resonant frequencies of the rigidly mounted MPU are presented in Table 5. As evident from the data in Table 5, the first seven resonant frequencies of forced vibrations for the non-damped MPU nearly coincide with its natural frequencies (considering the frequency increment used in the calculations).

Of particular note is the presence of a dense cluster of intense resonances near the rotational frequency of the steady-state regime, as well as a resonance at 38 Hz. When passing through this frequency during startup, the accelerating rotor system of the unit (with a rotational speed of 2280 rpm) generates significant centrifugal excitation forces.

4. DISCUSSION OF RESULTS

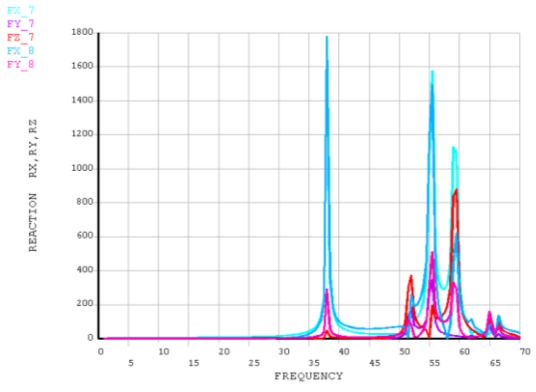
Analysis of the digital prototype results revealed that the maximum stresses from static loads occur at the outlet pipeline bend, reaching 285.0 MPa for the non-damped MPU and 289.0 MPa for the damped MPU. Resultant pipeline displacements were 1.72 mm for the non-damped configuration, 2.67 mm for MPUs on AMP-1700 dampers, 2.29 mm for A-2000 dampers, and 2.49 mm for AGP-2.1 dampers.



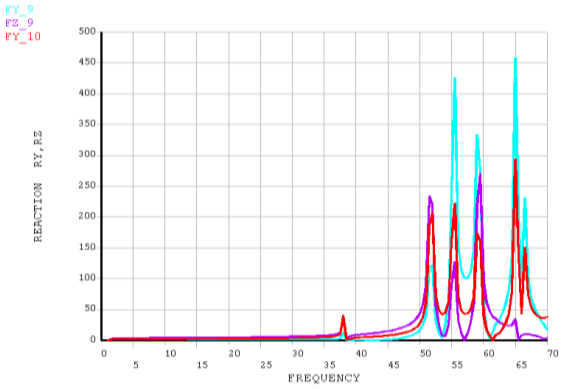
(a)

(b)

Figure 11. Vibration displacements at the outlet pipeline (a) and inlet pipeline (b) of the non-damped MPU

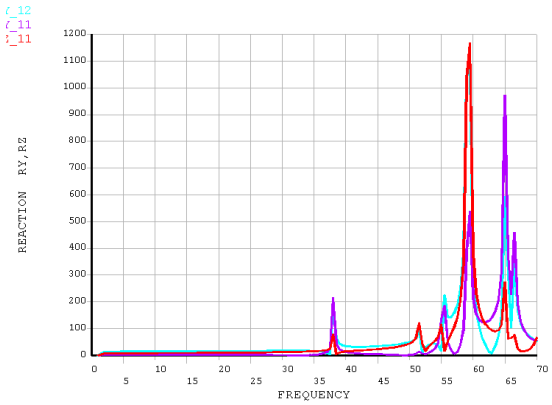


(a)

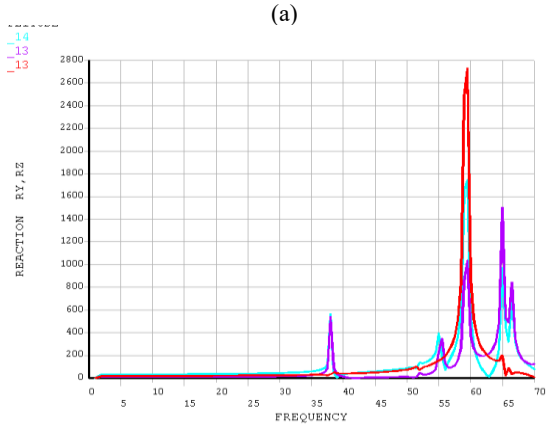


(b)

Figure 12. Vibration responses at points under the field-side motor bearing and under the motor bearing near the coupling of the non-damped MPU



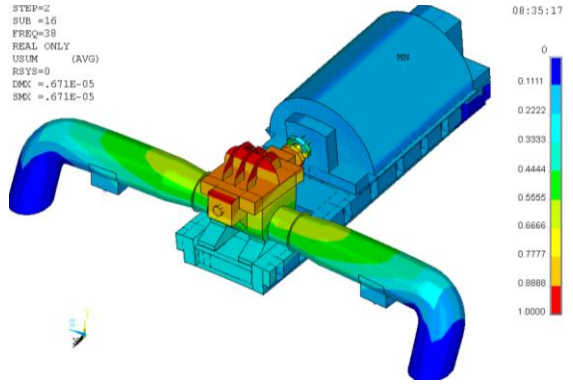
(a)



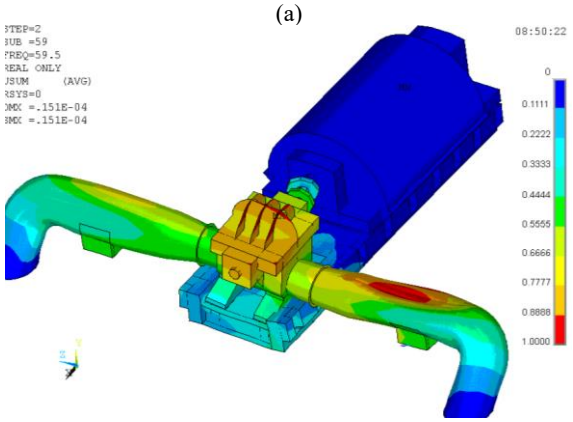
(b)

Figure 13. Vibration responses of the non-damped MPU at points under the pump bearing]

Under steady-state operation at the rotational frequency, the pump displacement was 8.0×10^{-6} m for the non-damped MPU, 6.0×10^{-6} m for MPUs on elastomeric AMP-1700 or A-2000 dampers, and 5.0×10^{-6} m for MPUs on AGP-2.5 hydrofilm dampers. Maximum stresses at the pump-pipe connection were 4.0 MPa for



(a)



(b)

Figure 13. Forced vibration mode shapes of the non-damped MPU at 38 Hz (a) and 59.5 Hz (b)

TABLE 5. Dynamic characteristics of elastic elements

Frequency, Hz						
f_1	f_2	f_3	f_4	f_5	f_6	f_7
38,2	52,2	55,3	59,1	62,4	65,7	66,8

both non-damped and damped configurations, and 3.0 MPa for the damped configuration.

Vibration velocity amplitudes at the rotational frequency under steady-state conditions were: 2.66 mm/s for the non-damped MPU; 2.04 mm/s for MPUs on elastomeric AMP-1700/A-2000 dampers; 1.57 mm/s for MPUs on hydrofilm dampers.

The maximum levels of forced pump vibrations during startup were also obtained and analyzed. Thus, for the non-damped MPU, the maximum forced vibration levels of the pump amount to $8{,}0\cdot10^{-5}$ m with a pipe stress of 16.0 MPa at the pump-pipe connection area; for the MPU on elastomeric dampers – $5{,}0\cdot10^{-5}$ m with a pipe stress of 9.0 MPa at the pump-pipe connection area; and for the MPU on hydrofilm dampers – $4{,}0\cdot10^{-5}$ m with a pipe stress of 6.0 MPa at the pump-pipe connection area.

The maximum amplitude of forced vibration velocity during startup for the non-damped MPU is 19.20 mm/s, for the MPU on elastomeric dampers – 11.30 mm/s, and for the MPU on hydrofilm dampers – 9.01 mm/s.

The vibrational and dynamic characteristics of the non-damped MPU, obtained through computational analysis, show good agreement with experimental data. This also confirms the reliability of the computational results for the damped MPU.

Thus, based on the data obtained in this study, it can be concluded that the numerical experiments confirm the effectiveness of hydrofilm dampers within the VCS, as they reduce the overall vibrational loading on the MPU by 18.5%.

Experimental research documented in the field (7-9) and measurements conducted on operational units corroborate the findings of this numerical study.

The obtained data on the assessment and reduction of vibrational stress using hydrofilm compensators are explained by their design features and hysteresis loop characteristics under cyclic loading conditions.

The key strength of this work lies in the first-time development of a digital prototype and computational methodology for the vibration-compensation system of compressor stations. This approach enables the investigation of numerous design parameters and non-standard operational modes affecting vibrational stress in compressor stations, constituting the core advantage of this digital experiment over established methodologies.

5. CONCLUSIONS

The results of the calculations and experiments demonstrated the following:

1. Equipping the unit with dampers shifts its primary resonances to the low-frequency range, which significantly increases inertial resistance to transient processes and excitation forces, shifts resonances to regions of relatively low centrifugal excitation forces generated by the accelerating rotor system, and consequently reduces vibration amplitudes at resonant frequencies during startup.

2. The maximum stresses in the MPU pipeline occur at the bend due to static loads, are virtually independent of damping, and reach 280.0 MPa. The wall thickness of the bend was assumed equal to the pipeline wall thickness.

3. The maximum stresses from dynamic loads in the MPU pipeline occur at the pump connection point. Under steady-state operation at the rotational frequency, these stresses are also independent of damping.

4. Maximum stresses and vibration velocity amplitudes in the MPU pipeline caused by dynamic startup loads are lower in units with hydrofilm dampers

compared to elastomeric dampers, and significantly lower than in non-damped units.

5. Frequency responses of the damped MPU exhibit a resonance at 51.5 Hz, dangerously close to the operational rotational frequency. This resonance is inherent to the frame design and independent of damper type.

7. Maximum resultant displacements of the MPU near the field-side motor bearing were recorded as 10.3 mm for A-2000 dampers and 11.1 mm for AGP-2.1 dampers.

8. The developed universal digital prototype of the mainline pumping unit with piping systems and vibration-compensation system, along with the novel methodology for forced vibration calculation, enables:

- investigation of transient operational modes of the MPU;
- accounting for thermal conditions, pipeline pressures, and other factors affecting system vibration strength;
- future optimization of the load-bearing frame design to reduce mass-dimensional parameters while maintaining vibration strength reliability.

Author contributions:

Budilov Igor Nikolaevich, Dr. Sci. (Tech), Prof. of the Department of Mechanics, Empress Catherine II Saint Petersburg Mining University – research concept and design;

Tyulyandin Sergey Nikolaevich postgraduate student of the Department of Transport and Storage of Oil and Gas, Empress Catherine II Saint Petersburg Mining University – data collection, processing, statistical analysis, text drafting, and editing.

6. REFERENCES

1. Tokarev A, Zotov A, Valeev A. The application of passive vibroprotective systems having power characteristics with hysteresis loops of rectangular shape for the main pumping units. Procedia Engineering. 2017;176:118-27. 10.1016/j.proeng.2017.02.279
2. Sukhorukov A, Koryagin N, Eroshkin SY, Kovkov D, editors. Statistical modeling of the process of generating analog information in the problems of the digital economy. 2017 Tenth International Conference Management of Large-Scale System Development (MLSD); 2017: IEEE. 10.1109/MLSD.2017.8109691
3. Cherepovitsyn AE, Tretyakov NA. Development of a new assessment system for the applicability of digital projects in the oil and gas sector. Записки Горного института. 2023(262 (eng)):628-42.
4. Mukherjee S. A theoretical study and 3D modeling of nonlinear passive vibration isolator. International journal of Applied and Advanced Scientific Research (IJAASR). 2017;2(2). 10.5281/zenodo.996576
5. Werner U. Active vibration control for rotating machines with current-controlled electrodynamic actuators and velocity feedback of the machine feet based on a generalized mathematical formulation. Control Theory and Technology. 2025;23(1):1-27. 10.1007/s11768-024-00230-w
6. Fomichev P, Fomicheva E, editors. Optimization problems of the characteristics of vibroprotective systems of a new type. IOP Conference Series: Earth and Environmental Science; 2019: IOP Publishing. 10.1088/1755-1315/272/2/022102
7. Li Y, Li J, Li W, Du H. A state-of-the-art review on magnetorheological elastomer devices. Smart materials and structures. 2014;23(12):123001. 10.1088/0964-1726/23/12/123001
8. Anvar V, Alexey Z, Artem T. Study of application of vibration isolators with quasi-zero stiffness for reducing dynamics loads on the foundation. Procedia Engineering. 2017;176:137-43. 10.1016/j.proeng.2017.02.281
9. Zotov A, Tokarev A. A Dynamic Vibration Absorber with Adjustable Stiffness. Journal of Machinery Manufacture and Reliability. 2023;52(6):525-31. 10.1134/S1052618823050175
10. Tokarev A, Usmanova A, Godovskiy D, editors. Conditions for the Occurrence of Resonant Oscillations of Vertical Booster Pumping Units. IOP Conference Series: Earth and Environmental Science; 2022: IOP Publishing. 10.1088/1755-1315/988/3/032063
11. Tokarev A, Yanbarisova A, Khatmullina R, editors. The Dependence of the Pump Piping Vibration from the Rotor Speed. IOP Conference Series: Earth and Environmental Science; 2021: IOP Publishing. 10.1088/1755-1315/666/4/042065
12. Javanmardi A, Ibrahim Z, Ghaedi K, Benisi Ghadim H, Hanif MU. State-of-the-art review of metallic dampers: testing, development and implementation. Archives of Computational Methods in Engineering. 2020;27(2):455-78. 10.1007/s11831-019-09329-9
13. Palaev A, Krasnikov A. Ultrasonic Treatment of Welded Joint from External, Internal and Two Sides on Reduction of Residual Welding Stresses. International Journal of Engineering Transactions B: Applications. 2024;37:2171-80. 10.5829/ije.2024.37.11b.04
14. Phu V, Canh T, Lieu P. Optimal parameters of dynamic vibration absorber for linear damped rotary system subjected to harmonic excitation. Vietnam Journal of Mechanics. 2020;42(4):385-400. 10.15625/0866-7136/14897
15. Li H, Yao H, Dou J, Jia R, Wei G. Torsional vibration suppression of cantilever beams system with the PNESI. Journal of Low Frequency Noise, Vibration and Active Control. 2024;43(4):1713-31. 10.1177/14613484241272264
16. Medel-Vera C, Ji T. Seismic protection technology for nuclear power plants: a systematic review. Journal of Nuclear Science and Technology. 2015;52(5):607-32. 10.1080/00223131.2014.980347
17. Wang H, Wang B, Hu L, Luo X. Modeling and Experimental Study on the Influence of Support Stiffness on the Vibration of Piston Pump. Journal of Vibration Engineering & Technologies. 2023;11(7):3069-80. 10.1007/s42417-022-00731-5
18. Shammazov I, Sidorkin D, Dzhemilev E, editors. Research of the dependence of the pipeline ends displacement value when cutting out its defective section on the elastic stresses in the pipe body. IOP Conference Series: Earth and Environmental Science; 2022: IOP Publishing. 10.1088/1755-1315/988/2/022077
19. Xu Z-D, Chen Z-H, Huang X-H, Zhou C-Y, Hu Z-W, Yang Q-H, et al. Recent advances in multi-dimensional vibration mitigation materials and devices. Frontiers in Materials. 2019;6:143. 10.3389/fmats.2019.00143
20. Pshenin VV DE, Rozanova LR, Komarovskii MS. Results of analysis of accident rates of gas distribution systems. Problemy

sbora, podgotovki i transporta nefti i nefteproduktov. 2022;4(138):89-101. 10.17122/ntj-oil-2022-4-89-101

21. Zemenkova MY, Chizhevskaya EL, Zemenkov YD. Intelligent monitoring of the condition of hydrocarbon pipeline transport facilities using neural network technologies. Записки Горного института. 2022;258:933-44. 10.31897/PMI.2022.105
22. Perveitalov OG, Nosov VV, Schipachev AM, Alekhin AI. Thermally Activated Crack Growth and Fracture Toughness Evaluation of Pipeline Steels Using Acoustic Emission. Metals. 2023;13(7):1272. 10.3390/met13071272
23. Shlyannikov V, Tumanov A, Zakharov A, Gerasimenko A. Surface crack growth subject to bending and biaxial tension-compression. Fracture and Structural Integrity. 2016;10(35):114-24. 10.3221/IGF-ESIS.35.14
24. Yusuf M, Latief Y, Rarasati AD, Trigunarsyah B, Laksono NB. Fuzzy Bayesian Belief Networks Method on Risk Assessment of EPC Pipeline Project. Civil Engineering Journal (Iran). 2025;11(3):1050-71. 10.28991/CEJ-2025-011-03-013
25. Baghdasaryan M, Hovhannisyn V. Stability assessment of an ore mill electric drive using machine learning. HighTech and Innovation Journal. 2024;5(2):213-30. https://doi.org/10.28991/HIJ-2024-05-02-01
26. Meena A, Parayil P, Suthar V. Simulation and Validation of FRP Base Plate for Electric Compressor. SAE Technical Paper; 2020. Report No.: 0148-7191.
27. Bolshunov AV, Ignatev SA, Gorelik GD, Krikun NS, Vasilev DA, Rakitin IV, et al. Comprehensive studies of the snow-firn layer in the area of the Russian Antarctic Vostok Station. Journal of Mining Institute. 2025.
28. Shojaee Barjoee S, Rodionov V, Vaziri Sereshk AM. Noise climate assessment in ceramic industries (Iran) using acoustic indices and its control solutions. Advances in Environmental Technology. 2025;11(1):91-115. 10.22104/AET.2024.6922.1899
29. Skamyin A, Shklyarskiy Y, Gurevich I. Influence of background voltage distortion on operation of passive harmonic compensation devices. Energies. 2024;17(6):1342. 10.3390/en17061342
30. Shibanov D, Ivanov S. Failure risks of mine excavator associated with its maintenance and repair. Russian Mining Industry. 2024;2:97-102. 10.30686/1609-9192-2024-2-97-102
31. Khamidov O, Shibanov D, Shishkin P, Kolpakov V. Efficiency of excavators application in open pit mines of Uzbekistan. Min Ind J. 2024;5:135-42. 10.30686/1609-9192-2024-5-135-142
32. Shammazov IA BA, Aleksandruk BS,. Algorithmic Models for Determining the Flow Patterns of Oil Pipelines in Gravity Sections. International Journal of Engineering Transactions A: Basics. 2025;38(10):2476-85. 10.5829/ije.2025.38.10a.22
33. Rousou K, Rani MA, Bahari A, Kyprianou A, editors. Exploratory structural modification and nonlinear based analysis of tuned mass dampers. Journal of Physics: Conference Series; 2024: IOP Publishing. 10.1088/1742-6596/2721/1/012016
34. Cao J, Liu B, Wu Q, Liu J. Research on the vibration control of micro-vibration using a novel hybrid isolator. Experimental Techniques. 2024:1-13. 10.1007/s40799-024-00772-3
35. Zang J, Cao R-Q, Zhang Y-W. Steady-state response of a viscoelastic beam with asymmetric elastic supports coupled to a lever-type nonlinear energy sink. Nonlinear Dynamics. 2021;105(2):1327-41. 10.1007/s11071-021-06625-7
36. Chang Y, Li Y, Zhou J, Wang K, Wang Q, Wen G. Compensation strategy for quasi-zero-stiffness vibration isolator under payload mismatch. Acta Mechanica Sinica. 2024;40(10):524033. 10.1007/s10409-024-24033-x
37. Faraj AK, Hussein HAHA. Application of finite element technique: a review study. Iraqi Journal of Chemical and Petroleum Engineering. 2023;24(1):113-24. https://doi.org/10.31699/IJCPE.2023.1.13
38. Fedorova E, Morgunov V, Pupysheva E. Effect of variation of internal diameter along the length of a rotary kiln on material movement. Non-ferrous Metals. 2024;56(1):28-34. 10.17580/nfm.2024.01.05
39. Chigarev A, Kravchuk A, Smalyuk A. ANSYS for engineers: reference manual. Mechanical Engineering. 2004;1.
40. Wang Z, Zhang Q, Zhang K, Hu G. Tunable digital metamaterial for broadband vibration isolation at low frequency. Advanced materials. 2016;28(44):9857-61. 10.1002/adma.201604009
41. Almashhor A, Asiri SA. Development of a new tuned vibration absorber based on one degree-of-freedom of translational motion. Cogent Engineering. 2021;8(1):1976929. 10.1080/23311916.2021.1976929
42. Pshenin V, Menshikov S, Komarovskiy M, editors. Study of static charge accumulation in HDPE gas pipelines. E3S Web of Conferences; 2023: EDP Sciences. 10.1051/e3sconf/202337802001
43. Schipachev A, Fetisov V, Nazyrov A, Donghee L, Khamrakulov A. Study of the pipeline in emergency operation and assessing the magnitude of the gas leak. Energies. 2022;15(14):5294. 10.3390/en15145294

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

اهمیت این تحقیق ناشی از ضرورت ارتقای سیستم‌های جبران ارتعاش موجود برای کاهش بارهای ارتعاشی بر روی واحدهای قدرت واحدهای پمپاژ خط اصلی (MPU) است. الگوریتم‌ها و نتایج محاسباتی وضعیت ارتعاش فضایی سیستم جبران ارتعاش (VCS) برای MPUها تحت رژیم‌های عملیاتی مختلف ارائه شده است. مدل‌های ریاضی و فضایی المان محدود، و همچنین نمونه‌های اولیه دیجیتال همه‌کاره VCS، توسعه داده شده‌اند. نوآوری یافته‌ها در نمونه اولیه دیجیتال توسعه‌یافته MPU نهفته است که قابلیت‌های گسترده‌ای را برای در نظر گرفتن عوامل طراحی-فنی ارائه می‌دهد، رژیم‌های عملیاتی متنوع را در خود جای می‌دهد و دقت و کارایی محاسباتی بالایی را نشان می‌دهد. اهمیت عملی این کار این است که روش شبیه‌سازی توسعه‌یافته مبتنی بر نمونه‌های اولیه دیجیتال MPU VCS، تجزیه و تحلیل اثرات ارتعاشی موتورهای الکتریکی، پمپ‌ها و قاب‌های تحمل بار را امکان‌پذیر می‌سازد. مطالعات محاسباتی جامعی در مورد وضعیت تنش-کرنش قاب VCS انجام شد و فرکانس‌های طبیعی سیستم در حین کار ارزیابی شدند. هدف از این کار، مطالعه رفتار واحد پمپاژ هنگام جایگزینی تکیه‌گاه‌های الاستومری در سیستم جبران با میراگرهای هیدروفیلمی است. یکی از مزایای غیرقابل انکار مدل دیجیتال توسعه‌یافته، پایه و اساس آن برای یک روش‌شناسی جهانی برای ارزیابی اثربخشی جبران‌کننده‌های ارتعاش در MPU VCS است. یک نمونه اولیه دیجیتال منحصر به فرد از MPU با در نظر گرفتن VCS مهندسی‌شده توسعه داده شد. امکان‌سنجی استفاده از یک نمونه اولیه دیجیتال مبتنی بر المان محدود برای ارزیابی بارگذاری ارتعاشی روی MPUها اثبات شده است. آزمایش‌های عددی با استفاده از دوقلوی دیجیتال MPU نشان داد که پیاده‌سازی میراگرهای هیدروفیلمی در VCS، تنش‌های دینامیکی اوج را ۱۸٪ کاهش می‌دهد. مقادیر جابجایی‌های محوری و منتج MPU با استفاده از میراگرهای هیدروفیلمی و الاستومری در VCS به ترتیب ۱۱/۱ میلی‌متر و ۱۰/۰ میلی‌متر اندازه‌گیری شدند.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1238-1253

International Journal of Engineering

Journal Homepage: www.ije.ir

Achieving Enhanced Efficiency and Security in IoT: A Hybrid Encryption Approach with ECC and Blowfish

M. Baghaei Jezehei[a], S. A. Olamaei*[a], A. Broumandnia[b]

[a] *Department of Electrical Engineering, ST.C., Islamic Azad University, Tehran, Iran*
[b] *Department of Computer Engineering, ST.C., Islamic Azad University, Tehran, Iran*

P A P E R I N F O

Paper history:
Received 10 February 2025
Received in revised form 06 June 2025
Accepted 15 July 2025

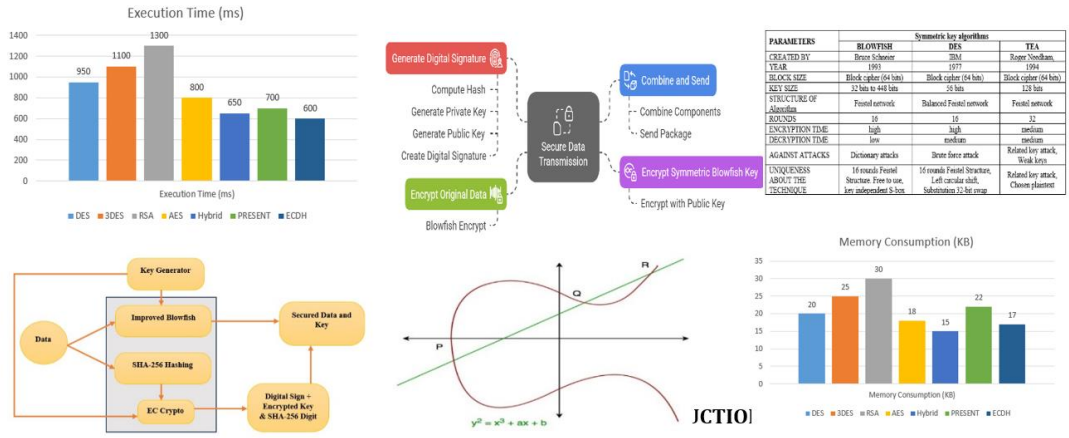
Keywords:
Internet of Things
Hybrid Encryption
Elliptic Curve Cryptography
Blowfish Algorithm
Data Security

A B S T R A C T

The rapid proliferation of the Internet of Things (IoT) has heightened the urgency to address security vulnerabilities inherent in resource-constrained connected devices. Protecting sensitive user information underscores the necessity of securing personal data. Unauthorized access threats require advanced encryption solutions to enhance security and optimize system performance. This study proposes a novel hybrid encryption framework that uniquely combines asymmetric Elliptic Curve Cryptography (ECC) with a symmetric Blowfish algorithm. Unlike conventional methods such as RSA, AES, or standalone ECC, our approach integrates these techniques innovatively to achieve superior efficiency and security. Furthermore, a new adaptive encryption protocol is developed, optimized for data volume and device capabilities. Blowfish efficiently encrypts large data (over 64 KB) while maintaining speed. Additionally, ECC employs smaller key sizes, enhancing security without significant processing overhead. This research used data from various IoT devices like sensors, wearable gadgets, smart home systems, traffic, and industrial tools, tested both in labs and real-world settings. The data included various types of sensitive information, with each dataset averaging over 50 KB. Furthermore, all experiments were implemented on a system with a dual-core Intel processor at 2.5 GHz, ensuring a consistent performance evaluation. Compared to traditional cryptographic solutions such as RSA, AES, and standalone ECC implementations, our approach demonstrates a significant improvement in execution time—greater than 15%—and substantially enhances overall efficiency. This research shows that our hybrid approach improves IoT security while saving processing power and energy, helping future device performance.

doi: *10.5829/ije.2026.39.05b.16*

Graphical Abstract



*Corresponding Author Institutional Email: Olamaei@iau.ac.ir (S. A. Olamei)

Please cite this article as follows: Baghaei Jezehei M, Olamei SA, Broumandnia A, Achieving Enhanced Efficiency and Security in IoT: A Hybrid Encryption Approach with ECC and Blowfish. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1238-53.

1. INTRODUCTION

The Internet of Things (IoT) connects many devices that collect and share data, emphasizing the need for strong security measures. Due to resource constraints, traditional cryptographic methods often fall short. This paper reviews lightweight protocols, including symmetric and asymmetric encryption, and block and stream ciphers, designed for IoT. While IoT enables real-time data sharing in sectors like smart cities, agriculture, and healthcare, it also introduces security challenges from cyberattacks.

The integration of networks and IoT enhances remote control but increases security risks, especially without proper safeguards. Effective authentication is vital; however, passwords are often inadequate for low-resource devices. A new lightweight authentication scheme is proposed to counter threats like replay and man-in-the-middle attacks. The paper also discusses challenges such as high costs, data management, and bandwidth overload affecting QoS. Its main contribution is a novel, resource-efficient authentication method tailored for low-resource environments, setting it apart from existing solutions (1, 2).

Smart healthcare systems that utilize ICT and IoT analyze real-time data to facilitate proactive care, particularly in the prediction of heart diseases. While these systems are highly effective, IoT devices are vulnerable to security threats often caused by poor configurations (3).

In wireless sensor network security, lightweight encryption algorithms were tested on the Atmel ATmega 128L microcontroller and Riverbed Modeler. The results showed that RECTANGLE was the fastest and most energy-efficient, while Camellia offered higher security. These findings highlight the importance of SPN-based algorithms for protecting sensor networks, especially in military and healthcare applications (4).

Vehicular Ad Hoc Networks (VANETs) facilitate communication between vehicles without fixed infrastructure. The GHRP protocol enhances security with source keys validated via one-way hashing, protecting against black hole attacks. Simulations show GHRP outperforms SAODV in packet delivery, with lower overhead and latency. A new protocol, SGHRP, further improves security and routing, demonstrating VANET insights can advance IoT security in smart homes and healthcare (5).

Networks on Chip (NoCs) boost inter-core communication. The ScRN algorithm manages traffic effectively, reducing packet latency by 38%, increasing throughput by 20%, and lowering energy consumption by 10%. As chip complexity rises, SoC architectures benefit from ScRN's improved latency and efficiency over older methods (6).

Securing data in IoT and cloud storage is challenging. Using NoC with Attribute-Based Encryption (ABE) enhances secure communication and access control through Blinding Value Encrypting Keys (BEKs) linked to user attributes. The system combines symmetric encryption, a polynomial tree, and authentication to ensure confidentiality and resist attacks like collusion and DoS. A lightweight authentication reduces resource depletion. Experiments show high efficiency and low overhead, with potential for improved privacy via multi-authority ABE (7).

IoT services are classified into low-level, data-oriented, and application-based categories, each with unique security risks requiring tailored solutions. All IoT devices must address the 'Big Three' security issues: confidentiality, integrity, and authentication, along with non-repudiation. Table ١ outlines these challenges.

This study highlights rising security and privacy concerns in IoT. It advocates for lightweight cryptography, analyzing algorithms like AES, DES, 3DES, and Blowfish. Tests on a Raspberry Pi 3B+ show Blowfish offers the best throughput despite DES and 3DES being faster. The research emphasizes the importance of secure IoT communication and calls for further work on optimizing power and memory for practical applications (8).

Present study emphasizes the importance of selecting appropriate encryption methods for resource-limited IoT devices. It explores neuromorphic computing and an adaptive routing protocol that reduces delays and power consumption. Additionally, a data-oriented RPL algorithm enhanced with Binary Gray Wolf Optimization and fuzzy logic improves network longevity and security. A comprehensive framework integrating cloud computing, blockchain, and advanced encryption techniques demonstrates superior data security and efficiency, addressing IoT challenges holistically (9-11).

Furthermore, this study compares the performance of RSA and ECC algorithms in IoT devices. Results show that ECC is more efficient in resource-constrained environments, with advantages in memory usage, energy consumption, and execution time, while RSA performs better in signature verification and encryption (12).

TABLE 1. Threats in IoT

Threats	Minimum Level	Data Related	App Level
Replication of Nodes	✓		
Tag Cloning		✓	
Eavesdropping		✓	
Packet Injection	✓	✓	✓
DoS Attack	✓		✓
Camouflage	✓		

To enhance IoT security without compromising performance, a novel encryption approach combines two Blowfish algorithms with ECC, resulting in reduced memory usage, faster processing, and lower complexity. This scheme, validated through literature review and testing, integrates Improved Elliptic Curve Cryptography (IECC) with deep learning (LSTM) for malware detection. It enhances security by accurately identifying threats and ensuring secure data transmission, achieving 95% accuracy and effectively addressing IoT vulnerabilities.

2. RELATED WORKS

Unlike previous approaches that focus solely on traditional encryption methods, the proposed hybrid approach introduces a novel combination of ECC and Blowfish to significantly improve both security and efficiency in resource-constrained IoT environments.

Mahandru and Kaur (13) introduced a new method for securing data in cloud computing utilizes various cryptographic techniques to tackle existing security challenges. Despite the advantages of online data storage, ensuring data security is crucial. This approach features a hybrid encryption scheme that combines the Blowfish algorithm and the AES with the Elliptic Curve Diffie-Hellman (ECDH) for secure key exchange. It addresses security concerns in public, private, and hybrid cloud models, enhancing speed and efficiency for better data security and resource management. However, its application in IoT environments, where resource limitations and scalability are key factors, requires further investigation.

Bhagwatrao and Lakshmanan (14) reported that to address these security challenges, a new solution aims to enhance remote communication by integrating Deep Q-Network (DQN) and Graph Convolutional Network (GCN) blocks. DQN is utilized to determine the optimal actions for maximizing rewards, while GCN analyzes graph data to bridge various encryption and hashing methods. This system, named RHSSAO, demonstrates significant improvements in terms of adaptability, speed, and energy efficiency. However, additional comparisons with more contemporary IoT-specific solutions are necessary to validate its utility in dynamic and heterogeneous networks.

The objective of the study conducted by Prashant et al. (15) is to ensure the secure transmission of data through AES encryption. Safely transferring information across different working devices has become quite an intimidating challenge. In this context, algorithms such as DES, RSA (Rivest-Shamir-Adleman) and AES are crucial. This paper introduces a mechanism for secure data transfer, whereby security is enhanced by AES encryption. Overall, RHSSAO represents a significant advancement in remote communication, contributing to the secure transmission of data within the broader IoT security framework. More attention should be given to the challenges of data transfer in IoT environments, considering the limited resources and dynamic nature of these networks. Although this mechanism strengthens data security for IoT devices, it fails to address issues related to limited memory and processing power common in IoT environments.

Raghad et al. (16) proposed the security of data transmission in the digital age is crucial, yet it faces risks such as hacking. Effective encryption methods, like symmetric encryption, are essential for protection. The Blowfish algorithm is a widely used symmetric cipher but has limitations, including complex calculations and fixed S-Boxes, complicating its use with complex data. This paper reviews improvements to the Blowfish algorithm, focusing on execution time and security based on around 40 studies. Key enhancements include using a 128-bit block cipher, reducing S-Boxes (Substitution Boxes), and implementing parallel processing. While some improvements increased speed and security, they also raised memory usage. The research emphasizes the need to optimize symmetric encryption algorithms for resource-constrained IoT devices, aiming to reduce computational complexity while maintaining security.

Kapalova et al. (17) proposed a lightweight algorithm for better protection in computing environments. The algorithm employs the Feistel network and permutation architecture in a 16-byte block-encryption scheme to add complexity to the mechanism. While this algorithm shows promise in achieving faster encryption, its dependency on private keys introduces vulnerabilities that are exacerbated in IoT systems requiring dynamic key management.

Trik and Boukani (18) employed a range of encryption techniques to enhance the security measures for storage systems in IoT and fog computing. The research team developed an asymmetric system that uses encryption founded on the AES algorithm in conjunction with asymmetric key transfer methods to ensure security and confidentiality in data exchange. These findings underscore the importance of balancing security and efficiency, especially given the latency-sensitive nature of IoT environments.

Obaidat et al. (19) provided a 2-stage encryption solution to assure the protection of stored information, which entails dividing the primary data into two segments, with each segment being encrypted using a shared key generated according to the chaos theory model. While this method demonstrates reduced processing times, its scalability for real-time IoT applications remains unclear.

Agbelusi and Matthew (20) also provided an in-depth examination of various symmetric encryption algorithms, specifically focusing on their performance in terms of run time and memory utilization. For IoT devices, where resource availability is constrained,

algorithms like Blowfish and DES should be further evaluated under IoT-specific scenarios.

The PRESENT encryption algorithm is a lightweight block cipher designed for securing data in low-resource systems and Internet of Things (IoT) devices. It employs a substitution-permutation network (SPN) structure and utilizes short keys, making it highly suitable for devices with limited computational power. Known for being fast, efficient, and resistant to certain cryptographic attacks, PRESENT's design aims to provide robust security in resource-constrained environments. However, as examined by Lo et al. (21), power analysis attacks— a form of side-channel attack— reveal vulnerabilities in the algorithm. The study indicates that specific components, such as the "Add Round Key" and "S-Box" modules, are susceptible to correlation analysis, which can potentially reveal some key bits. This highlights that, despite its efficiency and low resource requirements, the security of PRESENT against side-channel attacks is limited, and additional precautions may be necessary to ensure complete security in practical applications.

Khezri et al. (22), Khalafi and Boob (23) analyzed the leading encryption solutions relevant to Internet of Things (IoT) systems and distributed computing frameworks. These studies emphasize the pressing requirement for algorithms and solutions that not only ensure security but also maintain service quality parameters, thereby reducing the adverse effects of encryption methods on services offered within IoT environments. Further empirical research is necessary to validate the efficacy of these solutions in diverse IoT applications.

In contrast to the aforementioned methods, our proposed hybrid encryption framework combines ECC and Blowfish in an integrated manner that specifically targets resource constraints and operational efficiency in IoT environments, representing a significant step forward in providing lightweight yet highly secure encryption.

Khezri et al. (24) proposed a novel KGRD system for secure key management in IoT using MQTT, leveraging a TPM for cryptographic operations to enhance key distribution and renewal. It combines asymmetric and symmetric cryptography to improve security and efficiency. Unlike blockchain-based solutions, our method uses hybrid encryption, balancing security and resource consumption without requiring high computational power.

Ding et al. (25) also proposed a novel approach for real-time personalized route selection in urban traffic management. The method addresses issues caused by a rapid increase in vehicle ownership and congestion resulting from events and accidents. Limited road network capacity worsens travel conditions, leading to wasted time and resources. Since expanding infrastructure is difficult, the proposed solution employs a Nash equilibrium established through mutual information swapping and self-adaptive learning. Simulation results demonstrate that this algorithm outperforms existing methods in dynamic, personalized route optimization.

3. BLOWFISH AND FEISTEL NETWORK

In this section, we examine the Blowfish algorithm, a widely used symmetric encryption method based on a Feistel network. Developed by Bruce Schneier in 1993, Blowfish features a variable key length ranging from 32 to 448 bits, providing flexibility to balance security and performance. It is known for its high speed and ease of implementation, supporting a block size of 64 bits and making it suitable for a variety of applications. However, like many symmetric algorithms, Blowfish can be vulnerable to certain attacks, such as dictionary attacks, particularly if weak keys are used or key management practices are inadequate.

To facilitate a comprehensive comparison, Table 2 evaluates three prominent symmetric key algorithms: Blowfish, DES, and TEA. This analysis considers criteria such as year of development, block size, key length, internal structure, encryption and decryption times, and potential attack vectors. This comparison highlights the strengths and weaknesses of each algorithm, aiding in the selection of appropriate cryptographic solutions for IoT applications and beyond.

3. 1. Key Components

1. Block Size: 64 bits.
2. Key Size: Variable (32 to 448 bits).
3. Rounds: 16 rounds of processing.

3. 2. Feistel Network Structure The Blowfish cipher operates using a Feistel network design, which splits the input data into two halves and processes them through multiple iterations.

3. 2. 1. Block Division The plaintext P (64 bits) is divided into two halves:

$$P=L_0||R_0 \quad (1)$$

Where:

- L_0 (left half): 32 bits
- R_0 (right half): 32 bits

3. 2. 2. Round Process The encryption process consists of 16 rounds. For each round I (from 0 to 15), the transformation is defined as:

$$L_{i+1}=R_i \quad (2)$$

$$R_{i+1}=L_i \oplus F(R_i) \oplus K[i] \quad (3)$$

Where F is a function applied to the right half Ri, and K[i] is the sub-key for round i.

TABLE 2. Analysis of Symmetric Algorithms

PARAMETERS	Symmertic key algorithms		
	BLOWFISH	**DES**	**TEA**
CREATED BY	Bruce Schneier	IBM	Roger Needham
YEAR	1993	1977	1994
BLOCK SIZE	Block cipher (64 bits)	Block cipher (64 bits)	Block cipher (64 bits)
KEY SIZE	32 bits to 448 bits	56 bits	128 bits
STRUCTURE OF ALGORITHM	Feistel network	Balanced Feistel network	Feistel network
ROUNDS	16	16	32
ENCRYPTION TIME	High	High	Medium
DECRYPTION TIME	Low	Medium	Medium
AGAINST ATTACKS	Dictionary attacks	Brute force attack	Related key attack, Weak keys
UNIQUENESS ABOUT THE TECHNIQUE	16 rounds Feistel Structure. Free to use, key independent S-box	16 rounds Feistel Structure, Left circular shift, Substitution 32-bit swap	Related key attack, Chosen plaintext

3. 2. 3. Function F Function F is defined on a 32-bit input and consists of several steps:

3. 2. 3. 1. Input Splitting The input R is divided into four 8-bit segments:

$$R= [b_1, b_2, b_3, b_4] \quad (4)$$

3. 2. 3. 2. S-Boxes Lookup Each 8-bit part bi indexes one of four S-boxes

T_1, T_2, T_3, T_4: (5)

$$\begin{aligned} y_1&=T_1[b_1]\\ y_2&=T_2[b_2]\\ y_3&=T_3[b_3]\\ y_4&=T_4[b_4] \end{aligned} \quad (6)$$

3. 2. 3. 3. Combining Outputs As a precaution, it is necessary to analyze whether the modifications made to function F enhance speed without compromising security against analytical attacks, such as differential or linear cryptanalysis. This analysis was conducted during the redesign of the function to ensure it is less vulnerable to such attacks.

$$F(R)=(y_1+y_2) \bmod 256 \oplus y_3+y_4 \quad (7)$$

Where:

- The addition y_1+y_2 is taken modulo 256.
- The result of the modulo operation is XORed with y_3
- Finally, y_4 is added to the total.

3. 3. Subkey Generation To perform encryption, the Blowfish algorithm uses a key schedule to generate 18 subkeys and initializes four S-boxes.

3. 3. 1. P-array Initialization The P-array is initialized with specific constants:

$$\begin{aligned} &P[0] =0x243F6A88\\ &P[1] =0x85A308D3\\ &P[2] =0x13198A2E\\ &P[3] =0x03707344\\ &\vdots\\ &P[17] =0xA409B34C \end{aligned} \quad (8)$$

3. 3. 2. S-box Initialization Each of the four S-boxes is initialized with fixed values. Example Initialization:

$$\begin{aligned} &S_0[0] =0xD1310BA6\\ &S_0[1] =0x98DFB5AC\\ &S_0[2] =0x2FFD72DB\\ &S_0[3] =0xD01ADFB7\\ &\vdots\\ &S0[255] =0xCA62C1D6 \end{aligned} \quad (9)$$

3. 3. 3. Key Scheduling Steps

1. XOR Key into P-array: Each entry of the P-array is XORed with successive bytes of the key.

Suppose the key is K with T bytes (T=length of the key in bytes):

$$P[i]=P[i] \oplus Key\ [i \bmod T]\ (i \text{ from } 0 \text{ to } 17) \quad (10)$$

2. Encrypt Zero Input: A block of zero is used to fill the P-array:

temp=0x00000000||0x00000000

For each I (from 0 to 17), calculate:

[i]=Blowfish Encrypt (temp) (11)

3. Fill S-boxes: Use the outputs from step 2 to populate the S-boxes sequentially.
   Filling an S-box Entry Example: Using outputs sequentially:

S[j]=Blowfish Encrypt(temp) (j fills from the outputs)

3. 4. Encryption Process The encryption process can be defined with the following steps:

1. Start with halves L_0 and R_0 derived from plaintext P.
2. Repeat the function transformation for 16 rounds:

$$R_{i+1}=L_i \oplus F(R_i) \oplus K[i] \quad (12)$$

3. Upon completion of all rounds, the final output is constructed by swapping the halves:

$$\text{Ciphertext}=R_{16}||L_{16} \quad (13)$$

3. 5. Decryption Process Decryption mirrors the encryption process but applies the keys in reverse order:

1. Split the ciphertext C:

$$C=L_{16}||R_{16} \quad (14)$$

2. Perform rounds in reverse order ii from 15 down to 0:

$$R_i=L_{i+1} \quad (15)$$

$$L_i=R_i+1 \oplus F(L_{i+1}) \oplus K[15-i] \quad (16)$$

3. The output after processing through all rounds is:

$$\text{Plaintext}=R_0||L_0 \quad (17)$$

3. 6. Security Considerations

1. **Diffusion:** Each round significantly spreads the influence of the input bits over the output bits. A change in one input bit will result in a significant change in output bits.
2. **Confusion:** Due to the S-boxes and the function FF, the relationship between the key and ciphertext becomes highly complex.
3. **Resistance to Attacks:** Blowfish ensures strong resistance against various attack types, including differential and linear cryptanalysis.

The modifications to function F have been carefully analyzed to ensure that while speed is improved, security against differential and linear cryptanalysis is not compromised. The relationship between input changes and output results has been preserved to maintain the robustness of the encryption process.

3. 7. Summary of Key Mathematical Operations in Blowfish Here's a summary of the key mathematical operations used in the Blowfish algorithm:

- Round Transformation:
- The transformation steps in each round are defined as follows:

$$L_{i+1}=R_i \quad (18)$$

$$R_{i+1}=L_i \oplus F(R_i) \oplus K[i] \quad (19)$$

- Function F:

$$F(R)=(T_1[b_1]+T_2[b_2]) \bmod 256 \oplus T_3[b_3]+T_4[b_4] \quad (20)$$

In this equation:

- T1, T2, T3, and T4 are the four S-boxes that correspond to the 8-bit input segments R.
- The mod 256 operation is applied because the sum of T1[b1] and T2[b2] may exceed the range of 0 to 255, necessitating assurance that the result remains within the 8-bit range.
- Since y3 and y4 are directly sourced from the S-boxes and are naturally within the range of 0 to 255, they do not require the mod 256 operation. This design helps ensure that the values are used correctly while maintaining computational complexity.
- This use of mod 256 does not alter the standard functionality of Blowfish, as it continues to preserve the distribution pattern and information processing, ensuring the algorithm's security is not compromised.
- Key Scheduling:

$$P[i]=P[i] \oplus \text{Key}[i \bmod T] \quad (21)$$

Here:

- P[i] is used as an input for generating subkeys.
- Key is the main key that assists in generating and expanding the keys.

Blowfish Security: The security of the Blowfish algorithm arises from its complex key schedule and Feistel structure. This design makes it resistant to brute-force attacks and cryptanalysis. Key expansion ensures that changes in the key produce significant changes in the key array (K-array) and S-boxes (substitution-boxes), while the Feistel network and round function ensure that each bit of the plaintext influences many bits of the ciphertext.

4. NOVELTY AND COMPARISON WITH RELATED METHODS

4. 1. Introduction to the Proposed Method In this study, we introduce a novel hybrid encryption framework that combines Elliptic Curve Cryptography (ECC) with the Blowfish algorithm to address the dual challenges of security and efficiency in IoT environments. The primary aim of this approach is to leverage the strengths of both asymmetric and symmetric cryptography, ensuring a high level of security while maintaining low computational overhead suitable for resource-

constrained devices. ECC facilitates secure and efficient key exchange with minimal energy consumption, while Blowfish provides fast symmetric encryption for data transmission. By integrating these techniques, our framework offers a lightweight yet robust encryption solution specifically tailored for the dynamic and resource-limited context of IoT networks.

4. 2. Key Innovations and Contributions This study presents several significant innovations that enhance the efficiency and security of cryptographic solutions for IoT environments. The core advancement is an optimized Blowfish algorithm featuring a modified F function that enables parallel processing of addition operations, reducing processing time by up to sixteen times without compromising security. By integrating ECC with this improved Blowfish, the framework provides a high-security, resource-efficient encryption scheme suitable for constrained IoT devices. The design emphasizes minimizing energy consumption and memory usage, making it lightweight yet robust. Overall, the work advances the state of the art in hybrid cryptography for IoT, offering practical algorithmic improvements and a scalable solution to meet the evolving security demands of modern networks.

4. 3. Comparison with Existing Methods In this section, we compare our proposed hybrid encryption scheme with traditional cryptographic techniques such as RSA, AES, standard ECC, and standalone Blowfish, focusing on four critical aspects: security level, speed and computational efficiency, resource consumption (memory and energy), and suitability for IoT environments.

4. 3. 1. Security Level RSA and standard ECC provide robust security based on difficult mathematical problems. However, RSA generally requires larger key sizes, which can increase computational load. AES is known for high security in symmetric encryption but does not inherently facilitate secure key exchange. Our hybrid approach leverages ECC's strong security for lightweight key exchange and Blowfish's efficient symmetric encryption, offering a comprehensive security framework optimized for IoT devices.

4. 3. 2. Speed and Computational Efficiency Although RSA and AES are often faster than standard ECC in processing large data, RSA's encryption and decryption demands are still significant, especially for resource-limited devices. ECC achieves comparable security with smaller keys, but its computations remain relatively intensive. Standalone Blowfish, renowned for its speed, lacks secure key exchange capabilities. Our method enhances Blowfish with an optimized F function that enables parallel execution of addition operations, resulting in encryption and decryption speeds up to times faster than traditional implementations. This makes the scheme highly suitable for real-time IoT applications.

4. 3. 3. Resource Consumption (Memory and Energy) RSA's large key sizes and ECC's complex mathematical operations can impose heavy demands on device memory and energy, limiting practicality in IoT environments. AES, while more efficient, still consumes considerable resources. Blowfish, known for its lightweight design, excels in fast processing but does not address secure key management alone. By combining ECC for secure, low-energy key exchange with an optimized Blowfish for fast data encryption, our approach significantly reduces both memory usage and power consumption, making it ideal for constrained devices.

4. 3. 4. Suitability for IoT Environments Traditional encryption algorithms like RSA are often too resource-intensive for many IoT devices. AES is more practical but may still pose challenges in highly constrained scenarios. ECC offers a good compromise but can be demanding for extremely limited hardware. Blowfish's speed is advantageous, but without a secure key exchange method, it's insufficient alone. Our hybrid scheme provides a balanced solution: ECC enables secure and efficient key distribution with minimal energy, while the optimized Blowfish provides fast, secure data encryption. Overall, this makes our method highly suitable for IoT networks, where security, efficiency, and resource economy are paramount.

4. ۴. Summary of the Comparative Analysis Our proposed hybrid encryption scheme stands out among existing solutions due to its unique combination of ECC and an optimized Blowfish algorithm, specifically designed for the constraints of IoT devices. Unlike traditional methods that often sacrifice speed for security or vice versa, our approach achieves a balanced enhancement in both areas. It significantly reduces processing time and energy consumption while maintaining robust security. This makes it the most effective and practical solution for resource-limited IoT environments, setting a new benchmark in the field of lightweight cryptography.

5. METHODS

This involves establishing a comprehensive security structure that incorporates various safeguarding methods related to encryption. In this context, symmetric and asymmetric encryption techniques are integrated with an advanced signature mechanism to ensure the integrity of information using the SHA-256 (Secure Hash Algorithm 256-bit) hashing algorithm. All these factors lead to methods that seem to provide strong security for

computing environments. A critical parameter in cryptography is the speed of calculations, which directly affects network performance. One of the weaknesses of the asymmetric algorithm is that it takes more time to generate a key than the symmetric algorithm. Therefore, they are slower than symmetric key techniques like Blowfish. Thus, encrypting the original message for transmission, the symmetric key approach, due to its rapid processing capabilities, is considered the optimal choice. The main task of the Blowfish algorithm is to increase security in communication. Figure 1 illustrates the protection of network data using the SHA-256 hashing algorithm. Then, by using the elliptic curve encryption algorithm, we can improve security and create a digital signature as well as generate a private key in the network. Finally, to improve the two main parameters, i.e., speed and security, we use a symmetric algorithm and an advanced algorithm to encrypt the primary data.

5. 1. Digital Signature With the proposed solution, the hash function is combined with the digital signature to protect the integrity of the data. If an attacker attempts to access the data transmitted over the network using a fake digital signature, as shown in the literature (24, 25), they will not gain authorized access to the message according to the creation structure.

Thus, this digital signature is instrumental in fortifying the IoT environment. The various stages focused on maintaining data integrity through hash codes and digital signatures are illustrated below in the flowchart of the proposed solution, as presented in Figure 2.

A user should utilize Figure 2 to protect information from contamination and alteration.

5. 2. Data Encryption Module Though the provision for encrypted data is prioritized in production environments, the core aim is to realize efficient, secure encryption using a low execution time. The encryption of information on the network adopts an improved variant of the Blowfish algorithm to achieve this goal. The Blowfish algorithm supports key lengths starting from 32 bits and can accommodate key sizes up to 448 bits. However, more practically, it effectively operates with key lengths from 32 bits to 384 bits for most applications. This inconsistency arises because while the theoretical maximum is 448 bits, performance and efficiency considerations limit practical usage to 384 bits.

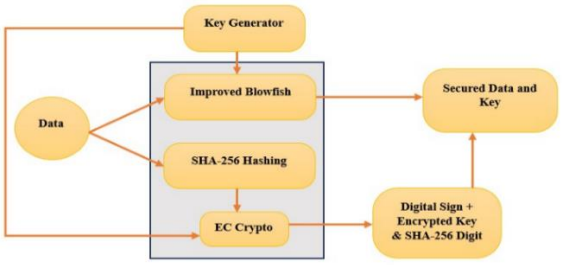


Figure 1. Structure of The Proposed Method

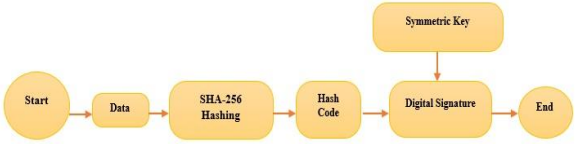


Figure 2. Digital Signature Flowchart

Furthermore, the generation of subkeys necessitates handling several keys within this range, which are created before any encryption or decryption of data. Blowfish distinguishes itself from other algorithms with higher speed, better security, and lower memory consumption. It uses four input boxes, each containing 256 keys with a maximum length of 32 bits. The first byte of the 32-bit input is fed into the first S-box, while the second byte accesses the second S-box, with subsequent inputs selecting S-boxes accordingly. The decryption process closely mirrors the encryption routine but applies subkeys in reverse. Each round revolves around the F module, which combines addition and shifting functions through modular arithmetic (26, 27). This paper aims to enhance execution speed by modifying the F module of the Blowfish algorithm. To optimize the algorithm's time complexity, the value of F is calculated using the standard Blowfish algorithm as described in Equations 22 and 23.

$$G(XL)=(X+Y) \bmod 2^{32} \quad (22)$$

$$F(XL)=G(XL) \oplus Z \quad (23)$$

Smaller values of F can be obtained with minimal integrity degradation, as shown in Equations 24 and 25:

$$Z=(U+V) \bmod 2^{32} \quad (24)$$

$$F'(XL)=G(XL) \oplus Z \quad (25)$$

Finally, the application of this modification allows the simultaneous execution of two addition operations represented as follows: The process of parallelization streamlines two addition operations:

$$G(X_L)=(X+Y) \bmod 2^{32} \text{ and } A=(U+V) \bmod 2^{32} \quad (26)$$

These operations run simultaneously, improving efficiency by consolidating tasks into one complete operation, resulting in a 16-fold reduction in execution time for encryption and decryption rounds. However, the algorithm's security relies on the strength of its keys. The pseudo-code for the revised Blowfish algorithm uses an adapted F function. The 64-bit input is divided into two 32-bit segments: left (L) and right (R). The left segment is processed with an XOR operation using an initial value P, generating an output for the F function. This output is then XORed again. The revised F function optimizes encryption, and decryption reverses these steps with keys applied in reverse order.

5. 3. Secure Techniques for Symmetric Keys In this section, the paper discusses the use of asymmetric encryption based on elliptic curves to overcome challenges related to transmitting private keys in the Blowfish algorithm. ECC (Elliptic Curve Cryptography) leverages the mathematical complexity of elliptic curves over finite groups to facilitate secure key distribution (28, 29). This approach provides high security with a 164-bit key and is more efficient than previous methods due to lower power consumption. It offers the highest level of confidentiality among asymmetric encryption techniques. Notably, by generating a private key through elliptic curve-based encryption and combining it with the Blowfish algorithm, both energy and memory usage are minimized without compromising security, which remains comparable to a 164-bit key, but with better overall performance.

Figure 3 illustrates an elliptic curve based on the standard form of the fundamental Equation 27:

$$y^2= x^3+ax+b \quad (27)$$

The plot clearly displays the x and y axes alongside the curve, enabling precise identification and analysis of points on it. The standard form of this elliptic curve is widely used in cryptography, including key exchange protocols, digital signatures, and encryption algorithms. A key point to note is that including the xy term in Equation 28:

$$y^2+xy-(x^3+ax+b) = 0 \quad (28)$$

indicates a more general and flexible form of elliptic curve. This form allows for broader mathematical analysis and transformations in cryptographic applications. Equation 28 can be converted into the standard form of Equation 27 through coordinate transformations. However, for simplicity and common cryptographic use, the standard form $y^2=x^3+ax+b$ is most frequently employed.

It offers the highest level of confidentiality among asymmetric encryption methods, while consuming little energy and requiring minimal memory resources. In summary, a significant feature of this encryption technique is its operation within finite fields. The logarithm of a discrete elliptic curve problem is known to be an NP-hard (Non-deterministic Polynomial-time hard) problem, and elliptic curves are widely used in cryptography based on this. The equation in Figure 3 describes an elliptic curve (30).

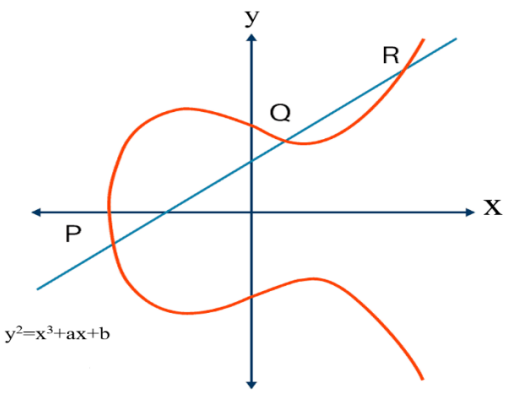


Figure 3. Elliptic Curve

This is where the essential features of elliptic curves come into play. Specifically, when we know two points on the curve, we can establish a rule for finding a third point by adding the two given points. The finite group formed by the points associated with an elliptic curve follows a specific addition rule. However, adding two points is not uniquely defined since they do not intersect at a single point. When defining an elliptic curve, we incorporate this distinguished point into the addition process. O is not an actual point on the curve; however, it possesses the characteristics of one, meaning that adding any point to O yields the original point. Within the framework of an elliptic curve, the total number of points is explicitly defined, taking into account both finite points and those regarded as infinite.

5. 4. Fixed Curve Selection As stated in Equation 27, elliptic curve cryptography (ECC) operates over a finite field GF(p), where p is a large prime number that defines the field. This allows cryptographic operations to be performed on a set of integers less than p.

In this context, m refers to the bit length of the key. Specifically, m indicates the number of bits used to represent the key in binary form. For example, a 256-bit key means m=256. This parameter is crucial because it directly influences the security level of the system and the computational effort required for cryptographic operations.

To ensure the mathematical validity of the elliptic curve, the condition of Equation 2٩ must hold:

$$4a^3 + 27b^2 \neq 0 \quad (29)$$

This guarantees that the curve is *nonsingular*, meaning it has no self-intersections or singular points, which could compromise security.

For cryptographic applications, standardized curves are selected following guidelines from organizations such as NIST or SECG. These curves have been thoroughly analyzed for vulnerabilities and are widely used due to their proven security.

In the process, a random integer k is chosen within the range: $1\leq k<p$ Selecting k uniformly at random within this interval ensures unpredictability, which is critical for security. The public key is then generated using:

$$Q=kG \quad (30)$$

Where G is a fixed point on the curve with a large prime order, ensuring the discrete logarithm problem remains hard and providing security against various attacks. The value m, representing the key length in bits, directly affects the effort needed to compute k from Q=kG. Based

on optimized algorithms, the number of operations required to compute kG is approximately

$$\text{Number of operations} \approx 3 * \left(\frac{m}{3}\right) = m \quad (31)$$

The phrase 'about 3(m/3) operations' indicates that the number of operations needed to compute kG is approximately equal to the key length m. This estimate is derived from optimized algorithms, which approximate the required computational effort as proportional to the size of the key.

This balance allows ECC to deliver high security with relatively small key sizes compared to other methods like RSA, while maintaining computational efficiency.

5. 5. ECC Implementation in Encryption Processes

Consider a scenario involving four users—Anna, Ben, Clara, and Daniel—who have mutually selected a non-confidential elliptic curve and a specific fixed point on that curve, referred to as G. To send a message to Ben, Anna must first acquire Ben's public key. Once she obtains it, she can use Ben's public key to encrypt the message and send the resulting ciphertext to him. In the ECC algorithm, Ben's public key corresponds to a curve point Q on the agreed-upon elliptic curve shared by Anna and Ben. Additionally, Ben has selected a secret integer H, which enables him to derive his public key. For securing the message, Anna begins by creating a random integer k. To ensure that k is generated securely, it is crucial that k be randomly selected from a valid range (between 1 and the order of the elliptic curve G) to prevent vulnerabilities associated with weak values of k. This validation ensures that k does not lead to predictable outcomes that could be exploited by an attacker. Following this, she identifies the poin T on the elliptic curve. Anna then computes the ciphertext, where N denotes the plaintext. Finally, she forwards the pair (T, C) to Ben as the encrypted payload.

$$Q=H\cdot G \quad (32)$$

$$T=kG \quad (33)$$

$$C=N+k\cdot Q \quad (34)$$

$$V=k.Q \quad (35)$$

The secret key H allows him to retrieve the original message corresponding to the ciphertext C. He computes V and finds it concerning C. Anna first performs the multiplication by H and then applies the result to the decryption process. Likewise, Ben can determine the same value for other reasons.

$$N=C-V=C-(kQ+N-kHG) \quad (36)$$

$$HAj=H\cdot Aj\cdot G=Aj\cdot H\cdot G=Aj\cdot Q \quad (37)$$

A strong assumption underpins the security of the method, the challenge of calculating k in relation to G and kG is significant. This is because the operations involved in this method are performed within the finite field. Thus, this method can leverage efficient algorithms for scaled multiplication to enhance the efficiency of elliptic curve-based encryption.

Therefore, this technique generally follows these steps:

- **Phase 1:** Before applying encryption, the original content undergoes hash code generation using the SHA-256 hashing algorithm.
- **Phase 2**: The resulting hash output is encoded and stored alongside the digital certificate and private encryption key.
- **Phase 3**: The private key of the Blowfish algorithm is encrypted using elliptic curve encryption.
- **Phase 4**: Before the data is transmitted, the original information is encrypted using a symmetric method, specifically the Blowfish algorithm, which employs the XOR function. Subsequently, the transmission data is combined with this encrypted data string before being sent out.
- **Phase 5**: The receiving side performs the reverse process. Here, decryption is executed, and the received data can be accessed using the private key.
- **Phase 6**: The secret key of the Blowfish algorithm is employed to recover the original data. After this, analysis and authentication take place using the hashing function (i.e., the digital signature).

5. 6. Pseudo Code for Hybrid Encryption Algorithm

Input:

- Original Data (M)
- Symmetric Blowfish Key (K)
- Elliptic Curve (EC)
- Generator Point (G)

Output:

- Encrypted Data (C)
- Digital Signature (DS)
- Encrypted Key (encrypted Key)

In Figure 4, the Secure Data Transmission Process illustrates the overall workflow involved.

Step-by-Step Explanations of the Pseudo Code:

Inputs:

- **Original Data (M):** The data intended for encryption.
- **Symmetric Blowfish Key (K):** The key used for encrypting and decrypting the data with the Blowfish algorithm.
- **Elliptic Curve (EC):** The mathematical structure used for generating public and private keys in asymmetric encryption.
- **Generator Point (G):** A fixed point on the elliptic curve used for key generation.

Outputs:

- **Encrypted Data (C):** The data obtained after encryption.

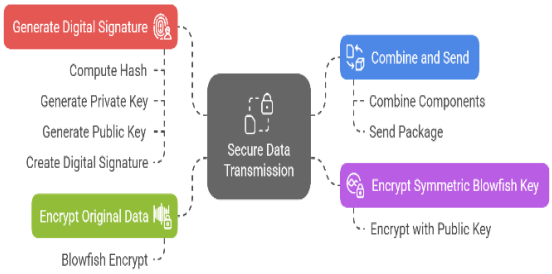


Figure 4. Secure Data Transmission Process

- **Digital Signature (DS):** The signature generated to ensure data integrity and authenticity.
- **Encrypted Key (encrypted Key):** The symmetric key that is securely transmitted.

Detailed Steps:

Step 1: Generate Digital Signature

- **1.1:** The hash value (H) of the original data (M) is computed using the SHA-256 algorithm, creating a unique representation of the original data.
- **1.2:** A private key for the elliptic curve is generated, which will be used for signing and encrypting.
- **1.3:** The public key is derived from the private key, used for encrypting the symmetric key.
- **1.4:** A digital signature (DS) is created for the hash value (H) using the private key, ensuring that the original data is verifiable by the holder of the private key.

Step 2: Encrypt Symmetric Blowfish Key

- **2.1:** The symmetric key (K) is encrypted using the public elliptic curve key, ensuring the symmetric key's security during transmission.

Step 3: Encrypt Original Data

- **3.1:** The original data (M) is encrypted using the Blowfish algorithm with the symmetric key (K). The result of this step is the encrypted data (C).

Step 4: Combine and Send

- **4.1:** The encrypted data (C), digital signature (DS), and encrypted key (encrypted Key) are combined into a single package for transmission.
- **4.2:** The Transmit Package is sent to the recipient, allowing them to retrieve the plaintext using the encrypted key and verify the signature.

6. ANALYSIS

The suggested approach was executed within the Eclipse development platform utilizing version 7 of the Java Development Kit (JDK). This JDK comprises two principal libraries: the Java Cryptography Architecture (JCA) and the Java Cryptography Extension (JCE). The functionalities of the JCA are seamlessly integrated into the fundamental Java APIs, offering vital cryptographic capabilities. The hardware used for the implementation was a dual-core Intel processor operating at 2.5 GHz, and all tests were conducted on the Windows 7 operating system. To evaluate the proposed method against various symmetric and asymmetric cryptographic techniques, a range of performance metrics, including processing time, data throughput, and memory utilization, were analyzed. Firstly, small-sized data (50 KB) solutions were evaluated to ascertain and analyze the operation using small-sized data. Thereafter, the solutions were analyzed by comparing their performance at sizes of 1024 KB (1 MB) and 2048 KB (2 MB) (15-17).

6. 1. Analysis Criteria There are specific metrics that can be used to determine the effectiveness of the solution and its viability. The following criteria have been established for use in this study:

- **Execution Time**: This is the time taken to accomplish all processes that require encryption functions. The execution time is calculated as shown in the following equation:

$$\text{Execution Time } (T_e) = T_f - T_i \quad (3^{\wedge})$$

Where:

- T_e = Execution Time
- T_i = Initial time (recorded at the start of the encryption process)
- T_f = Final time (recorded after the encryption process is completed)

To evaluate the execution duration, the initial moment of the encryption procedure is recorded, and the final moment is then deducted from that.

- **Throughput**: Throughput is defined as the number of tasks processed and encrypted within a specific time frame. The throughput for the proposed solutions is calculated using the following equation:

$$\text{Throughput(TP)} = \frac{\text{Data Size(D)}}{\text{Execution Time(Te)}} \quad (3^{9})$$

- Where:
- TP = Throughput
- D = Data Size (in bytes)
- T_e = Execution Time (in seconds)

In this context, throughput is assessed by measuring the amount of data processed over the time required for execution. The proposed method was evaluated with various data sizes, including small-sized data (50 KB), 1024 KB (1 MB), and 2048 KB (2 MB), to comprehensively analyze its performance.

7. RESULTS AND DISCUSSION

The first assessment is conducted over 50 KB of data, as shown in Figure 5. The execution times differ among the algorithms, with RSA taking 800 ms, AES 550 ms, DES 300 ms, 3DES 400 ms, and PRESENT at 800 ms (for

2SH). Overall, the execution times have been reduced, and the encryption tasks are accomplished more quickly than before. This improvement is attributed to the reduced execution time offered by Blowfish, which is recognized as one of the fastest encryption algorithms. This encoding method is renowned for its performance characteristics and demonstrates high efficiency in data processing. Additionally, enhancements in the speed and efficiency of other functionalities were made based on modifications in this area. The primary function of the EC algorithm is to encrypt the key and generate a hash, thereby reducing processing time. Compared to commonly used algorithms, it is significantly faster and consumes less memory.

In Figure 6, the comparison of solutions based on throughput is conducted with the 50 KB data size. The combination of Blowfish and ECC has been evaluated to confirm that it not only enhances speed and efficiency but also increases overall security by leveraging the strengths of both algorithms. However, it is crucial to ensure that this integration does not introduce new vulnerabilities that could compromise the encryption process. This has been tested through various simulations, demonstrating that the combined approach maintains robustness against potential threats. The proposed algorithm shows higher throughput than other solutions, as illustrated in this evaluation. Thus, the optimality rate reaches 100% when compared to RSA, 3DES, DES, AES, and ECDH. The time taken to implement this solution was exceptionally brief.

Given the limited performance observed in the evaluation, it is reasonable to expect this degree of success. In the next experiment, the amount of data being encrypted was identified as a crucial factor impacting the results during execution. Consequently, the size was increased to 1 megabyte to allow a fair comparison among all algorithms for this specified data size. Figure 7 demonstrates that our solution outperformed other algorithms, as indicated by the results presented by Khezri et al. (5).

Specifically, execution times improved by 240 ms, 240 ms, 130 ms, 80 ms, and 330 ms compared to other symmetric and asymmetric algorithms, highlighting the superior speed of the proposed solution. Likewise, the computation time has decreased by more than 731 ms compared to the RSA asymmetric algorithm, further establishing the faster implement ability of our solution. Furthermore, the throughput rate of the solutions was above average due to the reduced execution time, as shown in Figure 7.

At the same time, in the earlier scenario, RSA achieved performance that exceeded it by over 50%. The parallelization and optimization implemented in this segment significantly enhanced both execution time and throughput.

In this final assessment, we evaluate the algorithm using a sample data volume of 2 MB. Figures 9 and 10 clearly illustrate the advantages of this algorithm in comparison to the other two algorithms. It is essential to note that as the data size increases, there is a proportional rise in the time required for encryption. However, even with this increased demand, our algorithm consistently delivered faster operation times than the DES, 3DES, and RSA algorithms.

Specifically, the proposed solution was able to complete encryption tasks 470 ms and 320 ms faster than the two symmetric algorithms (3DES and DES),

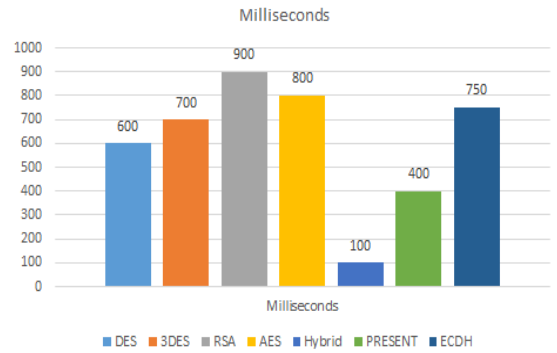


Figure 5. Execution time recorded in milliseconds (for a data size of 0.05 megabytes)

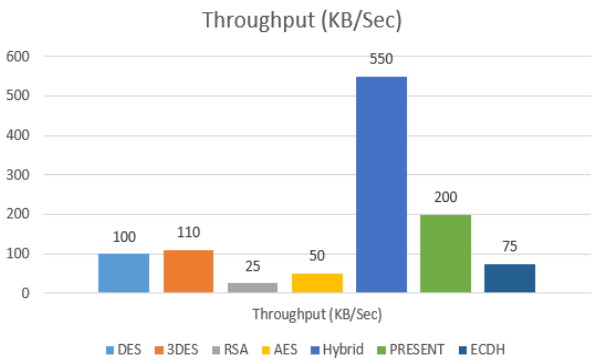


Figure 6. Efficiency measured (data size: 0.05 megabytes)

Figure 7. Execution duration measured in milliseconds (data size: 1 MB)

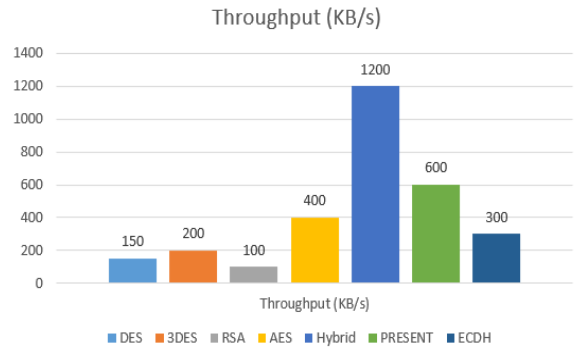


Figure 8. Throughput expressed in kilobytes per second (data size: 1 MB)

Figure 9. Operational time (data size: 2.0 MB)

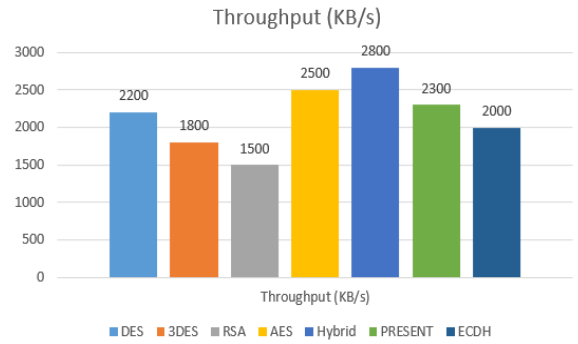


Figure 10. Efficiency indicated (data size: 2.0 MB)

respectively. Moreover, throughput demonstrated a significant increase, as depicted in Figures 8 and 9. During this experiment, the asymmetric RSA failed to show any improvement in performance compared to the symmetric AES.

It is important to highlight that the proposed solution encompasses additional features that extend beyond the capabilities of the conventional AES algorithm, including digital signatures and hashing. As a result, an increase in execution time is justified and is not expected to substantially affect the overall efficiency of the IoT infrastructure. Furthermore, the enhanced features of the proposed solution are instrumental in improving security and efficiency, indicating that the minor increase in execution time is acceptable within the broader context of IoT applications. Through these tests, we have demonstrated that the proposed algorithm not only meets but exceeds the performance metrics established by existing encryption methods.

7. 1. Memory Usage Analysis Another crucial element to consider is the memory capacity needed by encryption algorithms. The memory consumption induced by the developed technique is analyzed and compared with other commonly used encryption techniques. The outcomes of this evaluation are displayed in Figure 11. For this experiment, we considered a data size equal to 50 KB. Thus, as demonstrated in Figure 11 and Table 2, the lowest amount of memory consumption is associated with the suggested approach. The performance stems from the optimization of the cryptographic core; in this case, stemming from the Blowfish algorithm, achieving superior throughput while minimizing memory usage. Due to its reputation, the main features of the Blowfish algorithm include memory efficiency and speed of operations.

The innovative application of a 164-bit key makes this algorithm one of the most efficient asymmetric options regarding memory consumption. The memory requirements are so low that it can be effectively utilized within older systems and even integrated into the smallest embedded platforms.

In Table 3, the data indicates significant differences in memory requirements, with RSA demanding the most resources at 30.5 MB, while the hybrid approach is the most efficient, utilizing only 9.5 MB. Other algorithms such as 3DES and DES display moderate memory usage, with values of 20.5 MB and 18.0 MB, respectively.

In contrast, AES and ECDH each require 14.5 MB, while PRESENT's memory usage stands at 18.0 MB. This analysis highlights the varying efficiency of each method concerning memory utilization (18).

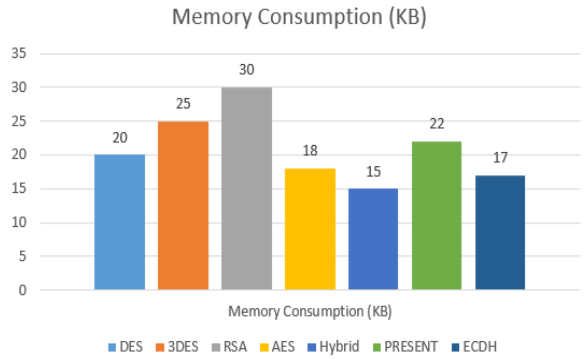


Figure 11. Memory usage by the solutions

TABLE 3. Memory Utilization Among the Methods Compared

Procedures for Algorithms	Memory Usage
DES	18.0
3DES	20.5
RSA	30.5
AES	14.5
Hybrid	9.5
PRESENT	18
ECDH	14.5

8. CONCLUSION AND FUTURE WORK

This research proposes a robust encryption strategy designed to enhance security in IoT-enabled computing environments, focusing on efficiency, security measures, and data integrity. The use of EC (Elliptic Curve) asymmetric encryption is pivotal for secure key sharing, addressing the limitations of symmetric algorithms. Additionally, SHA-256 digital signatures have proven effective in ensuring data integrity and validating transmitted information.

Compared to standard cryptographic algorithms like AES, DES, 3DES, and RSA, the proposed method demonstrates superior execution time within the IoT framework. ECC enables smaller key sizes without sacrificing security, showing that the proposed solution not only enhances efficiency but also provides comparable or better security, especially in resource-constrained scenarios.

However, the effectiveness of this method has yet to be tested in real-world IoT environments. Future work should include deploying the solution in practical settings to assess its performance under conditions typical of IoT devices, such as limited battery life, processing power, and network bandwidth.

Despite these advantages, the proposed method has limitations. Scalability is a significant concern, particularly due to the need for larger key sizes to maintain security, which may strain low-power IoT devices. Furthermore, using Blowfish encryption necessitates a single key for both encoding and decoding, imposing challenges for scalability with numerous connected devices. Addressing these issues will require optimizing encryption algorithms to efficiently support larger networks.

Additionally, using Blowfish in conjunction with EC encryption may introduce potential drawbacks, such as increased complexity and the challenge of managing multiple encryption schemes, which could affect overall security if not properly implemented. Innovations and further optimizations, including leveraging machine learning and artificial intelligence for threat detection, could enhance the efficiency and applicability of the proposed solution.

Key areas for future research include:

- **Improving scalability for algorithms**: Optimizing encryption algorithms to accommodate a larger number of connected devices.
- **Investigating larger key sizes**: Analyzing the impact of larger key sizes on security and performance.
- **Secure data transmission methods**: Exploring effective methods for secure transmission in dynamic, resource-constrained environments.
- **Optimizing encryption algorithms**: Further research on symmetric and asymmetric algorithms to boost efficiency.

These future endeavors aim to strengthen security and efficiency in IoT infrastructures while paving the way for innovative solutions to meet the diverse challenges ahead. Overall, this research highlights the necessity for comprehensive solutions in IoT security.

9. REFERENCES

1. Rana M, Mamun Q, Islam R. Lightweight cryptography in IoT networks: A survey. Future Generation Computer Systems. 2022;129:77-89. 10.1016/j.future.2021.11.011
2. Kumar V, Malik N, Singla J, Jhanjhi N, Amsaad F, Razaque A. Light weight authentication scheme for smart home iot devices. Cryptography. 2022;6(3):37. 10.3390/cryptography6030037
3. Liao Y, Tang Z, Gao K, Trik M. Optimization of resources in intelligent electronic health systems based on Internet of Things to predict heart diseases via artificial neural network. Heliyon. 2024;10(11). j. heliyon. 2024.e32090
4. Panahi U, Bayılmış C. Enabling secure data transmission for wireless sensor networks based IoT applications. Ain Shams Engineering Journal. 2023;14(2):101866. 10.1016/j.asej.2022.101866
5. Khezri E, Zeinali E, Sargolzaey H. SGHRP: Secure Greedy Highway Routing Protocol with authentication and increased privacy in vehicular ad hoc networks. Plos one. 2023;18(4):e0282031. 10.1371/journal.pone.0282031
6. Trik M, Molk AMNG, Ghasemi F, Pouryeganeh P. A hybrid selection strategy based on traffic analysis for improving performance in networks on chip. Journal of Sensors. 2022;2022(1):3112170. 10.1155/2022/3112170
7. Nasiraee H, Ashouri-Talouki M. DoS-Resistant Attribute-Based Encryption in Mobile Cloud Computing with Revocation. International Journal of Engineering Transactions C: Aspects. 2019;32(9):1290-8. 10.5829/ije.2019.32.09c.09
8. Devi RA, Arunachalam A. Enhancement of IoT device security using an Improved Elliptic Curve Cryptography algorithm and malware detection utilizing deep LSTM. High-Confidence Computing. 2023;3(2):100117. 10.1016/j.hcc.2023.100117
9. Trik M, Pour Mozaffari S, Bidgoli AM. Providing an adaptive routing along with a hybrid selection strategy to increase efficiency in NoC-based neuromorphic systems. Computational Intelligence and Neuroscience. 2021;2021(1):8338903. 10.1155/2021/8338903
10. Wang Z, Jin Z, Yang Z, Zhao W, Trik M. Increasing efficiency for routing in internet of things using binary gray wolf optimization and fuzzy logic. Journal of King Saud University-Computer and Information Sciences. 2023;35(9):101732. 10.1016/j.jksuci.2023.101732

11. Puneeth R, Parthasarathy G. Security and data privacy of medical information in blockchain using lightweight cryptographic system. International Journal of Engineering Transactions B: Applications. 2023;36(5):925-33. 10.5829/ije.2023.36.05b.09
12. Fadia T, Toufik L. Elliptic curves cryptography for lightweight devices in IoT system. Brazilian Journal of Technology. 2024;7(4):e73725-e. 10.38152/bjtv7n4-003
13. Mahandru T KR. A study on the impact of cloud computing on the digital transformation of organizations. International Journal of Computer Applications. 2018;179(22):41-4. 10.5120/ijca2018916422
14. Bhagwatrao G, Lakshmanan R. A Novel Approach to use Deep Dyna Q Learning for Enhancing Selection and Performace of Encryption and Hashing Techniques in Remote Healthcare Environment. International Journal of Engineering Transactions A: Basics. 2024;38(1). 10.5829/ije.2025.38.01a.07
15. Prashant MSH, Amrinder Kaur, Pankaj Yadav. . Comparative analysis of AES and RSA with other encryption techniques for secure communication. International Journal of Scientific Research in Computer Science Engineering and Information Technology. 2024;10(2):565-74. 10.32628/CSEIT2410263
16. Qader RAHA, AL-Wattar AHS. A Review of the Blowfish Algorithm Modifications in Terms of Execution Time and Security. Technium Science Journal. 2022;4(9):89-101. 10.47577/technium. v4i9
17. Kapalova N, Algazy K, Haumen A. DEVELOPMENT OF A NEW LIGHTWEIGHT ENCRYPTION ALGORITHM. Eastern-European Journal of Enterprise Technologies. 2023;123(9). 10.15587/1729-4061.2023.280055
18. Trick M, Boukani B. Placement algorithms and logic on logic (LOL) 3D integration. Journal of mathematics and computer science. 2014;8(2):128-36. 10.22436/jmcs.08.02.04
19. Obaidat M, Brown J, Obeidat S, Rawashdeh M. A hybrid dynamic encryption scheme for multi-factor verification: a novel paradigm for remote authentication. Sensors. 2020;20(15):4212. 10.3390/s20154212
20. Olutola A, Olumuyiwa M. Comparative analysis of encryption algorithms. European Journal of Technology. 2023;7(1):1-9. 10.47672/ejt.1312
21. Lo O, Buchanan WJ, Carson D, editors. Correlation power analysis on the PRESENT block cipher on an embedded device. Proceedings of the 13th International Conference on Availability, Reliability and Security; 2018. 10.1145/3230833.3232801
22. Khezri E, Zeinali E, Sargolzaey H. A Novel Highway Routing Protocol in Vehicular Ad Hoc Networks Using VMaSC-LTE and DBA-MAC Protocols. Wireless communications and mobile computing. 2022;2022(1):1680507. 10.1155/2022/1680507
23. Khalafi M, Boob D, editors. Accelerated primal-dual methods for convex-strongly-concave saddle point problems. International conference on machine learning; 2023: PMLR. 10.48550/arXiv.2209.04604
24. Khezri E, Yahya RO, Hassanzadeh H, Mohaidat M, Ahmadi S, Trik M. DLJSF: data-locality aware job scheduling IoT tasks in fog-cloud computing environments. Results in Engineering. 2024;21:101780. 10.1016/j.rineng.2024.101780
25. Ding X, Yao R, Khezri E. An efficient algorithm for optimal route node sensing in smart tourism Urban traffic based on priority constraints. Wireless Networks. 2024;30(9):7189-206. 10.21203/rs.3.rs-3276051/v1
26. Zhu J, Hu C, Khezri E, Ghazali MMM. Edge intelligence-assisted animation design with large models: a survey. Journal of Cloud Computing. 2024;13(1):48. 10.1186/s13677-024-00601-3
27. Hosseini A, Rahaeifard M, Mojahedi M. Analytical and numerical investigations of the ultrasonic microprobe considering size effects. Mechanics of Advanced Materials and Structures. 2020;27(24):2043-51. 10.1080/15376494.2018.1539890
28. Karabulut E, Gholizadeh F, Akhavan-Tabatabaei R. The value of adaptive menu sizes in peer-to-peer platforms. Transportation Research Part C: Emerging Technologies. 2022;145:103948. 10.2139/ssrn.4073150
29. Golovko G, Kalynovych M. SPECIFICS OF IMPLEMENTATION OF THE ASYMMETRIC ENCRYPTION ALGORITHM ON ELLIPTIC CURVES. Системи управління, навігації та зв'язку Збірник наукових праць. 2023;1(71):84-90. 10.26906/SUNZ.2023.1.084
30. Goyal TK, Sahula V, Kumawat D. Energy efficient lightweight cryptography algorithms for IoT devices. IETE Journal of Research. 2022;68(3):1722-35. 10.1080/03772063.2019.1670103

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

رشد سریع اینترنت اشیاء (IoT) اهمیت رسیدگی به آسیب‌پذیری‌های امنیتی در دستگاه‌های متصل با منابع محدود را افزایش داده است. حفاظت از اطلاعات حساس کاربران بر لزوم امن‌سازی داده‌های شخصی تأکید می‌کند. تهدیدهای دسترسی غیرمجاز نیازمند راه‌حل‌های پیشرفته رمزنگاری است تا امنیت را افزایش داده و عملکرد سیستم را بهینه کند. این مطالعه یک چارچوب رمزنگاری هیبریدی نوآورانه را پیشنهاد می‌دهد که به‌طور منحصربه‌فرد، رمزنگاری بر مبنای منحنی‌های بیضوی (ECC) و الگوریتم متقارن Blowfishرا ترکیب می‌کند. برخلاف روش‌های مرسوم مانندRSA ، AESیا ECC مستقل، رویکرد ما این تکنیک‌ها را به روشی نوآورانه ترکیب کرده است تا کارایی و امنیت برتر را فراهم کند. همچنین، یک پروتکل رمزنگاری تطبیقی جدید توسعه یافته است که برای حجم داده و قابلیت‌های دستگاه‌ها بهینه شده است Blowfish .با کارایی بالا، داده‌های بزرگ (بیشتر از ٦٤ کیلوبایت) را در کوتاه‌ترین زمان ممکن رمزگذاری می‌کند. علاوه بر این، ECCبا استفاده از کلیدهای کوچک‌تر، امنیت را بدون بار اضافی در پردازش افزایش می‌دهد. در این پژوهش، داده‌هایی از دستگاه‌های مختلف IoT مانند حسگرها، تجهیزات پوشیدنی، سیستم‌های هوشمند منزل، سیستم‌های ترافیک و ابزارهای صنعتی جمع‌آوری شد و در آزمایشگاه‌ها و محیط‌های واقعی مورد آزمایش قرار گرفتند. این داده‌ها شامل انواع مختلفی از اطلاعات حساس بودند و هر مجموعه داده به طور میانگین بیش از ۵۰ کیلوبایت حجم داشت. همچنین، تمام آزمایش‌ها در سیستم‌هایی با پردازنده دو هسته‌ای اینتل با سرعت ۲.۵ گیگاهرتز انجام شد تا ارزیابی عملکرد یکنواختی صورت گیرد. در مقایسه با راه‌حل‌های رمزنگاری سنتی مانندRSA ، AESو پیاده‌سازی‌های مستقلECC ، رویکرد ما زمان اجرای بیش از ۱۵ درصد سریع‌تر را نشان می‌دهد و کارایی کلی سیستم را به طور چشم‌گیری افزایش می‌دهد. این تحقیق نشان می‌دهد که رویکرد هیبریدی ما، امنیت IoT را بهبود می‌بخشد و در عین حال مصرف توان و انرژی را کاهش می‌دهد، که این امر به بهبود عملکرد دستگاه‌های آینده کمک می‌کند.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1254-1266

International Journal of Engineering

Journal Homepage: www.ije.ir

Numerical Evaluation of Behavior of Tapered Piled Raft in Loose Sand

M. S. Shokoohfar, M. Khanmohammadi*

Department of Civil Engineering, Isfahan University of Technology, Isfahan, Iran

P A P E R I N F O

Paper history:
Received 21 April 025
Received in revised form 19 May 2025
Accepted 04 July 2025

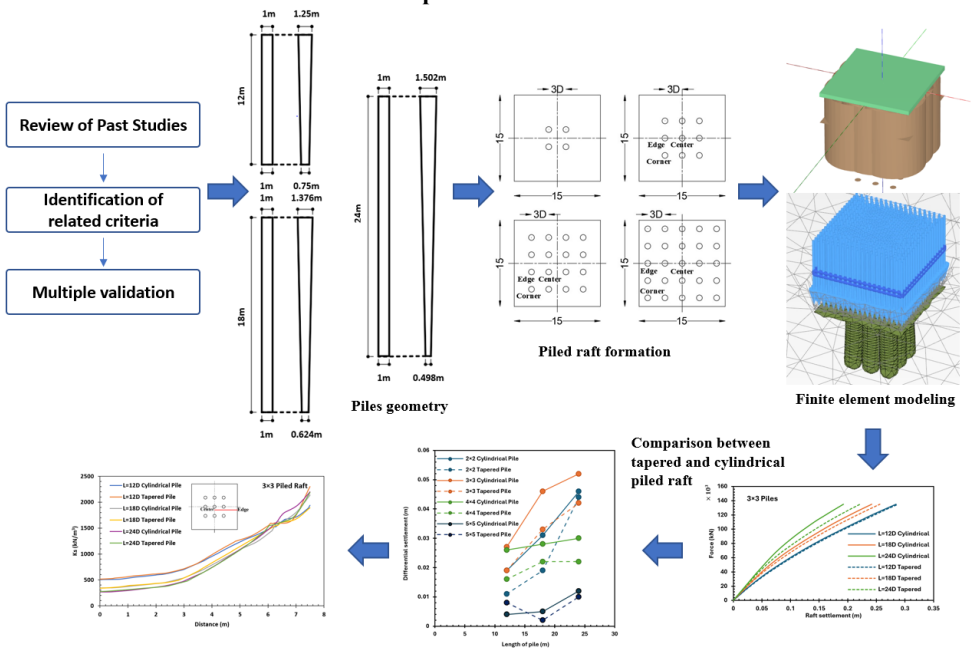
Keywords:
Piled Raft
Tapered Pile
Bearing Capacity
Finite Element
Settlement

A B S T R A C T

Tapered piles, with a larger diameter at the top than at the bottom, provide higher shaft resistance than conventional cylindrical piles and may offer a more economical alternative. Adding piles to a raft foundation improves its bearing capacity and reduces settlement. While tapered piles have been investigated for decades, limited research has focused on tapered piled raft systems. This study evaluates the behavior of piled rafts incorporating tapered and cylindrical piles of varying lengths and configurations in sandy soil. The model includes a raft measuring 15 × 15 × 0.7 meters, and piles with an average diameter of 1 meter. The taper angle of 1.20 degrees results in top diameters of 1.25, 1.376, and 1.502 meters and bottom diameters of 0.75, 0.624, and 0.498 meters for pile lengths of 12, 18, and 24 meters, respectively. A 3D finite element model was developed using PLAXIS 3D. Multiple simulations were performed with various pile types and arrangements. The model was validated against three experimental studies, confirming its reliability. Results show that tapered piles offer greater shaft resistance than cylindrical piles. Central piles carry less load compared to edge and corner piles. Increasing the number of piles reduces differential settlement. Additionally, tapered piled rafts demonstrated lower differential settlement than cylindrical piled rafts, highlighting their effectiveness in enhancing overall performance.

doi: *10.5829/ije.2026.39.05b.17*

Graphical Abstract



*Corresponding Author Email: mkhanmohammadi@iut.ac.ir (M. Khanmohammadi)

Please cite this article as: Shokoohfar MS, Khanmohammadi M. Numerical Evaluation of Behavior of Tapered Piled Raft in Loose Sand. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1254-66.

1. INTRODUCTION

Utilizing piles with raft is considered as a solution when the surface layer of soil does not have the necessary load bearing capacity. Meanwhile, tapered piles with increased shaft bearing capacity can be an ideal option to operate along the raft. Although the group performance of these piles has received the attention of several researchers. However, the utilization of tapered piles beneath the raft and the interaction between them have not been investigated. Therefore, the shortage of research about the group of tapered piles, especially tapered piled rafts, doubles the need to investigate this issue.

Previous studies indicate that tapered piles provide larger shaft bearing than cylindrical piles with equal volume (1-8). The performance of different geometrical parameters on tapered piles has been examined. El Naggar and Wei (5) concluded that the optimal tapered effect is considerable in the depth of 20 times pile diameter (L=20D) in compressive loads. The shaft bearing capacity of tapered piles was evaluated by Khan et al. (6) utilizing experimental and finite element analyses, leading to the conclusion that a greater taper angle increases the shaft resistance. Tavasoli and Ghazavi (7) have conducted experimental and numerical analysis to investigate the drivability of tapered and semi-tapered piles. They concluded that decreasing the pile tip diameter increased both the velocity and the settlement of the piles. El Naggar and Sakr (9) concluded that the length of the tapered piles should not exceed 20 to 25 times the diameter of the pile (L=25D) in order to achieve optimal efficiency.

In the last two decades, experimental research has been done on the tapered pile group (10, 11) and some numerical modelling (12, 13) to investigate the tapered effect on the pile group. Shafaghat et al. (13) extended the pile group efficiency relationship proposed by Sayed and Bakeer (14) to calculate the group efficiency for bored tapered piles.

Randolph (15) proposed three approaches for designing of piled raft. The first is the conventional method where the piles are built to support the majority of the load. The second method, where the piles are built to withstand 70-80% of operational load. The last method is where the piles are utilized to decrease the differential settlement. Research on piled raft foundations mainly includes numerical modelling (16-23) due to the difficulty of conducting experimental studies. However, only a few of these researches have focused on experimental studies (24-26). Sinha and Hanna (21) conducted a numerical study on the performance of the piled raft in cohesive soil. They observed that by increasing the pile spacing up to 6D, the bearing capacity of the piled raft decreases and after that, the raft carries almost the full load of the building. Lee et al. (19) evaluated the response of piled rafts under vertical loading. Their result showed that the loading type (distributed or concentrated) affects load distribution in the pile according to its arrangement in the raft. Luo et al. (27) studied the response of rigid piled rafts in clay soil. They reported that the normal settlement reduces with the increase of the safety factor, they also suggested a safety factor of 3 for the designing of the piled raft. Khanmohammadi and Fakharian (22) analyzed the response of piled rafts with various lengths, spacings, and arrangements. They concluded that increasing the pile length and reducing the pile spacing, decreases the settlement and increases the contribution of piles in load bearing. However, pile spacing for rafts on dense soils has less effect than rafts placed on looser soils. Banerjee et al. (28) have investigated the performance of piled raft in multi-layered soil using finite element analysis. They concluded that to achieve an efficient performance the pile group-raft area ratio is between 0.4 to 0.6. Jeong et al. (29) have investigated piled raft deformation and load distribution using numerical and analytical methods. They concluded that the settlement performance and load distribution of piled raft is reliant on foundation stiffness (pile-raft-soil interaction and raft flexibility).

The load applied to the piled rafts can be transmitted to the soil through the raft or the piles. To determine the effect of piles in the raft and also to determine their load-carrying contribution, the piled raft coefficient is characterized as follows:

$$\alpha = \frac{Q_P}{Q} \quad (1)$$

where Q_P is the part of the load supported by the piles and Q is the total applied load. It is noteworthy that α=0 is equal to a raft without any piles and α=1 is equal to a piled raft where the cap has no interaction with the surface of the soil (and provides no load bearing contribution).

In this study, finite element analysis has been conducted to evaluate the performance of tapered piled raft. The impact of pile length and arrangement for tapered and cylindrical piles has been evaluated.

2. MESHING AND BOUNDARY CONDITIONS

The performance of the tapered piled raft was determined by three-dimensional finite element analysis. PLAXIS 3D (2023) finite element software was chosen for modelling. Figure 1 shows the generated mesh for models. A 10-node tetrahedral elements fine mesh was employed for all models, and by using the coarseness factor for the surfaces of the piles and raft, smaller meshes have been created around the piles and raft than in other parts of the models. A square raft with a width and thickness of 15 meters and 0.7 meters respectively and piles with an average diameter of 1 meter with

various lengths (length to pile diameter ratio of 12, 18, and 24) were positioned in the center of the raft. The piles were modeled using volume elements. This approach allows for accurate modelling of the geometry of both tapered and conventional piles and simulates the pile-soil interaction around the pile circumference. Figure 2 shows the raft geometry and positioning of the piles.

The relative flexibility of the rectangular raft was calculated with the criterion provided by Horikoshi and Randolph (30) which is specified in the following relationship:

$$K_{rs} = 5.57 \frac{E_r(1-v_s^2)}{E_s(1-v_r^2)} \left(\frac{B}{L}\right)^{\alpha} \left(\frac{t_r}{L}\right)^3 \quad (2)$$

where E_r, v_r, E_s and v_s are raft Young's modulus, raft Poisson ratio, soil Young's modulus and soil Poisson ratio respectively, B and L are the lateral and longitudinal dimension of the raft, t_r is the raft thickness and α is aspect ratio. Differential settlement is dependent on the raft aspect ratio (B/L). To minimize this dependency, $\alpha = 0.5$ has been considered. By using the materials properties specified in Table 2 the value of K_{rs} will be equal to 0.575 which represents a flexible piled raft. Raft relative flexibility is important because it directly affects raft center settlement and differential settlement.

Khanmohammadi and Fakharian (31) concluded that the soil in proximity of the pile is affected by pile installation. Therefore, to avoid the influence of model dimensions on finite element analysis, the boundaries considered on each side of the raft are 4 times the length of the raft, and the depth of the model is 8 times the length of the rafts, which results in a cube with dimensions of 135 meters as illustrated in Figure 3. The boundary conditions for the vertical boundaries are (normally fixed), which expresses the freedom of movement in the Z direction, and for the bottom horizontal boundary, boundary conditions are fully fixed.

3. CONSTITUTIVE MODELLING

The behavior of loose sand has been modelled by the Mohr-Coulomb model, and it is expressed as follows:

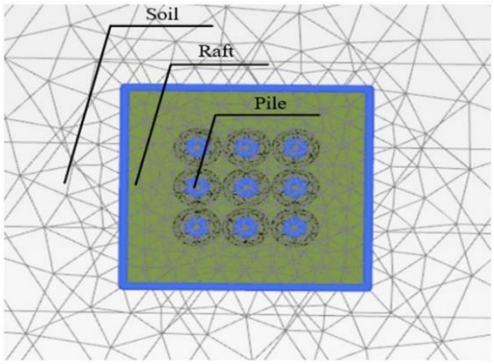


Figure 1. Finite element mesh around the piled

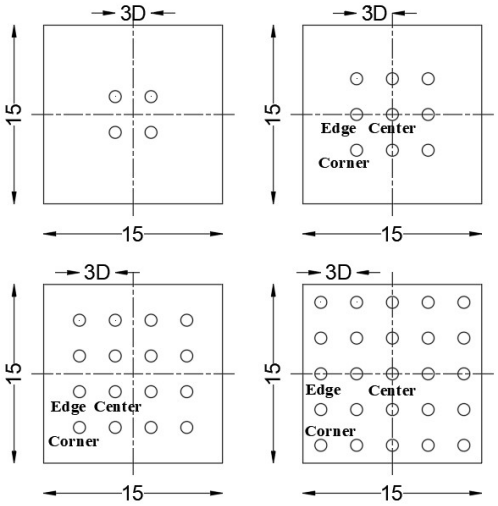


Figure 2. Plies arrangement used in a finite element analysis

$$\tau = c + \sigma \tan\varphi \quad (3)$$

where τ and σ are the shear stress and normal stress at the failure plane, c is cohesion and φ is the soil friction angle. Also, the behavior of concrete used for modelling of piles and raft was modeled through the linear elastic model according to Hooke's isotropic elasticity law. The Mohr-Columb model has also been used to simulate interfaces including soil-raft and soil-piles. The interface element between the soil and the piled raft will cause small displacements and sliding between the two surfaces. The parameters of the interface will be determined using the parameter R_{inter} and it is calculated with the defined parameters of the soil around the interface element as follows (32):

$$c_i = R_{inter} \times c_{soil} \quad (4)$$

$$\tan\varphi_i = R_{inter} \times (\tan\varphi_{soil}) \quad (5)$$

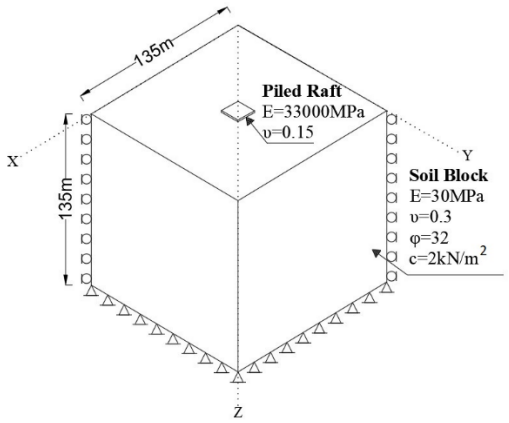


Figure 3. Schematic sketch of the numerical model and boundary condition

where c_i is the cohesion of the interface and φ_i is the internal friction angle of the interface.

The piled raft and the soil were modelled using characteristics described in Table 1. The considered soil is loose sand modelled by the Mohr-Coulomb model. The characteristics of loose sand considered in this study have a good match with values considered for loose sand by Bowles and Guo (33).

A similar value for Rinter has been considered for better comparison with similar studies which is used for validation as well. Additionally, to achieve the precise modelling of the soil-pile interface, the R_{inter} coefficient for sand in contact with concrete considered in this study was compared with the studies of Potyondy (34), Praporgescu and Popa (35) a good match was obtained. The average diameter for all the piles is considered 1 meter and the raft dimensions are 15×15×0.7 and the loading was distributed uniformly on the raft. The cross-sectional area of tapered piles changes linearly from top to bottom (for example, for a 12 meter pile, the cross-sectional area of the pile starts at a diameter of 1.25 meter at the top, reaches 1 m in the middle of the pile, and reaches 0.75 meter at the pile tip). The variables are the number, length, and tapered angle of the piles. The geometric specifications of the models are specified in Table 2.

TABLE 1. Material properties used in finite element analysis

Parameter	loose Sand	Concrete
Material Model	Mohr-Coulomb	Linear Elastic
Unit Weight, γ (kN/m^3)	17	25
Elastic modulus, E(MPa)	30	33000
Poisson's ration, υ	0.3	0.15
Cohesion, C (kPa)	2	-
Friction angle, φ˚	32	-
Dilatancy angle, ψ˚	2	-
R_{inter}	0.95	-

TABLE 2. Geometry properties used in finite element analysis

Model	Pile Formation	L/D	α˚	D_{top}(m)	D_{bot}(m)
1	2×2	12	0˚	1	1
2		18		1	1
3		24		1	1
4		12	1.2˚	1.25	0.75
5		18		1.376	0.624
6		24		1.502	0.498
7	3×3	12	0˚	1	1
8		18		1	1
9		24		1	1
10	4×4	12	1.2˚	1.25	0.75
11		18		1.376	0.624
12		24		1.502	0.498
13		12	0˚	1	1
14		18		1	1
15		24		1	1
16		12	1.2˚	1.25	0.75
17		18		1.376	0.624
18		24		1.502	0.498
19	5×5	12	0˚	1	1
20		18		1	1
21		24		1	1
22		12	1.2˚	1.25	0.75
23		18		1.376	0.624
24		24		1.502	0.498

4. VALIDATION

Validation of three experimental studies on piles (tapered and cylindrical single pile, tapered pile group, and piled raft) have been carried out to ensure the accuracy of modelling for this study. The first validation was done to ensure the correct modelling of tapered piles. Wei and El Naggar (4) investigated the axial performance of two types of tapered piles with tapered angles of 0.6 and 0.95 degrees and cylindrical piles in sandy soil inside a chamber with a diameter of 1.5 meters and a depth of 1.445 meters. For accurate modelling of the confining pressure of the soil, a hardening soil model was utilized. The material characteristics utilized in the validations are shown in Table 3. These findings are correlated with the load-settlement performance acquired from the current finite element study employing PLAXIS 3D, in Figure 4a, and it shows a good match.

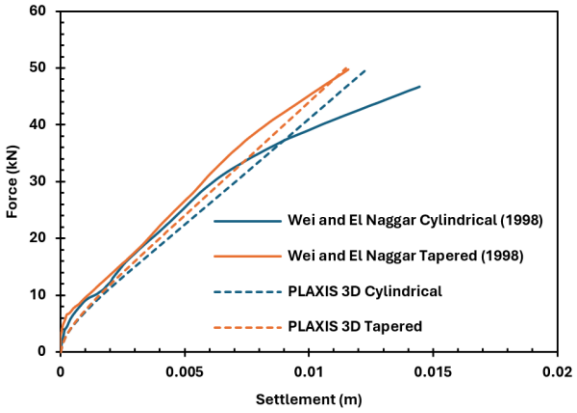
To ensure the correct modelling of the interface between tapered piles, a second validation based on Nasrollahzadeh and Hataf's (11) study was used.

They investigated the performance of the tapered pile group in a cubical chamber with dimensions of 1×1×1 m. Piles in two types of tapered (with an angle of 1°) and cylindrical with a buried length of 0.6 meters and an average diameter of 7.3 cm were used. The material characteristics are provided in Table 3. The load settlement performance from experimental research shows good agreement with the graph obtained from numerical modelling in Figure 4b.

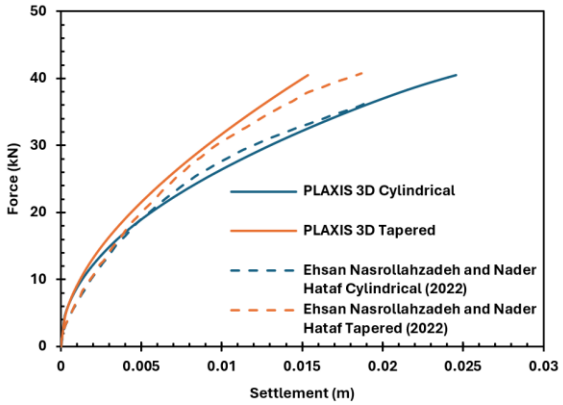
Poulos's (36)study was used to validate the accuracy of piled raft performance on sand. Poulos analyzed, employing various software like FLAC 3D and GASP, a

TABLE 3. Model parameters and properties for various materials used for validations

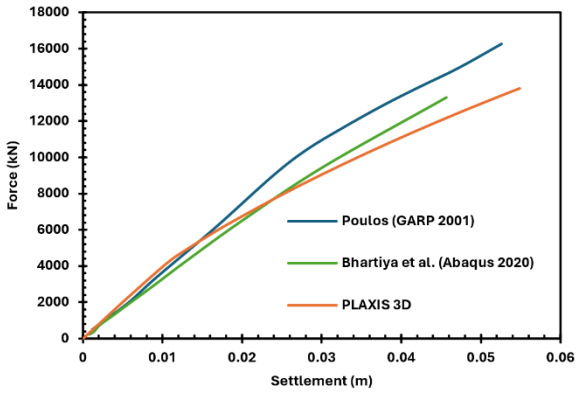
Parameters	Wei and El Naggar (4)	Nasrollahzadeh and Hataf (11)	Poulos (36)
Material Model	Hardening soil	Mohr-Coulomb	Mohr-Coulomb
Unit weight, γ (kN/m^3)	18	17	15.9
Saturated unit weight, γsat (kN/m^3)	20	20	-
Elastic modulus, E (MPa)	-	25	30
Secant stiffness, E50 (MPa)	30	-	-
Tangent stiffness, Eoed (MPa)	30	-	-
Unloading / reloading stiffness, Eur (MPa)	90	-	-
Poisson's ration, υ	0.3	0.35	0.3
Friction angle, φ	35	32	30
Cohesion, C (kPa)	2	4	2
Dilatancy angle, ψ	5	2	0
	Steel Properties	**Concrete Properties**	
Material Model	Linear Elastic	Linear Elastic	
Unit weight, γ (kN/m3)	78.5	24	
Elastic modulus, E(MPa)	2.00E+05	3.00E+04	
Poisson's ration, υ	0.3	0.2	



a) load-settlement responses obtained by Wei and El Naggar for single tapered and cylindrical piles.



b) Load-settlement responses obtained for tapered and cylindrical pile groups by Nasrullahzadeh and Hatef.



c) Load-settlement responses obtained for piled raft analyzed by Poulos.

Figure 4. Validation of PLAXIS 3D finite element analysis

piled raft supported by 9 piles with the diameter of 0.5 and the length of 10 meters respectively, placed under a raft of 0.5×6×10 meters. Table 3 illustrates the material properties used in this validation. The results show a slight difference in the settlement of the middle of the piled raft between Poulos and the present study. Also, Bhar tiya et al. (37) analyzed the piled raft foundation modelled by Poulos with ABAQUS software. Figure 4c shows the load-settlement performance for Poulos and the present study. The obtained load-settlement response for this study has a good match with the previous results, the slight differences can happen due to various assumptions in the elasto-plastic characteristics of sand.

5. RESULTS AND DISCUSSION

5. 1. Piled Raft Bearing Capacity Figure 5 illustrates the load-settlement response for center of piled rafts with various pile length ratios. As expected, the foundation settlement decreases with a rise in the pile length. This process is evident for all modelled piled rafts with any quantity of piles. Also, with a rise in the number of piles, the settlement of the raft center decreases, which is due to the participation of more piles in the load bearing process. This decrease in settlement is greater for piles with a longer length than piles with a shorter length (which leads to the separation of the load-settlement response for piles with various lengths).

By comparing the tapered and cylindrical piled rafts, it is evident that the tapered angle increases the settlement of the piled rafts. Also, a rise in pile length, causes greater settlement for tapered piled rafts than cylindrical ones.

Figure 6 illustrates the load-settlement performance for the three types of piles according to their position in the raft (center, edge, and corner) for the arrangement of 9, 16, and 25 piled rafts is shown. It can be seen that the tapered and cylindrical piles in the center, edge, and corner undergo more settlement respectively. A similar performance can be seen in piled rafts by raising the quantity of piles from 9 to 16. However, the behavior of the piles is different for the 25 piled rafts. Piles position is inclined to the edge of the raft because of the rise in their number (the piles were positioned in the center of the rafts due to the assumed 3D spacing between piles). This causes the piles to have more interaction with adjacent soil (which is not under direct loading). This increase in interaction for corner piles has come from both sides. Figure 6 shows the diagram of the center, edge, and corner piles for a 25 piled raft. It is clear that unlike the raft with 9 and 16 piles, piles in these positions experience less settlement respectively. Khanmohammadi and Fakharian (22) also concluded similar results by analyzing the load-settlement performance of the 64 piled rafts for the piles in the mentioned position. The ratio of the settlement of the edge and corner piles to the center pile is given in Table 4. In this table, the settlement of the center pile is given in Q_{135MN} column (the settlement for 135 MN loading on the raft), and the settlement ratio of the edge and corner piles with the center pile (in constant loading) is determined.

5. 2. Tapered Effect On Differential Settlement of The Piled Raft The settlement difference between the corner and the center of the piled raft is defined as the differential settlement. The number and arrangement of piles are influential in the differential settlement. A rise in the volume of the piles in the center of the piled raft can lead to a rise in the differential settlement. Depending on the arrangement, geometry, and type of the load applied, a piled raft can experience less center settlement, while the corners of the raft (given their distance from the

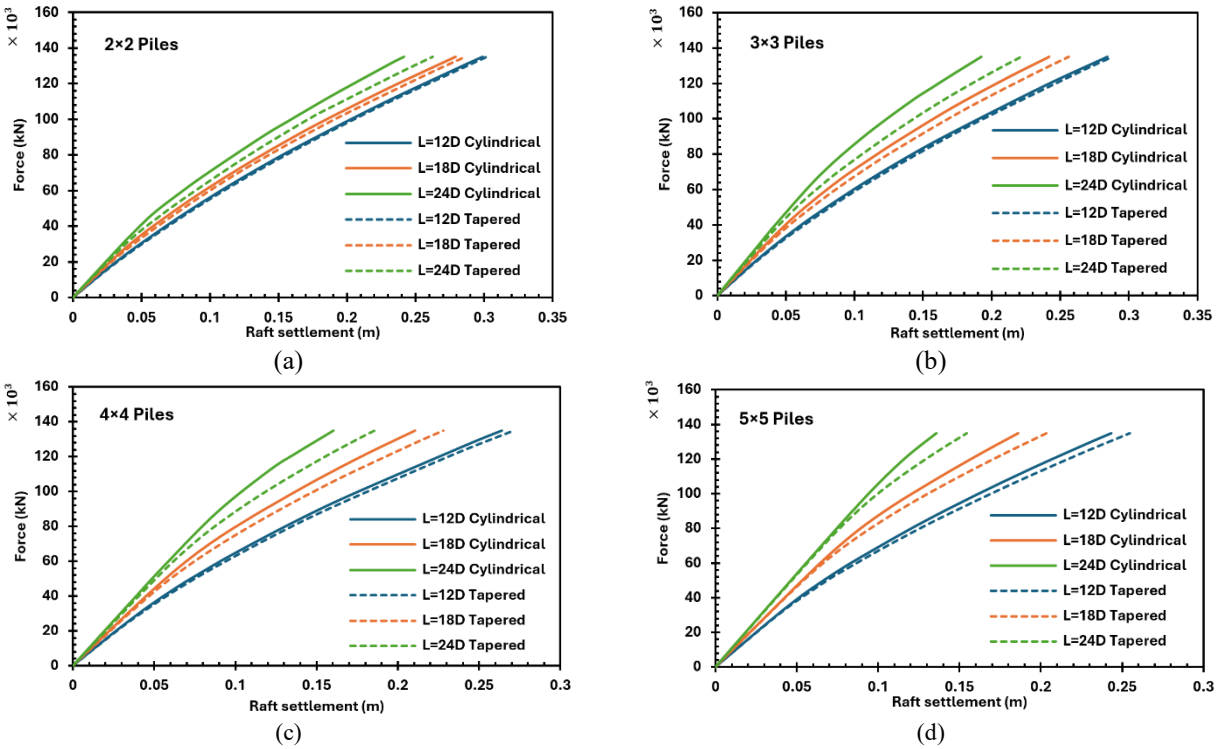


Figure 5. Load-settlement response for the tapered and cylindrical piled raft with a) 4 piles, b) 9 piles, c) 16 piles and d) 25 piles

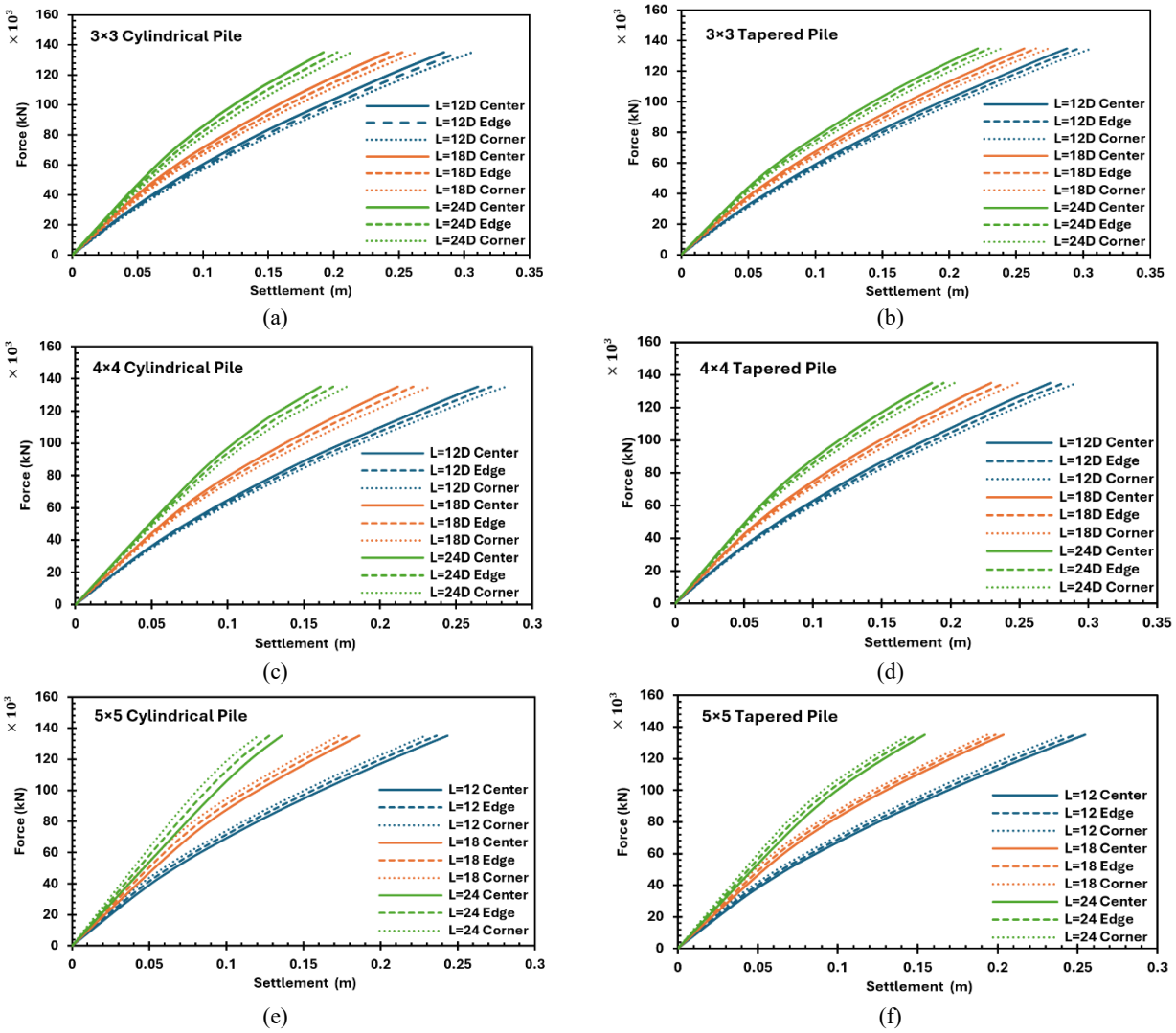


Figure 6. Load-settlement responses for the tapered and cylindrical piled raft for center, edge and corner piles for a) 3×3 cylindrical pile b) 3×3 tapered pile c) 4×4 cylindrical pile d) 4×4 tapered pile e) 5×5 cylindrical pile f) 5×5 cylindrical pile

TABLE 4. Settlement ratio of piles considering their position in raft

Case	Pile Formation	L/D	α°	Q_{135MN}(m)	Q_{edge}/Q_{center}	Q_{corner}/Q_{center}
1	2×2	12	0°	0.299	1	1
2		18		0.279	1	1
3		24		0.242	1	1
4		12	1.2°	0.302	1	1
5		18		0.287	1	1
6		24		0.262	1	1
7	3×3	12	0°	0.284	1.041	1.077
8		18		0.242	1.045	1.089
9		24		0.192	1.054	1.110
10		12	1.2°	0.288	1.030	1.063
11		18		0.256	1.033	1.070
12		24		0.221	1.038	1.079

13	4×4	12	0°	0.264	1.034	1.070
14		18		0.212	1.048	1.099
15		24		0.161	1.051	1.106
16		12	1.2°	0.272	1.035	1.070
17		18		0.229	1.043	1.085
18		24		0.186	1.044	1.088
19	5×5	12	0°	0.243	0.971	0.940
20		18		0.186	0.964	0.929
21		24		0.136	0.940	0.888
22		12	1.2°	0.255	0.976	0.949
23		18		0.204	0.975	0.951
24		24		0.154	0.969	0.939

piles) can experience more settlement, which will lead to differential settlement. Also, a piled raft can experience uniform settlement throughout the piled raft (regardless of the total settlement) by carefully placing the piles in the raft. Figure 7 illustrates that for the piled raft with a small number of piles (4 piles and 9 piles), raising the length of the piles results in an increase in differential settlement, while for more piles, the differential settlement is negligible. It should be mentioned that the differential settlement for the 9 piled raft exceeds the 4 piled raft, which is due to a rise in the volume of the piles and the decrease of the settlement at the center of the raft. Because of the constant distance between piles, increasing the length of the piles leads to a rise in the volume of the piles in a small area beneath the raft, which reduces the settlement, while the settlement in the raft's edges increases in piled rafts with fewer piles. But in the case of more piles and a wider arrangement the differential settlement decreases.

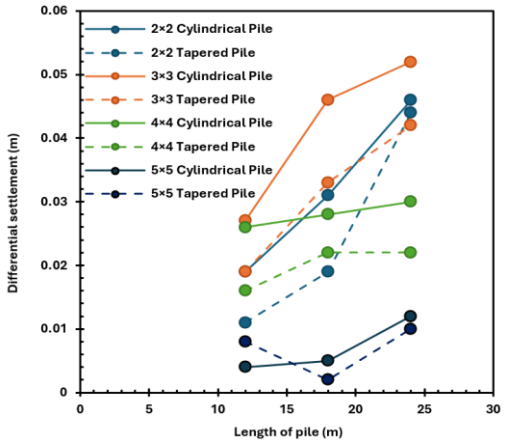


Figure 7. Differential settlement for tapered and cylindrical piled raft

It is also evident that the differential settlement of the tapered piled rafts is less than cylindrical piled rafts, which indicates the positive effect of tapered piles on differential settlement reduction.

5. 3. Tapered Effect on Piled Raft Coefficient The comparison of piled raft coefficient with increasing pile length, for tapered and cylindrical piles is shown in Figure 8. The findings illustrate that the piled raft coefficient increases with a rise in the quantity of the piles, this is because more piles participate in the load bearing and the raft portion of the load decreases. Also, the piled raft coefficient increases with a rise in the length of the piles, which indicates the increase in the shaft bearing capacity of each pile. It is evident that in 25 piled rafts with a rise in pile length from L/D=18 to L/D=24, the piled raft coefficient does not experience a significant increase. Whereas the other piled rafts with relatively lower piled raft coefficients showed more increase. Figure 8 also shows that tapered piles have a lower piled raft coefficient, which is a result of multiple interactions in piled raft foundation.

5. 4. Axial Force In Piles Piles can perform differently considering their position in the raft. Figure 9 to 11 show axial force in tapered and cylindrical piles for center, edge, and corner piles. It is evident that tapered piles have higher shaft bearing capacity at any position than cylindrical piles, which match with previous studies (1, 4, 6).

It is evident that piles located at the center of the raft absorbed lesser loading in comparison with other piles in the raft. This is revealed when the length of piles in the rafts reaches from 12 meters to 18 and 24 meters(according to the quantity of piles in the raft). For example, in Figure 9, with an increase in the pile length from 12 to 18 and 24 meters, it is found that the loading

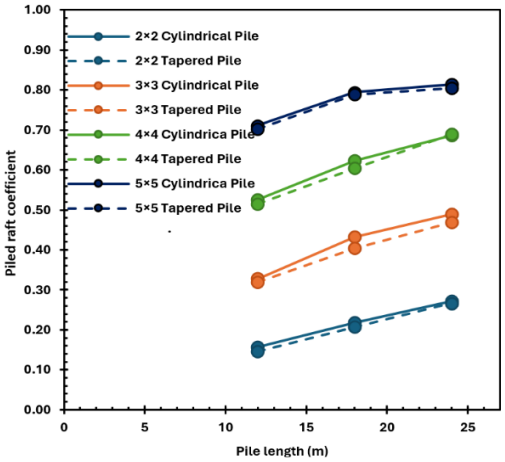


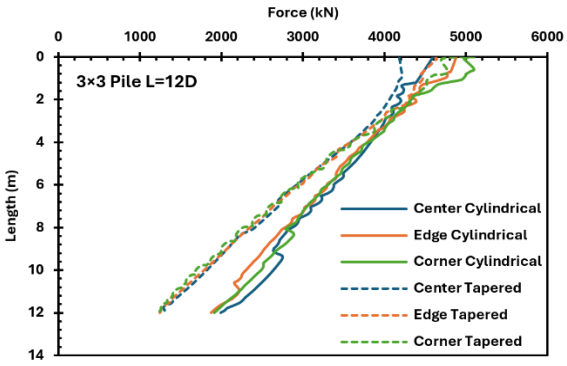
Figure 8. Piled raft coefficient for tapered and cylindrical piled raft

share of piles increases, and it is also observed that the absorbed loading for the center pile has lesser increase than the rest of the piles. Lee et al. (19) indicated that the loading share of the center pile was lower than other piles, which was dependent on the loading level of the piled raft.

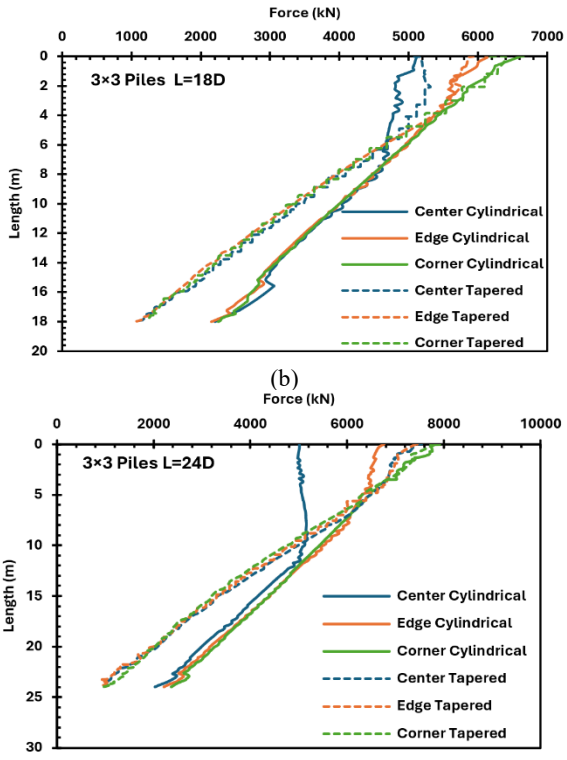
According to Figure 9-11, with an increase in pile length for center piles, the axial force remains relatively constant within a specific depth range (no force transition occurs between pile and soil). Figure 12 illustrates the settlement of the center pile and soil with the length of 12 and 24 m for 4×4 piled raft. It is evident that, the neutral plane for piles with no force transition between pile and soil occurs in deeper layers of soil.

Also, tapered piles have lower tip bearing capacity due to the lesser cross section at the pile's toe than cylindrical ones.

The total stress developed in the soil in the cross-section due to the absence of pore water pressure (the groundwater level was considered to be at the lowest possible level) is equal to the effective soil stress. Figure 13 shows a vertical cross-section of a 5×5 piled raft with a length of 12 times the pile diameter, for both tapered



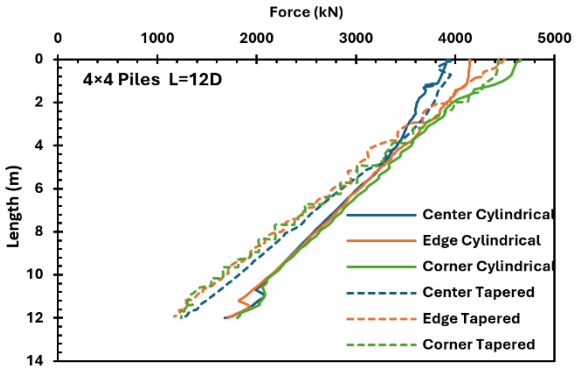
(a)



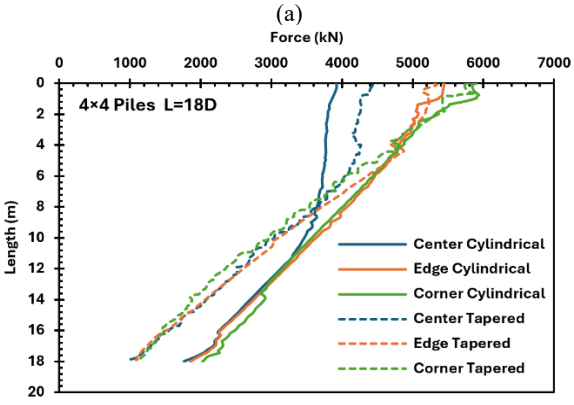
(b)

(c)

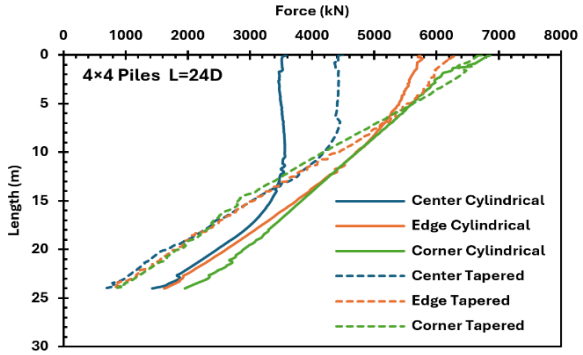
Figure 9. Axial force for tapered and cylindrical piles in center, edge and corner of 3×3 Piled raft with a) L=12D b) L=18D c) L=24D



(a)

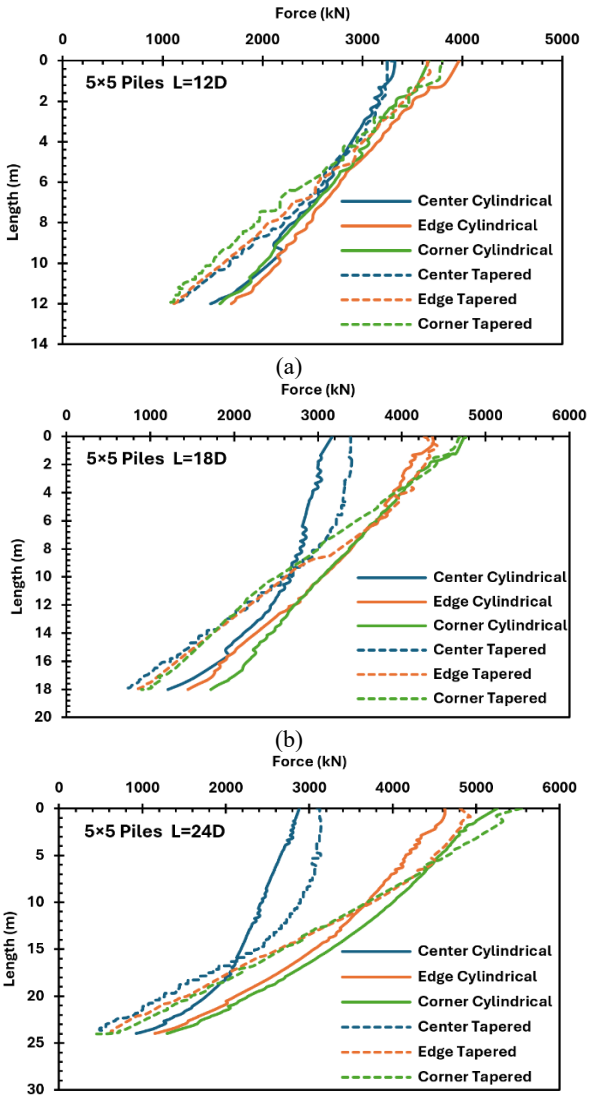


(b)



(c)

Figure 10. Axial force for tapered and cylindrical piles in center, edge and corner of 4×4 piled raft with a) L=12D b) L=18D c) L=24D



(a)

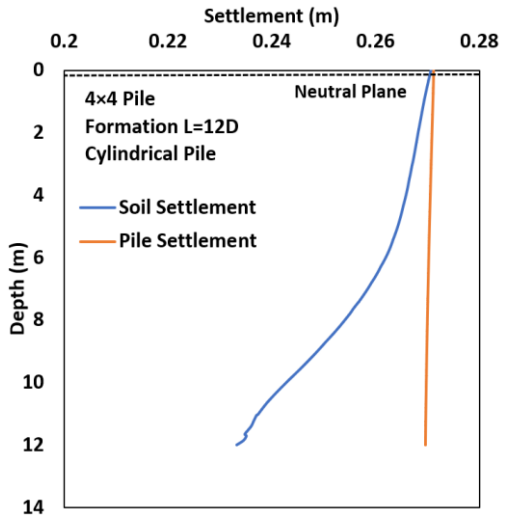
(b)

(c)

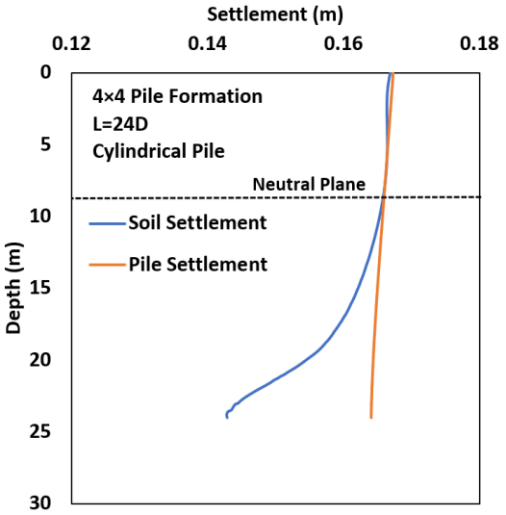
Figure 11. Axial force for tapered and cylindrical piles in center, edge and corner of 5×5 piled raft with a) L=12D b) L=18D c) L=24D

and cylindrical piles. This cross-section, which passes symmetrically through the center of the raft, depicts the stresses resulting from loading and initial soil layer pressure.

According to Figure 13, the stress value under the cylindrical pile tip is -1325 kN/m², while this value reaches -1872 kN/m² for tapered piles. This increase in stress beneath the tapered pile tip is attributed to its smaller cross-sectional area compared to the cylindrical pile (note that part of the stress contour directly under the pile tip has been excluded in the Figure 13 to provide a clearer visualization of the overall stress distribution). It is worth noting that, the stress distribution contour under piles tip is higher for cylindrical piles than the tapered piles beneath the piled raft. On the other hand, the stress distribution contour around the piles shaft shows a greater value for tapered piles.

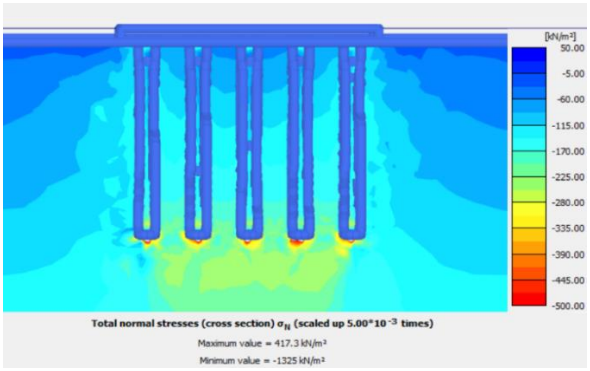


(a)

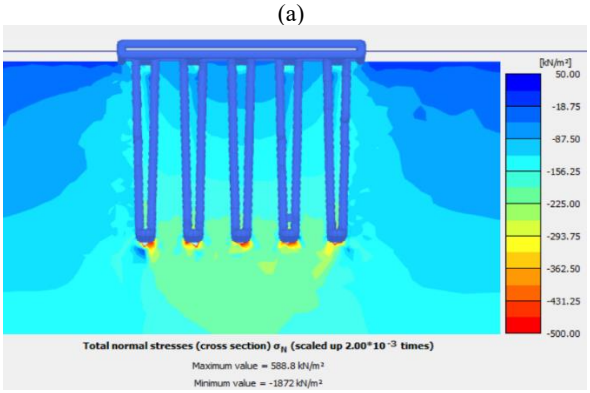


(b)

Figure 12. Illustration of pile-soil settlements and Neutral plane for 4×4 piled raft with a) L=12D b) L=24D



(a)



(b)

Figure 13. Stress distribution for 5×5 piled raft for a) cylindrical piles b) tapered piles

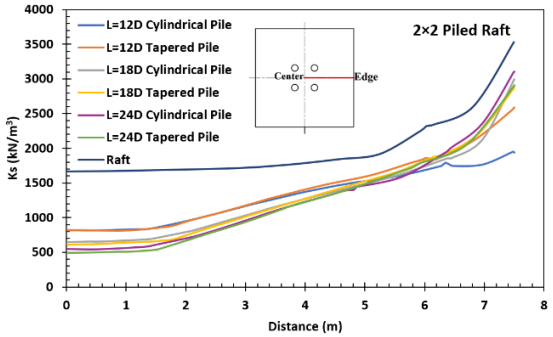
5. 5. Soil Subgrade Modulus Soil subgrade modulus is calculated by obtaining the stress and displacement of the soil in the vertical direction. Bowles and Guo have proposed the following equation to calculate K_s (33):

$$K_s = \frac{\Delta\sigma}{\Delta\delta} \quad (6)$$

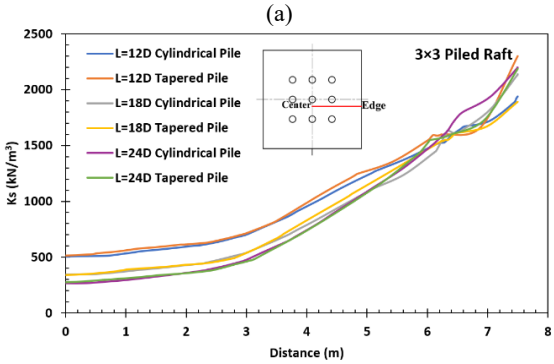
where $\Delta\sigma$ is the increment of contact pressure and $\Delta\delta$ is the corresponding deformation in the vertical direction.

Figure 14 shows the soil subgrade modulus along the raft (from the center to the side for rafts with different piles). It can be seen that the subgrade modulus does not change much with the change in the pile length, which is in accordance with the observations of Khanmohammadi and Fakharian (22). With an increase in the pile number the magnitude of K_s decreases, but it maintains its increasing trend from the middle to the edge of the raft. As the number of piles in the soil increases, the contact area of the pile with the soil decreases. This has led to a decrease in the soil subgrade modulus. Bhartiya et al. (37) mentioned the K_s ratio of the edge to the center of the raft is equal to 5, which indicates the increasing trend in the mentioned study.

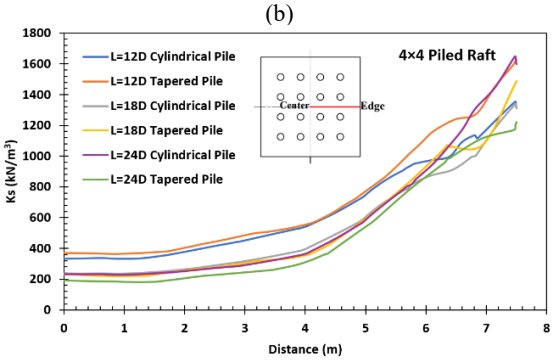
In the present study, this value is equal to 5.8 in its highest state in Figure 13 (for the raft with 9 tapered piles with a length of 24), which is in agreement with the value mentioned by Bhartiya et al. (37) The difference between the present study and the mentioned report is the slope of the graph, in the present study the graph slope is milder than in those cases.



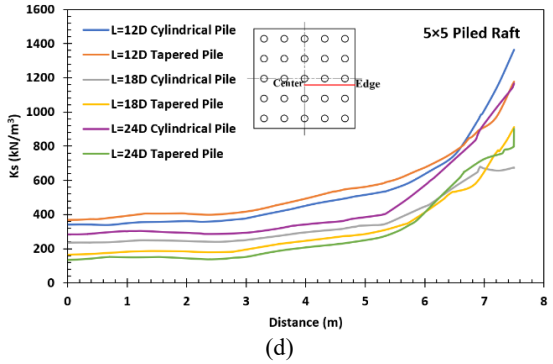
(a)



(b)



(c)



(d)

Figure 13. Soil subgrade modulus for the tapered and cylindrical piled raft in a) 4 piled raft, b) 9 piled raft, c) 16 piled raft and d) 25 piled raft.

6. CONCLUSIONS

In this study, a series of finite element analyses were conducted using PLAXIS 3D software to investigate the effect of tapered angle on piled raft foundations in sandy soil subjected to vertical loading. For this purpose, piles with different numbers and lengths, tapered and cylindrical, were placed at a fixed distance in the center of the raft.

The accuracy of the analyses has been evaluated through multiple validations. The main results obtained are as follows:

- As expected, for both tapered and cylindrical piled rafts, settlements decrease with a rise in length and quantity of piles in a constant loading. Also, it can be seen that the tapered piled rafts have more settlement than cylindrical piled rafts, this becomes clearer when the pile length increases. This increase in settlement is due to stress transmission from the pile to the surrounding soil because shaft bearings generated more widely along the tapered piles.
- By examining the differential settlement of piled rafts, it is evident that rafts with tapered piles have lesser differential settlement compared to cylindrical piles.
- By examining the axial force along cylindrical and tapered piles in raft, it is evident that the tapered piles absorb more of the loading than the cylindrical ones for center piles. Also, it can be seen that tapered piles have larger shaft bearing capacity but smaller end bearing.
- It is obvious from this research that as the quantity of piles increases soil subgrade modulus reduces. It was observed that K_s value increases with a moderated slope from the center of the raft to the edge of the raft.

As it was concluded the settlement of tapered piled rafts is slightly higher than cylindrical piled rafts and it is due to insufficient tip bearing of piles and more contribution of raft. But considering the high shaft bearing capacity, tapered piled rafts promise better performance with floating piles.

7. REFERENCES

1. Robinsky E, Sagar W, Morrison C. Effect of shape and volume on the capacity of model piles in sand. Canadian Geotechnical Journal. 1964;1(4):189-204. https://doi.org/10.1139/t64-015
2. Robinsky E, Morrison C. Sand displacement and compaction around model friction piles. Canadian Geotechnical Journal. 1964;1(2):81-93. https://doi.org/10.1139/t64-002
3. Kodikara JK, Moore ID. Axial response of tapered piles in cohesive frictional ground. Journal of Geotechnical Engineering. 1993;119(4):675-93. https://doi.org/10.1061/(ASCE)0733-9410(1993)119:4(675)
4. Wei J, El Naggar MH. Experimental study of axial behaviour of tapered piles. Canadian Geotechnical Journal. 1998;35(4):641-54. https://doi.org/10.1139/t98-033
5. El Naggar MH, Wei JQ. Axial capacity of tapered piles established from model tests. Canadian Geotechnical Journal. 1999;36(6):1185-94. https://doi.org/10.1139/t99-076
6. Khan MK, El Naggar MH, Elkasabgy M. Compression testing and analysis of drilled concrete tapered piles in cohesive-frictional soil. Canadian Geotechnical Journal. 2008;45(3):377-92. https://doi.org/10.1139/T07-107
7. Tavasoli O, Ghazavi M. Effect of tapered and semi-tapered geometry on the offshore piles driving performance. Ocean Engineering. 2020;201:107147. https://doi.org/10.1016/j.oceaneng.2020.107147
8. Ghazavi M, Lavasan AA, editors. Bearing capacity of tapered and step-tapered piles subjected to axial compressive loading. The 7th international conference on coasts Ports & marine structures, ICOPMAS, Tehran, Iran; 2006.
9. Naggar MHE, Sakr M. Evaluation of axial performance of tapered piles from centrifuge tests. Canadian Geotechnical Journal. 2000;37(6):1295-308. https://doi.org/10.1139/t00-049
10. Shabanpour A, Ghazavi M, El Naggar MH. Behavior of axially loaded tapered pile groups using centrifuge test. Innovative Infrastructure Solutions. 2023;8(6):168. https://doi.org/10.1007/s41062-023-01114-9
11. Nasrollahzadeh E, Hataf N. Experimental and numerical study on the bearing capacity of single and groups of tapered and cylindrical piles in sand. International Journal of Geotechnical Engineering. 2022;16(4):426-37. https://doi.org/10.1080/19386362.2019.1651042
12. Hataf N, Shafaghat A. Numerical comparison of bearing capacity of tapered pile groups using 3D FEM. Geomechanics and Engineering. 2015;9(5):547-67. https://doi.org/10.12989/gae.2015.9.5.547
13. Shafaghat A, Khabbaz H, Fatahi B. Axial and lateral efficiency of tapered pile groups in sand using mathematical and three-dimensional numerical analyses. Journal of Performance of Constructed Facilities. 2022;36(1):04021108. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001680
14. Sayed SM, Bakeer RM. Efficiency formula for pile groups. Journal of geotechnical engineering. 1992;118(2):278-99. https://doi.org/10.1061/(ASCE)0733-9410(1992)118:2(278)
15. Randolph M, editor Design methods for pile group and piled rafts. Proc 13th Int Conf on SFE, 1994; 1994.
16. El-Garhy B, Galil AA, Youssef A-F, Raia MA. Behavior of raft on settlement reducing piles: Experimental model study. Journal of Rock Mechanics and Geotechnical Engineering. 2013;5(5):389-99. https://doi.org/10.1016/j.jrmge.2013.07.005
17. Nguyen DDC, Jo S-B, Kim D-S. Design method of piled-raft foundations under vertical load considering interaction effects. Computers and Geotechnics. 2013;47:16-27. https://doi.org/10.1016/j.compgeo.2012.06.007
18. de Sanctis L, Mandolini A. Bearing capacity of piled rafts on soft clay soils. Journal of Geotechnical and Geoenvironmental Engineering. 2006;132(12):1600-10. https://doi.org/10.1061/(ASCE)1090-0241(2006)132:12(1600)
19. Lee J, Kim Y, Jeong S. Three-dimensional analysis of bearing behavior of piled raft on soft clay. Computers and Geotechnics. 2010;37(1-2):103-14. https://doi.org/10.1016/j.compgeo.2009.07.009
20. Prakoso WA, Kulhawy FH. Contribution to piled raft foundation design. Journal of geotechnical and geoenvironmental engineering. 2001;127(1):17-24. https://doi.org/10.1061/(ASCE)1090-0241(2001)127:1(17)

21. Sinha A, Hanna A. 3D numerical model for piled raft foundation. International Journal of Geomechanics. 2017;17(2):04016055. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000674
22. Khanmohammadi M, Fakharian K. Evaluation of performance of piled-raft foundations on soft clay: A case study. Geomechanics and Engineering. 2018;14(1):43-50. https://doi.org/10.12989/gae.2018.14.1.043
23. Fakharian k, Khanmohammadi M. Evaluation of the effect of geometric characteristics of a piled-raft on its behavior on soft clay under drained conditions. Sharif Journal of Civil Engineering. 2013;29(2):71-6.
24. Fioravante V, Giretti D, Jamiolkowski M, editors. Physical modeling of raft on settlement reducing piles. From research to practice in geotechnical engineering; 2008. https://doi.org/10.1061/40962(325)2
25. Horikoshi K, Matsumoto T, Hashizume Y, Watanabe T, Fukuyama H. Performance of piled raft foundations subjected to static horizontal loads. International Journal of Physical Modelling in Geotechnics. 2003;3(2):37-50. https://doi.org/10.1680/ijpmg.2003.030204
26. Long PD, Vietnam V. Piled raft—a cost-effective foundation method for high-rises. Geotechnical Engineering. 2010;41(1):149. https://doi.org/10.14456/seagj.2010.7
27. Luo R, Yang M, Li W. Normalized settlement of piled raft in homogeneous clay. Computers and Geotechnics. 2018;103:165-78. https://doi.org/10.1016/j.compgeo.2018.07.023
28. Banerjee R, Bandyopadhyay S, Sengupta A, Reddy G. Settlement behaviour of a pile raft subjected to vertical loadings in multilayered soil. Geomechanics and Geoengineering. 2022;17(1):282-96. https://doi.org/10.1080/17486025.2020.1739754
29. Jeong S, Park J, Chang D-w. An approximate numerical analysis of rafts and piled-rafts foundation. Computers and Geotechnics. 2024;168:106108. https://doi.org/10.1016/j.compgeo.2024.106108
30. Horikoshi K, Randolph M. On the definition of raft—soil stiffness ratio for rectangular rafts. Géotechnique. 1997;47(5):1055-61. https://doi.org/10.1680/geot.1997.47.5.1055
31. Khanmohammadi M, Fakharian K. Numerical modelling of pile installation and set-up effects on pile shaft capacity. International Journal of Geotechnical Engineering. 2019. https://doi.org/10.1080/19386362.2017.1368185
32. PLAXIS 3D Reference Manual: bentley; 2023.
33. Bowles JE, Guo Y. Foundation analysis and design: McGraw-hill New York; 1996.
34. Potyondy JG. Skin friction between various soils and construction materials. Geotechnique. 1961;11(4):339-53. https://doi.org/10.1680/geot.1961.11.4.339
35. Praporgescu A, Popa A, editors. Different aspects regarding slope stability improvement using piles. IOP Conference Series: Materials Science and Engineering; 2024: IOP Publishing. https://doi.org/10.1088/1757-899X/1304/1/012009
36. Poulos H. Piled raft foundations: design and applications. Geotechnique. 2001;51(2):95-113. https://doi.org/10.1680/geot.2001.51.2.95
37. Bhartiya P, Chakraborty T, Basu D. Nonlinear subgrade modulus of sandy soils for analysis of piled raft foundations. Computers and Geotechnics. 2020;118:103350. https://doi.org/10.1016/j.compgeo.2019.103350

COPYRIGHTS

©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

رادیه شمع‌ها با انتقال بخشی از بارگذاری از سوی رادیه به خاک، گزینه‌ای اقتصادی‌تر برای تحمل بارهای وارد شده، در مقایسه با دیگر پی‌های شمعی در نظر گرفته می‌شوند. هرچند که افزایش باربری جداری شمع‌های مخروطی در دهه‌های گذشته مورد بررسی قرار گرفته است اما تحقیقات کافی بر عملکرد گروهی این شمع‌ها به خصوص رادیه شمع‌های مخروطی صورت نگرفته است. در پژوهش حاضر رادیه شمعی با شمع‌های مخروطی و استوانه‌ای با طول‌ها و آرایش مختلف در خاک ماسه‌ای مورد بررسی قرار گرفت. مدل شامل یک فونداسیون به ابعاد ۱۵ × ۱۵ × ۰/۷ متر و شمع‌هایی با قطر متوسط ۱ متر است. زاویه مخروطی ۱/۲۰ درجه منجر به قطرهای بالایی به ترتیب ۱/۲۵، ۱/۳۷۶ و ۱/۵۰۲ متر و قطرهای پایینی ۰/۷۵، ۰/۶۲٤ و ۰/٤۹۸ متر برای طول‌های ۱۲، ۱۸ و ۲٤ متر می‌شود. نرم‌افزار اجزاء محدود پلکسیس سه بعدی برای مدل‌سازی انتخاب شد و از مدل‌های موهر –کلمب و الاستیسته‌ی خطی برای مدل رفتاری خاک و بتن استفاده شد. یک سری تحلیل عددی با تغییر در طول و چیدمان برای شمع‌های استوانه‌ای و مخروطی در رادیه‌ی مربعی انجام شد. نتایج نشان داد که شمع‌های مخروطی باربری جداری بیشتری نسبت به شمع‌های استوانه‌ای دارند. همچنین، رادیه شمع‌های مخروطی نشست تفاضلی کمتری نسبت به رادیه شمع‌های استوانه‌ای از خود نشان دادند.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1267-1274

International Journal of Engineering

Journal Homepage: www.ije.ir

Machining of Austenite Steels Using the Method of Preliminary Local Cryogenic Treatment

V. V. Maksarov, V. D. Nguyen*

Empress Catherine II Saint Petersburg Mining University, Russia

P A P E R I N F O

Paper history:
Received 15 May 2025
Received in revised form 23 June 2025
Accepted 15 July 2025

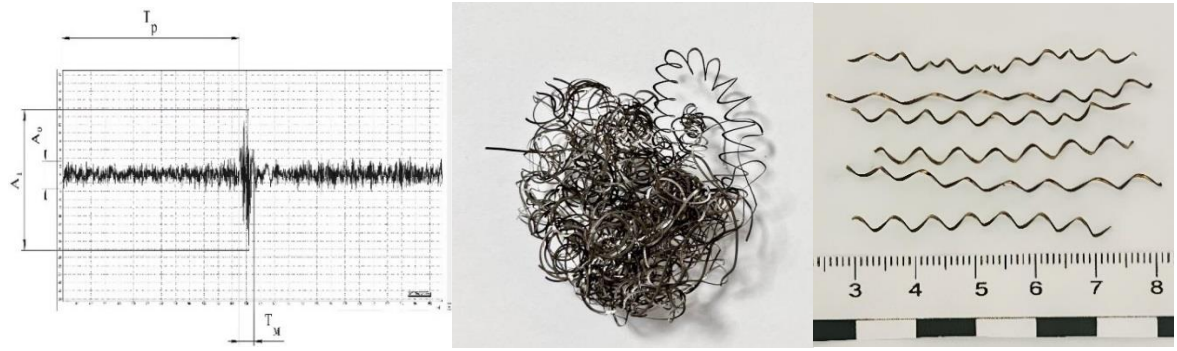
Keywords:
Machinability of Austenitic Steel
Cryogenic Treatment
Structural Martensitic Transformation
Hardening of Surface Layer
Chip Segmentation
Auto Oscillations

A B S T R A C T

The oil and gas industry uses austenite stainless steels because of their high strength, flexibility, and resistance to the impacts of the atmosphere, damp environment, acids, salts, and other harsh chemicals. Problems in machining of such steels are associated with the formation of drain chip, their adhesion to the edge of the tool with the formation of build-up and the emergence of vibrations, which complicates the machining process and reduces its quality. The complexity of machining places greater demands on the tool used, the rigidity of technological equipment mandates the use of coolant, and various ways of chip removal. The scientific novelty consists in the fact that development of a method matches local cryogenic treatment with mechanical processing, ensuring automatic chip breaking and increased productivity. The influence of the method of cryogenic treatment on the structure, properties and machinability of austenitic steels was evaluated on the most common steel grade 08Kh18N10T. The formation of a magnetic martensitic phase with increased hardness at the boundaries of austenitic grains, which causes a change in the mechanism and conditions of chip formation in the local zone, acts as a stress concentrator, resulting in drain chip segmentation and improved machinability. The main findings include an analysis of the method's impact on the dynamic stability of the mechanical processing technological system; reduction of surface roughness by 2 times during mechanical processing with local cryogenic treatment compared to conventional cutting; ensuring stable segmentation and uniform fragmentation of the chip.

***doi**: 10.5829/ije.2026.39.05b.18*

Graphical Abstract



1. INTRODUCTION

Oil is an important strategic resource for the world economy not only because of its role in providing energy, but also because of its extensive use in chemical and other industrial sectors (1-3). Countries with large oil reserves, such as Saudi Arabia, Iran, Russia and the United States, not only influence the global energy

*Corresponding Author Email: nguyenvandao091097@gmail.com
(V. D. Nguyen)

Please cite this article as: Maksarov VV, Nguyen VD. Machining of Austenite Steels Using the Method of Preliminary Local Cryogenic Treatment. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1267-74.

market, but also have significant geopolitical influence by regulating oil production and exports, strongly influencing national policies and strategies. Although many countries are seeking to reduce their dependence on oil and switch to renewable energy sources, oil and natural gas will continue to play an important role in the coming decades, remaining indispensable energy sources in modern life (4).

Therefore, the oil and gas industry plays a key role in providing vital national energy security and defense capabilities, generates large foreign exchange earnings from exports, and makes an important contribution to national budgets. In addition, this industry also creates millions of jobs and contributes to the development of infrastructure and new technologies. The oil and gas industry is an important sector of the economy specializing in the extraction, processing and distribution of oil and natural gas products.

Stainless steels with special corrosion resistance and high strength are widely used in the oil and gas industry (5, 6). The main group of stainless steels, most widely used in industry, is formed by corrosion-resistant austenitic chromium-nickel steels. Typical austenitic stainless steels in addition to iron contain chromium from 16 to 20% and nickel from 9 to 12%, and are known in world practice as steels of type 18-10. Chromium provides the ability of steel to passivate and makes it resistant to aggressive environments, and nickel stabilizes the γ-phase (austenite) in iron, in which the steel is at room temperature, and transforms it into the austenite class. Additionally, steel may contain small amounts of such elements as manganese, silicon and molybdenum (7, 8). In the world practice austenitic stainless steels have unique service properties, are able to retain their mechanical properties at low temperatures, which makes them suitable for use in extreme climatic conditions of Arctic regions and deep sea, as cryogenic materials (9). High stability of characteristics and durability, including at elevated temperatures due to the preservation of the austenitic phase and low tendency to grain growth, allows the steel to be used also for the production of heat-resilient and heat-resistant equipment, in the chemical industry (10-12).

However, machining of austenitic stainless steels faces many difficulties due to high ductility and viscosity of the steel, its tendency to harden during deformation and the formation of drain chip (13). Due to the low thermal conductivity of austenitic steels, there is also an increase in temperature in the cutting zone, which is accompanied by increased adhesion of steel to the tool, its adhesion to the cutting wedge and the occurrence of uneven vibrations (14-16). All this leads to high loads (cutting forces) and accelerated wear and destruction of the tool, as well as to the deterioration of the quality of the surface layer of parts. To eliminate these disadvantages it is necessary to reduce the cutting speed, which leads to a decrease in the productivity of the process. Poor machinability of austenitic stainless steels imposes higher requirements on the used tools and rigidity of the equipment, causes the need to use coolant and chip breaking (17, 18).

At present, a number of methods of machining of difficult-to-machine materials, including austenitic class steels, developed by domestic and foreign researchers are used in research and production practice. Among them can be mentioned machining with vibrations, laser machining, machining with the use of coolants, etc. Each of these methods has both its advantages and limitations. However, so far no universal method has been developed that allows to completely eliminate technological problems arising in the machining of austenitic steels and provides an optimal combination of productivity, quality and economic efficiency (19, 20).

One of the methods used to increase the efficiency of mechanical processing of products from austenitic stainless steels is the method of local cryogenic treatment on their surface during processing. The essence of the method lies in the fact that during thermal exposure of austenitic steel with liquid nitrogen temperature (-196 °C) in the zone of exposure due to a sharp temperature gradient, steel properties change, including phase and structural transition from γ-austenite to α-martensite (21-23), which leads to the formation of structural heterogeneity and increase in micro-hardness in the zone of exposure. During subsequent machining, such areas serve as stress concentrators that ensure chip breaking and reduce the cutting force during machining (24-26).

Based on the above analysis, a method of machining of parts made of austenitic steels by means of preliminary local exposure to ultra-low temperatures was developed and applied in order to ensure automatic chip breaking and elimination of build-up. The advantages of the method of machining on the basis of local cryogenic treatment with automatic chip breaking are the following: reduction of cutting force and friction, increase of cutting tool life, increase of efficiency and quality of machining.

2. METHODOLOGY AND EXPERIMENTAL PART

The essence of the method of mechanical processing of parts of the type “bodies of rotation” with chip breaking on the basis of preliminary cryogenic treatment is that before machining by cutting is carried out application of liquid nitrogen on the machined surface of the workpiece in the form of a straight line, the impact of which leads to changes in the structure and properties of the surface layer (27, 28). The subsequent process of machining by cutting the workpiece is accompanied by periodic intersection of the cutting plane with the created local zone with a high stress-strain state, which provides segmentation of the drained chip, removal of build-up,

which leads to a decrease in the magnitude of self-oscillations of the technological system and a decrease in the roughness values of the machined surfaces of the obtained item (29-31).

To evaluate the effect of the method of cryogenic treatment on the structure and properties of austenitic steels, experiments were conducted on the basis of 08Kh18N10T steel. The experiments were conducted in such a way that the samples of 08Kh18N10T steel were completely immersed for a time starting from 10 s and up to 180 s in a wide-mouth Dewar vessel with liquid nitrogen one by one so that they did not touch each other and did not create an additional layer of vapor jacket. The results of the experiments are presented below:

– 10 s (sample 1) - active boiling of nitrogen is observed, quenching occurs only on the surface and is accompanied by intensive heat dissipation from the central layers; apparently the transformation has not yet started, as the nitrogen boils vigorously even with intensive twisting of the sample to break the steam jacket, the steam creates a thermal shell and does not let the cold to the metal

– 30 s (sample 2) - nitrogen is still boiling vigorously even with vigorous twitching of the sample to break the vapor jacket, but gradually after about 20 s (the process was monitored with a stopwatch), the boiling starts to subside - the metal cools down and a shear, in all probability martensitic process, starts from the surface to the depth

– 60 s (sample 3) - boiling stopped for about ≈ 43 s, the metal cooled down completely - sheared in the volume, all potential martensite was formed and the process stopped

– 180 s (sample 4) - nothing changes compared to 60 s.

Thus, the entire observed transformation process took place within 30 s to 60 s. During such a time interval, only shear martensitic process takes place. Since we are interested in the process of cryogenic treatment on the surface, quenching of the surface in the interval of 30...40 s is sufficient. The formation of martensitic phase, which is magnetic in contrast to austenite, can be confirmed by changes in the magnetic permeability and hardness of the samples.

The study of the magnetic permeability was carried out with a portable magnetometer Magnetoscop 1.070 from FOERSTER. The Magnetoscop 1.070 is a microprocessor-controlled portable magnetometer system with probes for rapid measurement of magnetic flux density as an absolute or differential value (gradient), which is used among others for quality control of stainless steel and confirmation of structural transformations by comparative change of relative magnetic permeability.

The magnetic permeability was measured using a PP-2-5 probe in accordance with ASTM A342/A342M method 4 (measuring within μ from 1.00000 to 2.00000), where μ is the magnetic permeability. Standard sample (instrument calibration μ = 1.05).

Magnetic permeability is determined by the formula:

$$\mu = 1 + 4\pi\chi \quad (1)$$

where π – 3,14; χ – magnetic susceptibility - what we actually measure; ρ – density.

Measurements were made at 4 positions - at opposing points on the diameter. Magnetic permeability before austenitization was 1.097 (residual stresses, etc.), after austenitization 1.085. The measurement results after quenching are summarized in Table 1.

The hardness was determined using a hardness tester HT-1208. The results are summarized in Table 2.

3. RESULTS AND DISCUSSION

Based on the obtained results, the graphs of dependence of magnetic permeability and hardness values on the quenching time were plotted (Figure 1).

The increase in magnetic permeability and hardness during shock quenching can be related both to the formation of new solid ferromagnetic phases of the martensitic type, most probably α- or ε-martensite, and to the preliminary release of excess phases of the carbide type (32, 33). The rate of their formation is determined

TABLE 1. Magnetic permeability values as a function of quenching time

Quenching time t, s	Magnetic permeability, μ				Average magnetic permeability values, μ
	1	2	3	4	
After hardening	1,083	1,085	1,087	1,085	1,085
10	1,088	1,089	1,091	1,093	1,091
30	1,119	1,137	1,141	1,153	1,142
60	1,216	1,214	1,210	1,215	1,214
180	1,218	1,221	1,239	1,221	1,220

TABLE 2. Value of microhardness (HV) of austenitic steel at different distance from the surface depending on the quenching time

Quenching time t, s	Distance from the outer surface, mm				
	0	0,5	1	1,5	2
After hardening	135	132	135	139	146
10	148	146	142	153	155
30	178	171	184	183	182
60	200	195	196	201	203
180	197	187	193	195	194

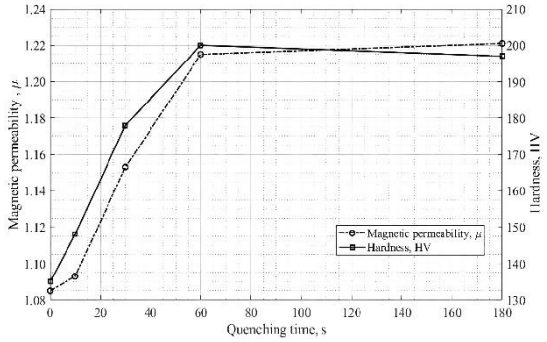


Figure 1. Variation of austenitic steel properties (magnetic permeability, hardness) under cryogenic treatment depending on the quenching time

by the rate of metal freezing and corresponds to the shear martensitic mechanism. The amount of martensite, judging by the increase in magnetic permeability, is very small - less than 5% - magnetic susceptibility changes insignificantly, but since the transformation occurs mainly along the boundaries of austenitic grains, and the formed martensite is brittle, a partial change of the breakaway mechanism to spalling is possible even at very low contents, which leads to a change in the conditions of chip in the local area.

The microstructure of 08Kh18N10T steel was investigated on the optical microscope "Zeiss Axio" SN 38350000483. The results of the studies are shown in Figure 2.

Additionally, the effects of plastic deformation on the structure and properties of stainless steel, quenched for 10 and 180 s, were investigated. Pressure was applied using a device with a ball of 5 mm diameter with a force from 0.2 MPa to 1.0 MPa. The sample was cut along the indentation axis with the maximum deformation force using an electrical discharge cutting unit and further fixed with a two-component self-curing epoxy resin.

Micro-hardness was determined using a Reichert-Jung Meaf-3A microscope with a Micro-Duramat 4000E micro-hardness measuring attachment. ThixometPro program was used for image processing. The results of micro-hardness measurement in the area of the indentation zone are shown in Figure 3.

The hardness in the base material away from the deformation zone is 20.4 HRC, which practically corresponds to the initial hardness of steel after 10 seconds of exposure to nitrogen. In the surface layer under the influence of the ball deformation the hardness increases to 25...25.8 HRC, i.e. by 5.0 units - this is the total increase from deformation and martensitic transformation, however, martensite is not visible in the structure due to its insignificant amount. During deformation of steel after quenching for 60...180 s the hardness in the base is 22.3 HRC, and in the surface layer - from the ball impact - the hardness increased significantly up to 30.6...30.7 HRC. The difference of 8.3...8.4 HRC is the sum of the increase from deformation and martensitic transformation, and the white layer observed in the surface structure is martensite. In the second case, the quenching martensite formed during shock quenching served as a catalyst for the formation of deformation martensite under the same deformation conditions as in the first variant. Similarly, the force impact during cutting causes redistribution of stress-strain states in the stress concentration zones caused by the formation of martensite component at grain boundaries during shock quenching of surface areas, and

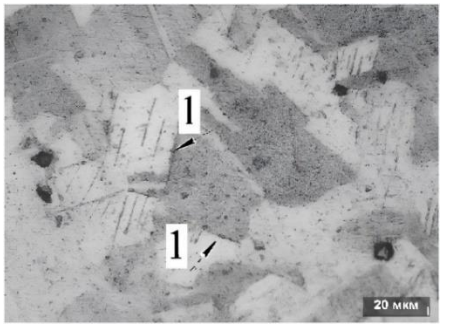
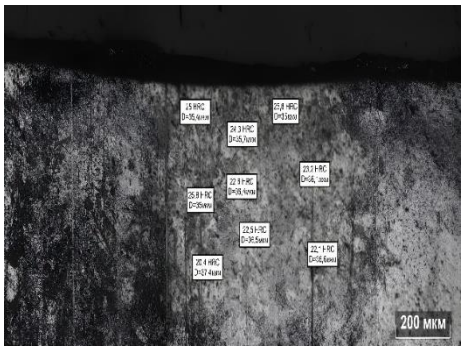
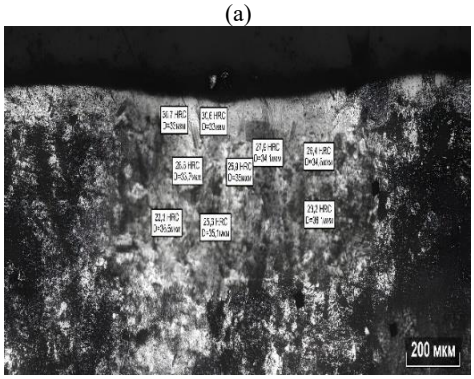


Figure 2. Microstructure of 08Kh18N10T steel after cryogenic treatment, where: 1 - martensite along the austenite grain boundaries



(a)



(b)

Figure 3. Microstructure and micro-hardness of austenitic steel in the zone of deformation by a metal ball after quenching for: a - 10 seconds (x100); b - 60 seconds (x100)

leads to a change in cutting conditions, separation of drain chip into segmented pieces and their removal from the cutting zone (34-36).

Experimental work during machining of 08Kh18N10T steel parts with the use of a two-channel vibration analyzer model VIBXpert EX, provided an opportunity through the measurement of vibration acceleration readings to evaluate the system's auto oscillations, which confirmed the presence of suppression of the oscillatory process during machining of the workpiece on the basis of local cryogenic treatment (33). The result of the experiments are shown in Figure 4.

In the course of experimental studies aimed at measuring vibration accelerations during mechanical processing of parts of the type “bodies of rotation”, it was found that when the tool is introduced into a metastable zone with altered physicomechanical properties, as opposed to the original material, a sharp increase in amplitude A_1 is observed during the time interval T_M. During the period T_M, in the processing zone, the build-up of deposits on the cutting edge of the tool is eliminated, and its free movement occurs without material removal. As a result, the system transitions to a stable state with an amplitude of A_0 and a period of T_P. Based on this, it can be said that the method of local cryogenic treatment has no significant effect on the dynamic stability of the technological system for the mechanical processing of parts of the type “bodies of rotation” made from austenitic steels.

Based on the theory of experimental design for constructing mathematical models that describe the relationship between varying factors and measurable parameters, with the aim of minimizing the number of necessary experiments, it is proposed to apply a central composite design. Technological parameters of the process of machining of parts of the type “bodies of rotation” made from austenitic steels in the course of experimental studies were divided into three groups: variable, static and measurable parameters.

The parameters varied during the experiment were: rotational speed n, min-1; depth of cut t, mm and feed rate S, mm/rev. The values of the workpiece speed n were limited to the values obtained on the lathe model JET GHB 1340A DRO. Static parameters were considered to be those parameters that were not changed during the experimental studies. As a result, it is recommended to apply a method of mechanical processing of parts of the type “bodies of rotation” based on local cryogenic treatment with chip breaking in real production conditions with the following range of technological parameters: local cryogenic treatment time t_{LCT} = 30 – 40 sec.; cutting speed n = 750 – 1250 min^{-1}; the cutting depth t = 0.15 – 0.5 mm; feed S = 0.12 – 0.20 mm/rev.These include: dimensional values of the workpieces to be machined and the geometry of the cutting tool. The measured parameters were the roughness of the machined surface R_a, which was determined on a Surftest SJ-210 profilometer. The conducted studies showed that with the conventional cutting method (without local cryogenic treatment), the surface roughness value R_a is 3.082 μm, while with mechanical machining by cutting based on local cryogenic treatment, the roughness is reduced by more than 2 times to a value of R_a = 1.400 μm. Visual results of experiments on the segmentation and fragmentation of the cutting chip are presented in Figure 5.

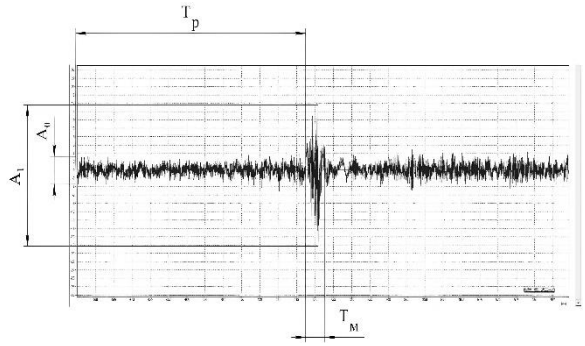
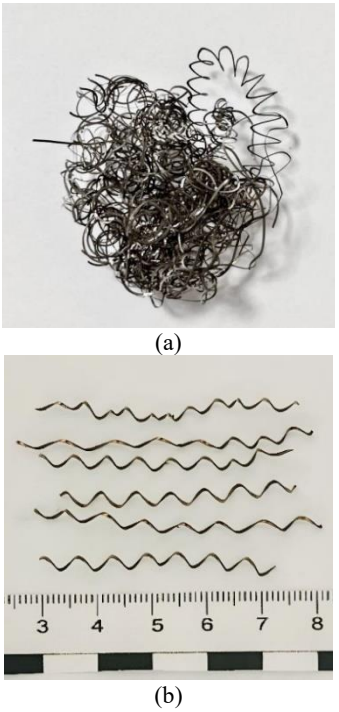


Figure 4. Vibration accelerations during turning of 08Kh18N10T steel previously subjected to local cryogenic treatment: V = 70 m/min; S = 0.12 mm/rev; t = 0.4 mm; n = 750 min^{-1}, where T_P – machining process in the original material; $T_м$ – machining process in the hardened zone with martensitic structure along the austenitic grain boundaries



(a)

(b)

Figure 5. Experimental results: a – chip formed during cutting without local cryogenic treatment; b – chip formed during cutting with local cryogenic treatment

The results of the research showed that in a wide range of variation of varying parameters there is a two-fold decrease in surface roughness when using local cooling, which is due to the reduction of vibration activity of the technological system.

4. CONCLUSIONS

1. When compared to other traditional processing methods, an effective solution for machining rotating-type parts made of austenitic stainless steels is the method of local cryogenic treatment, which ensures stable segmentation, automatic chip breaking and removal of burr formation, and also contributes to increased productivity, product quality (surface roughness of the product), and tool life. Despite some challenges in terms of cost and technology, the prospects for applying this method in the manufacturing industry are very promising, especially given the constantly increasing demands for product quality.
2. Under the influence of ultra-low temperatures on the structure of austenitic steels in the zone of local impact on the machined surface of the work piece martensitic transformation occurs, characterized by a different crystal structure compared to austenite and high hardness. Due to the high stability of austenite in stainless steels, the formation of marten site occurs mainly at the boundaries of austenite grains, which creates a metastable structure that has modified mechanical properties compared to the original structure.
3. Significant (or substantial) change in the structure and physical and mechanical properties of austenitic steel parts in the local area of shock quenching leads to periodic changes in the conditions of metal shear by the cutting wedge during subsequent machining and to changes in the cutting conditions, which is accompanied by a decrease in the vibration activity of the technological system, and thus reduces the values of roughness of machined surfaces and improves the quality of the machined surface.
4. It has been established that during the mechanical processing of parts of the type “bodies of rotation” made from austenitic steels based on local cryogenic treatment, the contact time of the work piece in the local area with liquid nitrogen is 30 – 40 sec., which allows achieving a surface roughness of R_a = 1.4 μm, which is 2 times lower compared to conventional processing methods. The application of the local cryogenic treatment method for the mechanical processing of parts of the type “bodies of rotation” made from austenitic steels ensures stable segmentation of chip into equal segments in the range of 150…200 mm, improves the surface roughness of the finished products in the final operation, and does not significantly affect the dynamic stability of the cutting tool.

5. REFERENCES

1. Zhixin W, Jianjun W, Zhaoming W, Zhengjun H, Chengpeng S, Ruiyin C, et al. Analysis of the current status of global oil, gas, and associated resources exploration in 2023. Petroleum Exploration and Development. 2024;51(6):1465-79. https://doi.org/10.1016/S1876-3804(25)60553-2
2. Lapinskas AA. Influence of mining rent on the efficiency of using natural potential: the paradox of plenty and its Russian specifics. Записки Горного института. 2023(259 (eng)):79-94. 10.31897/PMI.2023.13
3. Zhukov I, Martyushev N, Zyukin D, Azimov A, Karlina A. Modification of hydraulic hammers used in repair of metallurgical units. Metallurgist. 2023;66(11):1644-52. 10.1007/s11015-023-01480-w
4. Burkhanov RN, Lutfullin AA, Raupov IR, Maksyutin AV, Valiullin IV, Farrakhov IM, et al. Localization and involvement in development of residual recoverable reserves of a multilayer oil field. Записки Горного института. 2024(268 (eng)):599-612.
5. Khrustaleva IN, Lyubomudrov SA, Larionova TA, Brovkina YY. Increasing the efficiency of technological preparation for the production of the manufacture components equipment for the mineral resource complex. Записки Горного института. 2021;249:417-26. 10.31897/PMI.2021.3.11
6. Ilyushin Y, Talanov N. Development of methods and models for assessing technical condition of mines and underground structures. International Journal of Engineering Transactions A: Basics. 2025;38(7):1659-66. 10.5829/ije.2025.38.07a.16
7. Dzyura V, Maruschak P, Zakiev I, Sorochak A. Analysis of Inner Surface Roughness Parameters of Loadcarrying and Support Elements of Mechanical Systems. International Journal of Engineering Transactions B: Applications. 2017;30(8):1170-5. 10.5829/ije.2017.30.08b.0
8. Shareef I, Potluri DKR, Horton G. Effect of materials and process parameters on machinability of stainless steels. Manufacturing Letters. 2024;41:696-707. https://doi.org/10.1016/j.mfglet.2024.09.088
9. Ilyushin Y, Nosova V. Development of Mathematical Model for Forecasting the Production Rate. Int J Eng. 2025;38:1749-57.
10. Li W, Zhou B, Xing L, He H, Ni X, Li M, et al. Influence of cutting parameters and tool nose radius on the wear behavior of coated carbide tool when turning austenitic stainless steel. Materials Today Communications. 2023;37:107349. https://doi.org/10.1016/j.mtcomm.2023.107349
11. Voloigzhanina S, Igolkin A, Peregudov A, Baranov I, Martyushev N. Effect of the deformation degree at low temperatures on the phase transformations and properties of metastable austenitic steels. Obrabotka Metallov/Metal Working and Material Science. 2022;24(1):73-86. 10.17212/1994-6309-2022- 24.1-73-86
12. Ermakov S, Ermakov B, Vologzhanina S, Sleptsov O. Investigation of material properties for cryogenic products, produced by additive manufacturing techniques. Metallurgist. 2023;67(5):644-51. 10.1007/s11015-023-01552-x
13. Chandrasekaran K, Marimuthu P, Raja K. Prediction model for CNC turning on AISI316 with single and multilayered cutting tool using Box Behnken Design. International Journal of Engineering Transactions A: Basics. 2013;26(4):621-30. 10.5829/idosi.ije.2013.26.04a.09
14. Sharma G, Dubey V, Sharma AK, Ramkumar J. Study of cutting force in turning of AISI-304 steel using mono and hybrid graphene nanoparticles enriched cutting fluid. Materials Today: Proceedings. 2023. https://doi.org/10.1016/j.matpr.2023.05.190
15. Pang X, Zhang B, Li S, Zeng Y, Liu X, Shen P, et al. Machining performance evaluation and tool wear analysis of dry cutting

austenitic stainless steel with variable-length restricted contact tools. Wear. 2022;504:204423. https://doi.org/10.1016/j.wear.2022.204423

16. Singh P, Dubey V, Sharma AK. Analysis of tool wear and chip morphology under different machining environments in turning operation. Materials Today: Proceedings. 2023. https://doi.org/10.1016/j.matpr.2023
17. Principal T, Barnabas K. Optimization of minimum quantity liquid parameters in turning for the minimization of cutting zone temperature. International Journal of Engineering Transactions C: Aspects. 2012;25(4):327-40. 10.5829/idosi.ije.2012.25.04c.08
18. Kumar A, Sharma R, Kumar S, Verma P. A review on machining performance of AISI 304 steel. Materials Today: Proceedings. 2022;56:2945-51. 10.1016/j.matpr.2021.11.003
19. Major Z, Bodnár L, Merczel DB, Szép J, Lublóy É. Analysis of the heating of steel structures during fire load. 2024. 10.28991/ESJ-2024-08-01-01
20. Bazhin VY, Tran D, Makushin D, Krylov K. Influence of temperature regime of the combined process of casting and rolling of strips from high-alloy aluminium alloys. Non-Ferrous Metals. 2024;2:45-51. 10.17580/nfm.2024.02.07
21. Fricke LV, Basten S, Nguyen HN, Breidenstein B, Kirsch B, Aurich JC, et al. Combined influence of cooling strategies and depth of cut on the deformation-induced martensitic transformation turning AISI 304. Journal of Materials Processing Technology. 2023;312:117861. https: //doi.org/10.1016/j.jmatprotec.2023.117861
22. Fernández-Pisón P, Vishnu A, Vadillo G, Rodríguez-Martínez J. Modeling strain-induced martensitic transformation in austenitic stainless steels subjected to cryogenic temperatures using incremental mean-field homogenization schemes. Mechanics of Materials. 2024;189:104891. https://doi.org/10.1016/j.mechmat.2023.104891
23. Glatt M, Hotz H, Kölsch P, Mukherjee A, Kirsch B, Aurich JC. Predicting the martensite content of metastable austenitic steels after cryogenic turning using machine learning. The International Journal of Advanced Manufacturing Technology. 2021;115(3):749-57. 10.1007/s00170-020-06160-6
24. Fan Z, Gong X, Li B, Yu P, Liu X, Zhou H, et al. The formation of strain-induced martensite and its influence on hydrogen compatibility of metastable austenitic stainless steels: A state-of knowledge review. Journal of Science: Advanced Materials and Devices. 2025;10(1):100842. https://doi.org/10.1016/j.jsamd.2024.100842
25. Deng X, Cheng M, Zhang S, Song H, Taha MA. Residual stresses and martensite transformation in AISI 304 austenitic stainless steel. Materials Research Express. 2018;6(1):016503. 10.1088/2053-1591/aae292
26. Jalili B, Ghafoori H, Jalili P. Investigation of carbon nano-tube (CNT) particles effect on the performance of a refrigeration cycle. International Journal of Material Science Innovations. 2014;2(1):8-17.
27. Chausov M, Maruschak P, Pylypenko A, Shmanenko V, Brezinová J, Lisnichuk M. The effects of cryogenic treatment and impact-oscillatory loading on changes in the mechanical properties and structural condition of stainless steel 12Kh18N10T. Materials Science and Engineering: A. 2024;913:147007. https://doi.org/10.1016/j.msea.2024.147007
28. Vologjanina SA, Igolkin AF, Petkova AP. STUDY OF EFFECT OF LOW TEMPERATURE AND DEFORMATION ON PROPERTIES OF AUSTENITIC STEEL 12KH18N10T/ИССЛЕДОВАНИЕ ВЛИЯНИЯ НИЗКИХ ТЕМПЕРАТУР И ДЕФОРМАЦИЙ НА СВОЙСТВА АУСТЕНИТНОЙ СТАЛИ 12Х18Н10Т. St Petersburg Polytechnic University Journal of Engineering Sciences and Technology. 2019;25(4):83. 10.18721/JEST.25407
29. Keksin A, Sorokopud N, Zakirov N. Peculiarities of abrasive finishing of surfaces of parts made of aluminium alloy of amts grade in magnetic field. International Journal of Engineering, Transactions C: Aspects. 2024;37(06):1098-105. 10.5829/ije.2024.37.06c.06
30. Timofeev DY, Khalimonenko AD, Nacharova MA, editors. Preliminary local thermal impact as a surface quality assurance factor. Materials Science Forum; 2021: Trans Tech Publ. 10.4028/www.scientific.net/MSF.1031.125
31. Kim J-H, Kim E-J, Lee C-M. A study on the heat affected zone and machining characteristics of difficult-to-cut materials in laser and induction assisted machining. Journal of Manufacturing Processes. 2020;57:499-508. 10.1016/j.jmapro.2020.07.013
32. Jovicevic-Klug P, Lipovsek N, Jovicevic-Klug M, Mrak M, Ekar J, Ambrozic B, et al. Assessment of deep cryogenic heat-treatment impact on the microstructure and surface chemistry of austenitic stainless steel. SURFACES AND INTERFACES. 2022;35. https://doi.org/10.1016/j.surfin.2022.102456
33. Al-Tameemi AF, Al-Fayyadh ZT, Al-Mosawe H, Wang Y, Moyet AA. Sustainable Asphalt Mixtures Comprising Steel Slag Filler and SBS-Modified Binder: An Experimental Investigation. Civil Engineering Journal. 2025;11(04). 10.28991/CEJ-2025-011-04-04
34. de Oliveira DA, Brito PP, de Castro Magalhães F, Cangussu VM, Azzi PC, Ardisson JD, et al. Influence of deep rolling on the α'-martensite formation in the subsurface, geometrically necessary dislocations and corrosion resistance of austenitic stainless steel AISI 304. Surface and Coatings Technology. 2024;487:131011. https://doi.org/10.1016/j.surfcoat.2024.131011
35. Zheng C, Liu C, Ren M, Jiang H, Li L. Microstructure and mechanical behavior of an AISI 304 austenitic stainless steel prepared by cold-or cryogenic-rolling and annealing. Materials Science and Engineering: A. 2018;724:260-8. https://doi.org/10.1016/j.msea.2018.03.105
36. Zhao X, Gu K, Zhang M, Weng Z, Pan R, Liu B, et al. Evolution of surface gradient hardening layer of AISI 304 stainless steel induced by cryogenic conventional shot peening. Materials Science and Engineering: A. 2025;925:147919. https://doi.org/10.1016/j.msea.2025.147919

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

صنعت نفت و گاز به دلیل استحکام بالا، انعطاف‌پذیری و مقاومت در برابر اثرات جو، محیط مرطوب، اسیدها، نمک‌ها و سایر مواد شیمیایی خشن، از فولادهای ضد زنگ آستنیتی استفاده می‌کند. مشکلات ماشینکاری چنین فولادهایی با تشکیل براده‌های تخلیه، چسبندگی آنها به لبه ابزار با تشکیل رسوب و ایجاد ارتعاشات مرتبط است که فرآیند ماشینکاری را پیچیده کرده و کیفیت آن را کاهش می‌دهد. پیچیدگی ماشینکاری، ابزار مورد استفاده را با تقاضای بیشتری مواجه می‌کند، صلبیت تجهیزات تکنولوژیکی، استفاده از مایع خنک‌کننده و روش‌های مختلف حذف براده را ایجاب می‌کند. نوآوری علمی این روش در این واقعیت نهفته است که توسعه آن، عملیات برودتی موضعی را با پردازش مکانیکی تطبیق می‌دهد و شکستن خودکار تراشه و افزایش بهره‌وری را تضمین می‌کند. تأثیر روش عملیات برودتی بر ساختار، خواص و قابلیت ماشینکاری فولادهای آستنیتی بر روی رایج‌ترین درجه فولاد Kh18N10T۰۸ ارزیابی شد. تشکیل فاز مارتنزیتی مغناطیسی با افزایش سختی در مرزهای دانه‌های آستنیتی، که باعث تغییر در مکانیسم و شرایط تشکیل تراشه در منطقه محلی می‌شود، به عنوان یک متمرکزکننده تنش عمل می‌کند و منجر به تقسیم‌بندی تراشه تخلیه و بهبود قابلیت ماشینکاری می‌شود. یافته‌های اصلی شامل تجزیه و تحلیل تأثیر روش بر پایداری دینامیکی سیستم تکنولوژیکی پردازش مکانیکی؛ کاهش زبری سطح به میزان ۲ برابر در طول پردازش مکانیکی با عملیات برودتی موضعی در مقایسه با برش معمولی؛ تضمین تقسیم‌بندی پایدار و تکه‌تکه شدن یکنواخت تراشه است.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1275-1292

International Journal of Engineering

Journal Homepage: www.ije.ir

Adaptive Fractional-order Differentiation for Enhanced Image Contrast Utilizing Caputo Masks

Z. Mortezaie*[a], Z. Amiri[b], M. Mortezaee[c]

[a] *Department of Mathematics and Computer Sciences, Hakim Sabzevari University, Sabzevar, Iran*
[b] *Department of Computer Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran*
[c] *Faculty of Computer Engineering, Shahrood University of Technology, Shahrood, Iran*

P A P E R I N F O

Paper history:
Received 27 April 2025
Received 07 June 2025
Accepted 11 July 2025

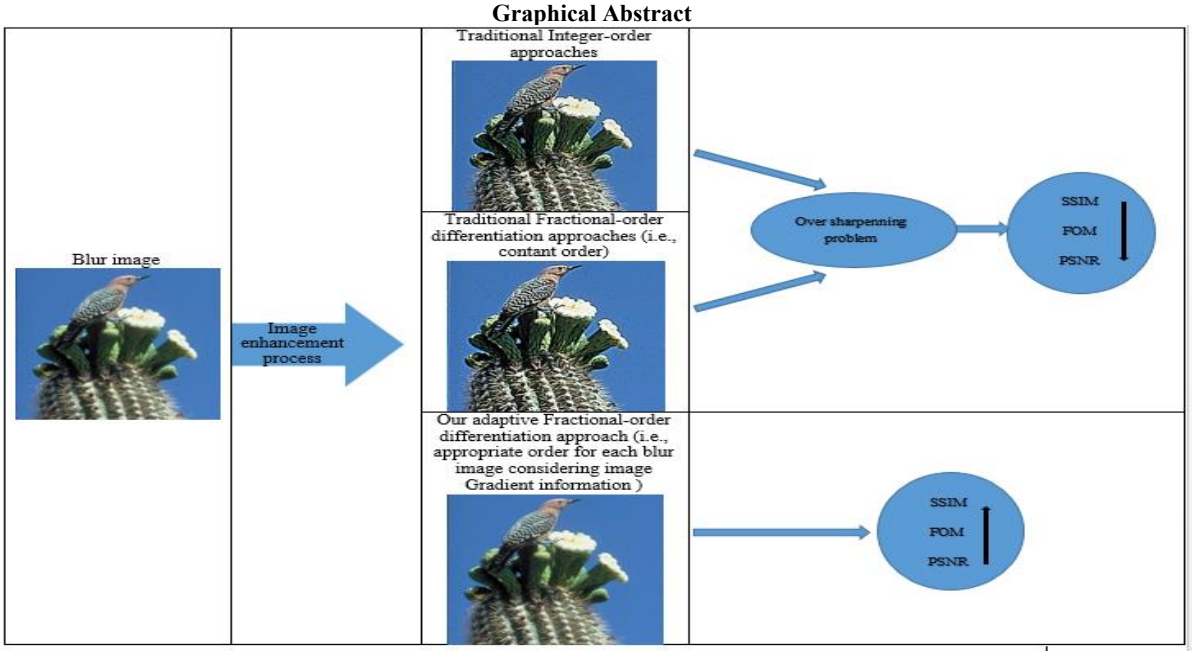
Keywords:
Fractional-order Differentiation
Caputo Fractional Differential Mask
Image Enhancement
Image Processing
Blur Images

A B S T R A C T

Image enhancement remains a cornerstone in digital image processing, aiming to improve visual clarity through various methods. Spatial domain techniques include integer-order and fractional-order differentiation. Although widely used, traditional integer-order differentiation techniques suffer from limitations such as indiscriminate spatial frequency treatment and noise amplification, leading to degraded image quality. This paper proposes an adaptive fractional-order differentiation approach employing Caputo fractional differential masks to selectively enhance image details. This approach uses image gradient information to determine the appropriate fractional order. By dynamically adjusting the fractional order based on specific image requirements, the method achieves superior contrast improvement while preserving fine details and minimizing noise. Experimental results, evaluated using metrics such as Pratt's Figure of Merit (FOM), Structural Similarity Index (SSIM), and Peak Signal-to-Noise Ratio (PSNR), demonstrate that this approach outperforms comparable techniques, highlighting its effectiveness in image enhancement.

***doi**: 10.5829/ije.2026.39.05b.19*

Graphical Abstract



*Corresponding Author Email: zm.mortezaie@gmail.com (Z. Mortezaie)

Please cite this article as: Mortezaie Z, Amiri Z, Mortezaee M. Adaptive Fractional-order Differentiation for Enhanced Image Contrast Utilizing Caputo Masks. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1275-92.

1. INTRODUCTION

Digital image processing relies heavily on image enhancement methods because they are fundamental tools for achieving better visual outcomes. Different image refinement techniques have evolved over time, including basic approaches such as histogram equalization (1) and modern methods based on machine learning (2). The two basic enhancement strategies for digital image processing fall under spatial domain methods and techniques from the frequency domain. Spatial domain methods modify pixel values using local filters or masks that affect small clusters of nearby pixels (3). Image enhancement methods using the frequency domain analyze the Fourier transform of images to modify their frequency characteristics (4).

Multiple spatial domain differentiation techniques are used to highlight significant image features through unsharp masking (5), the Laplacian filter (6), the Canny edge detector (7), and anisotropic diffusion (8). First- and second-order derivative-based operations are traditional techniques used for edge detection and sharpening tasks. Integer-order methods have significant performance constraints as their primary limitation. These methodologies fail to discriminate spatial frequencies because they treat all equally, leading to detail loss (9). High-pass filters created with integer-order differentiation processes tend to produce noise increases that degrade image quality (10). Integer-order enhancement techniques demonstrate an inflexible nature that restricts their ability to support diverse image features and enhancement needs (11).

Fractional-order differentiation methods have proven to be viable solutions to address these problems. The use of fractional derivatives expands the pixel area from which information is gathered, leading to a better understanding of extended image dependencies [9]. Fractional-order differentiation offers two important advantages. The first advantage stems from its continuous differentiation order palette, which interlinks integer-order operator functions (12). The computation process using fractional derivatives preserves historical image information, thus maintaining contextual details better (13). Fractional differentiation methods demonstrate impressive capabilities in multiple enhancement operations, including contrast improvement (14), edge detection (15), and texture analysis (14). The fractional-order control enables image enhancement that improves contrast precision while maintaining all fine details. These derivatives provide better performance for edge detection in noisy conditions (15). These derivatives offer significant benefits in texture analysis and segmentation while performing edge detection because they efficiently acquire low-frequency information and high-frequency details (16).

The application of fractional-order differentiation extends numerous benefits, but multiple obstacles remain when implementing it for image enhancement. Strong research demands better theoretical models that explain how fractional derivatives interact with image elements. Scientists work toward improving computational efficiency in fractional derivative implementation processes (17). The research community requires more study to find optimal fractional orders that suit different image enhancement goals (18).

Considering the importance of fractional orders in image enhancement, this paper presents an adaptive fractional-order differentiation method that adjusts the fractional order of the Caputo fractional differential mask (19) according to different image enhancement requirements for various image types.

The contributions of this paper are summarized as follows:

a) This paper fills the research gap through an adaptive fractional-order differentiation approach to enhance images with different contrast levels.
b) The adaptive Caputo fractional-order differential mask is used to enhance various images with different enhancement requirements.
c) To achieve this goal, the gradient information of each image is used to adaptively select the appropriate fractional-order.
d) The performance of our proposed approach is compared with some existing image enhancement methods using FOM, SSIM, and PSNR metrics.

The experimental results show the superiority of the proposed adaptive fractional-order differentiation method compared to other existing image enhancement approaches.

This paper follows the below structure: Section 2 provides a brief review of current image enhancement techniques using fractional-order differentiation approaches. Section 3 explains the Caputo fractional derivative while presenting our adaptive fractional-order differentiation strategy. The manuscript concludes with its findings and conclusion in Sections 4 and 5, respectively.

2. RELATED WORKS

Fractional calculus is used in image processing tasks such as image edge detection (20-29), image enhancement, and de-noising (9, 16, 29-33). This section briefly describes some existing approaches in this field.

Wang et al. (29) proposed an enhancement method that combines fractional-order phase stretch transform (30) with relative total variation (31) to enhance both low and medium image frequencies. The algorithm removes initial image noise through the implementation of low-pass filtering techniques. Edge extraction from the

filtered image is achieved through applications of the fractional-order phase stretch transform. Specific edges obtained from this process are reunited with the original image through relative total variation processing to preserve important structures and improve image quality.

The fractional differential-inverse-distance-weighted enhancement algorithm, presented Huang et al. (32), serves as a depth image enhancement algorithm. This method utilizes fractional differentiation to improve the detection of texture and edge qualities for enhanced feature detection. The refinement process and noise management during enhancement occur through the combination of convolution templates with fractional differential operators, which protect important structural details. The algorithm applies a fractional differential-inverse-distance-weighting method to modify pixel values as a function of their relative position to important features and edges.

Ruiyin and Bo (33) demonstrated an image enhancement method by merging fractional calculus with the Retinex theory (34). Fractional differentiation serves to improve image texture before obtaining the final input by preprocessing the initial image. After image preprocessing through a single-scale Retinex algorithm, the method divides images into illumination and reflection elements. The process continues with fractional integration, which performs noise reduction on the reflected components. Contrast Limited Adaptive Histogram Equalization (CLAHE) (35) serves as the last step to boost image contrast levels.

Huang et al. (9) demonstrated how fractional calculus enhanced images while performing de-noising applications. The proposed methodology consists of two concurrent tasks using fractional-order derivative properties to achieve image enhancement and image de-noising capabilities. The method uses fractional-order differential operators as an enhancement mechanism for images. Researchers apply experimental variations to differentiation orders to improve image clarity and enhance edge visibility during the enhancement process. The method implements fractional-order integral operators to perform image de-noising functions.

Jalab and Ibrahim (16) presented a texture enhancement methodology in which incorporated the fractional-order Savitzky-Golay differentiator introduced by Chen et al. (36). Researchers used this technique to upgrade digital images by improving texture visibility without changing the essential image characteristics. The Savitzky-Golay filter receives a generalized form through fractional calculus, employing Srivastava-Owa fractional operators (37). The method applies sliding weight window operations to calculate generalized fractional-order derivatives of images. Applying the fractional-order Savitzky-Golay operator allows the method to improve edges and prominent features present in textures.

Kaur et al. (38) combines fractional Fourier transforms with Riesz fractional derivatives (39) to detect more accurate image edges before using these edges to enhance image quality. This image analysis method utilizes the fractional Fourier transform to process images by preserving spatial and frequency data points simultaneously. Image edges are extracted through the application of Riesz fractional derivatives following image transformation. The quality improvement process reintegrates edges derived from Riesz fractional derivatives into the original image.

Mortazavi et al. (40) described how fractional derivatives were incorporated into sparse representation methods to achieve image resolution enhancement. The method utilizes fractional derivative mathematical properties to make super-resolved images more detailed by better preserving image information. This method enhances images by extracting dictionaries from low-resolution images to create high-resolution images. Fractional derivatives integrated into the proposed method help improve the representation of image characteristics.

Azarang and Ghassemian (41) integrated panchromatic (PAN) and multispectral (MS) images via fractional-order differentiation methods for image fusion. The purpose of this method is to improve both quality and details in the merged images for multiple remote sensing applications. The procedure starts by loading entire PAN and MS image collections from the dataset. As a first step in image fusion, the data requires preprocessing through both down-sampling and normalization techniques. The main concept of this method requires producing fractional-order differentiations. This operation generates a superimposed filter which extracts necessary image features from the original data. A primitive detail map generation procedure occurs for each spectral band of MS images through this method. The map indicates essential image characteristics vital for analytic purposes. The superimposed mask, which has been generated, applies onto the primitive detail maps to perform additional refinement. The process produces one fused image which unites PAN image high-resolution details with the spectral information of MS images.

Motłoch et al. (13) developed an approach to upgrade fused images by implementing fractional derivative techniques. Fractional-order derivatives are used to increase specific details during image fusion processes. This approach utilizes eight-directional non-integer order filters which extract directional attributes from data. The method consolidates PAN and MS image data through fractional-order derivative masks which extract intricate characteristics from each image type. The computational efficiency increases through Fast Fourier Transform (FFT) that executes fractional-order derivative approximation in the frequency domain, leading to effective image characteristic manipulation.

Nchama et al. (42) presented an algorithm that effectively improves image contrast through fractional calculus technique by applying the Caputo-Fabrizio fractional differential mask found by Caputo and Fabrizio (43). The technique provides regulated image contrast improvement because users can modify the fractional differential order as needed to personalize the enhancement operation. The proposed mask delivered the best contrast results after being used with a fractional order value of 10^{-10}.

Most of the examined methods fail to recognize how fractional orders affect image enhancement results. This paper fills the research gap through an adaptive fractional-order differentiation approach to enhance images with different contrast levels. The developed image enhancement method uses Caputo fractional differential masks whose fractional order depends on the necessary degree of contrast adjustment for diverse image categories (10, 44-54).

3. PROPOSED METHOD

3. 1. Fractional-Order Differentiation Fractional-order differentiation offers unique benefits in image enhancement through its ability to provide a wide spectrum of differentiation orders between integers. Image processing algorithms become more flexible through the use of fractional-order differentiation systems because this property surpasses traditional integer-order methods. The use of fractional-order differentiation provides multiple stable approaches applicable to image enhancement. Multiple acceptable frameworks emerged from mathematicians seeking to utilize the capabilities of this method despite its undefined standardized definition. Image processing utilizes three main definitions for fractional differentiation: Grunwald-Letnikov (G-L) (55), Riemann-Liouville (R-L) (56), and Caputo (19). The fundamental concept of integer-order differentiation supports the G-L definition to enhance edge detection and feature extraction processes. The R-L definition performs well at image pattern detection via its integer-order integration foundation to analyze fragile image features. The Caputo definition, derived from integer-order integration, finds practical uses in real-world modeling and can serve as an effective tool for adaptive image enhancement techniques.

Although fractional-order differentiation is increasingly used in image processing, the vast majority of related studies focus on the G-L and R-L definitions. In contrast, the Caputo derivative has received relatively limited attention in image processing applications. Therefore, in our proposed adaptive fractional-order differentiation approach, the Caputo derivative is employed to demonstrate its potential in image enhancement. Furthermore, the proposed adaptive strategy can be applied to other fractional definitions, helping to expand the scope of adaptive fractional-order techniques in image processing.

3. 2. Caputo Definition The Caputo fractional differentiation of the function v with respect to t of order α is defined as Equation 1 (19):

$$D^{\alpha}v(t) = \frac{1}{\Gamma(1-\alpha)}\int_0^t \frac{v'(\mathcal{T})}{(t-\mathcal{T})^{\alpha}}\, d\mathcal{T}, 0 < \alpha \leq 1. \quad (1)$$

The finite difference approximations for the Caputo fractional derivative are provided as follows. The time interval $[0,T]$ is divided into M equal subintervals, resulting in a step size $\delta t = \frac{T}{M}$. The time points is denoted as $t_k = k\delta t$ where $k = 0,1,\ldots,M$ and let $v_k = v(t_k)$. The Caputo derivative in Equation 1 at $t = t_k$ can be approximated as Equation 2:

$$D^{\alpha}v(t_k) = \frac{1}{\Gamma(1-\alpha)}\sum_{j=1}^{k}\int_{t_{j-1}}^{t_j}(t_k-\mathcal{T})^{-\alpha}\frac{\partial v(\mathcal{T})}{\partial \mathcal{T}}\, d\mathcal{T} \approx \frac{1}{\Gamma(1-\alpha)}\sum_{j=0}^{k}\frac{v_j-v_{j-1}}{\delta t}\int_{t_{j-1}}^{t_j}(t_k-\mathcal{T})^{-\alpha}\, d\mathcal{T}. \quad (2)$$

The integral shown in Equation 2 is computed as Equation 3:

$$\int_{t_{j-1}}^{t_j}(t_k-\mathcal{T})^{-\alpha}\, d\mathcal{T} = \frac{\delta t^{(1-\alpha)}}{1-\alpha}\left[(k-j+1)^{(1-\alpha)} - (k-j)^{(1-\alpha)}\right], \quad (3)$$

Considering $\omega_{\alpha} = \frac{(\delta t)^{-\alpha}}{\Gamma(2-\alpha)}$ and $\varpi_{\alpha,j} = (j+1)^{1-\alpha} - j^{1-\alpha}$, the discretized form of the Caputo fractional derivative is as Equation 4:

$$D_t^{\alpha}v_k \approx \omega_{\alpha}\left[v_k - v_{k-1} + \sum_{j=1}^{k-1}\varpi_{\alpha,j}\left(v_{k-j} - v_{k-j-1}\right)\right], k = 1,\ldots,M. \quad (4)$$

Note that, in image processing, δt should be set to 1, as the smallest scale is one pixel. Differentiating the continuous function $v(t)$ with respect to t allows us to express the difference formula for the fractional differentiation of $v(t)$ as shown in Equation 4.

In this context, j is obtained based on the function and indicates the j-th neighborhood value. For easier calculations, Equation 4 is extended to two dimensions, and for $j=1$ Equations 5 and 6 are obtained as follows:

$$D_x^{\alpha}v(x,y) = \frac{1}{\Gamma(2-\alpha)}\left[v(x,y) + (2^{1-\alpha}-2)v(x-1,y) - \left(2^{(1-\alpha)}-1\right)v(x-2,y)\right] \quad (5)$$

$$D_y^{\alpha}v(x,y) = \frac{1}{\Gamma(2-\alpha)}\left[v(x,y) + (2^{1-\alpha}-2)v(x,y-1) - \left(2^{(1-\alpha)}-1\right)v(x,y-2)\right] \quad (6)$$

The implementation of anti-rotation abilities in image processing requires developers to integrate several directional masks thoughtfully because this process presents multiple challenges. These special fractional differentiation masks operate on both horizontal (x-axis)

and vertical (y-axis) directions, enabling systematic detection and compensation of input image rotational changes.

Let $k_1 = \frac{1}{\Gamma(2-\alpha)}$, $k_2 = \frac{(2^{(1-\alpha)}-2)}{\Gamma(2-\alpha)}$, and $k_3 = \frac{(2^{(1-\alpha)}-1)}{\Gamma(2-\alpha)}$, denote the horizontal, vertical, and cross-directional fractional differentiation masks, respectively. The eight distinct masks represented in Figure 1 correspond to various directional orientations: the positive x-direction (Figure 1(a)), positive y-direction (Figure 1(b)), negative x-direction (Figure 1(c)), negative y-direction (Figure 1(d)), left upper diagonal (Figure 1(e)), left lower diagonal (Figure 1(f)), right upper diagonal (Figure 1(g)), and right lower diagonal (Figure 1(h)).

Masks form the foundation for developing the end product of an anti-rotation mask. A configuration as shown in Figure 2 emerges after performing partial fractional differentiation within each of the eight separate orientations. To complete the procedure we need weight normalization for each element of the combined mask by performing a total sum division of all its components. The normalization method guarantees that every directional component creates an identical influence on the end results. By employing this combined approach, the system obtains simultaneous multi-perspective imaging evaluation, which improves its capacity to detect subtle variations and retain critical image details.

3. 3. Proposed Image Enhancement Approach

The implementation of the Caputo fractional differential mask depicted in Figure 2 improves image quality as it convolves with low contrast images. Image quality is significantly enhanced only through selecting the proper value of the fractional order parameter (α). The graphical representation in Figure 3 shows how contrasting image elements change with different fractional order inputs.

0	k_1	0
0	k_2	0
0	k_3	0

(a) Positive x-coordinate

0	0	0
k_1	k_2	k_3
0	0	0

(b) Positive y-coordinate

0	k_3	0
0	k_2	0
0	k_1	0

(c) Negative x-coordinate

0	0	0
k_3	k_2	k_1
0	0	0

(d) Negative y-coordinate

k_3	0	0
0	k_2	0
0	0	k_1

(e) Left upper diagonal

0	0	k_1
0	k_2	0
k_3	0	0

(f) Left lower diagonal

0	0	k_3
0	k_2	0
k_1	0	0

(g) Right upper diagonal

k_1	0	0
0	k_2	0
0	0	k_3

(h) Right lower diagonal

Figure 1. The eight distinct masks in different directional orientations

This figure contains an original input image in part (a) which looks blurry, followed by images (b) through (d) that depict the processed data after using the Caputo fractional differential mask together with three different fractional order procedures, including low, high, and optimal settings. The specified visualizations display how fractional order variations produce considerable effects on image clarity and their contrast capacity. Using a low fractional order imposes more blurriness on the input image, while a high fractional order causes an over-sharpening issue.

Based on the aforementioned points, in this section, we propose a novel adaptive fractional-order differentiation technique designed to enhance image contrast. The proposed method adaptively adjusts the fractional order of the Caputo fractional differential mask in accordance with the required level of enhancement for blur images. In our proposed approach, using Equation 7, the input image is convolved separately with 20 different Caputo fractional differential masks considering 20 different fractional orders from 0.01 to 1, at intervals of 0.5, resulting in 20 different convolved images. Then, the gradient variation of each convolved image compared with the input image is computed using Equations 8 and 9:

$$C(x,y) = I(x,y) * M, \tag{7}$$

$$G(x,y) = I(x,y) - C(x,y), \tag{8}$$

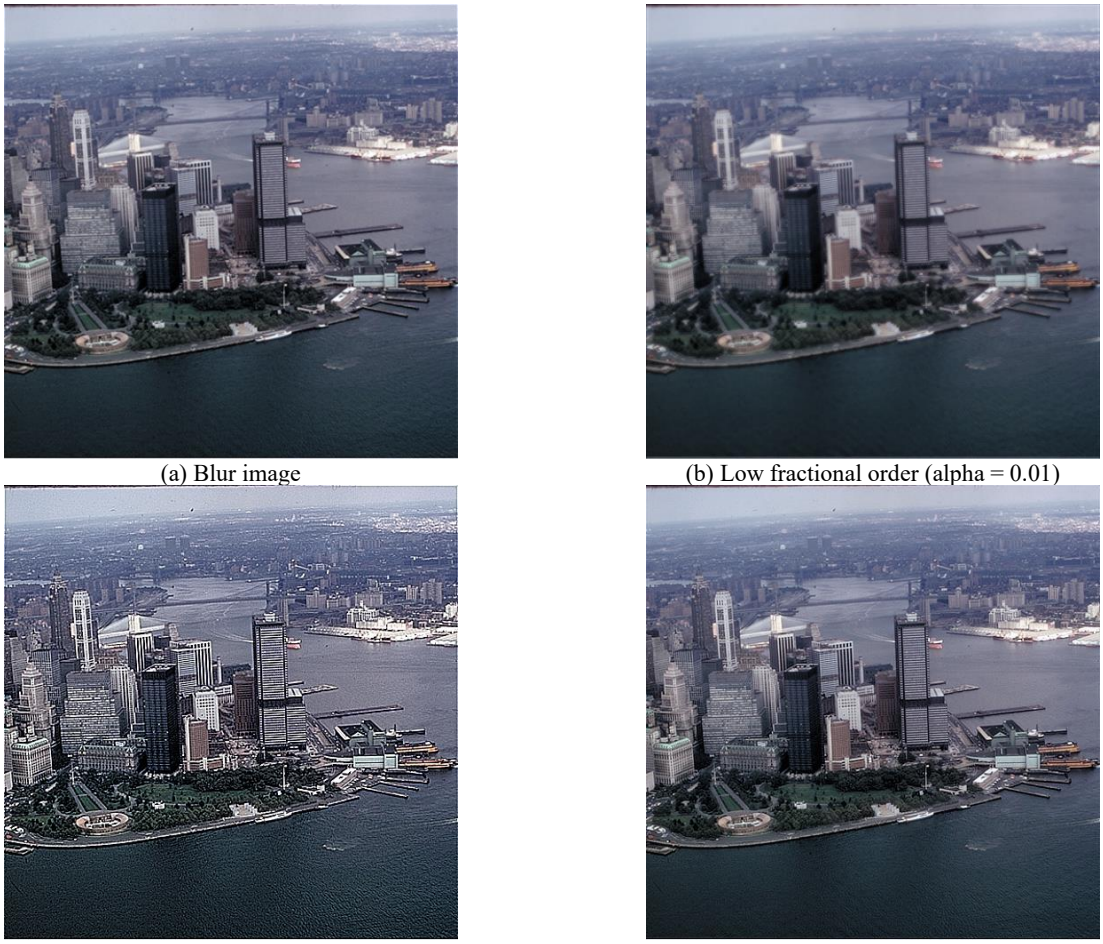
$$\|G\|_n = \left(\sum_{x,y}|G(x,y)|^n\right)^{\frac{1}{n}}, \tag{9}$$

where M and $*$ respectively denote the Caputo fractional differential mask shown in Figure 2 and the convolution operator; $I(x,y)$, $C(x,y)$, and $G(x,y)$ respectively represents an input image, convolved image, and gradient image. Also, $|\,.\,|$ and $\|\,.\,\|_n$ symbols are the absolute value and l_n norm respectively.

Our analysis indicates that weak edges and details can be more effective than strong ones in determining the optimal fractional order. Hence, the proposed method aims to enhance the influence of weak edges and details compared to stronger ones during the fractional order selection process. We determine the $l_{0.8}$ norm of gradient information through the input exponent value n equal to 0.8 in Equation 9. This approach allows us to balance the contribution of both strong and weak details in determining the optimal fractional order.

k_3	0	k_3	0	k_3
0	k_2	k_2	k_2	0
k_3	k_2	$8 \times k_1$	k_2	k_3
0	k_2	k_2	k_2	0
k_3	0	k_3	0	k_3

Figure 2. The final mask constructed by combining eight distinct masks shown in Figure 1



(a) Blur image (b) Low fractional order (alpha = 0.01)

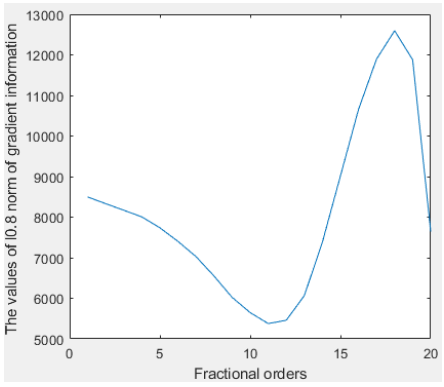
(c) High fractional order (alpha = 0.9) (d) Appropriate fractional order (alpha = 0.7)

Figure 3. The results of convolving the input image with Caputo fractional differential mask using different fractional order

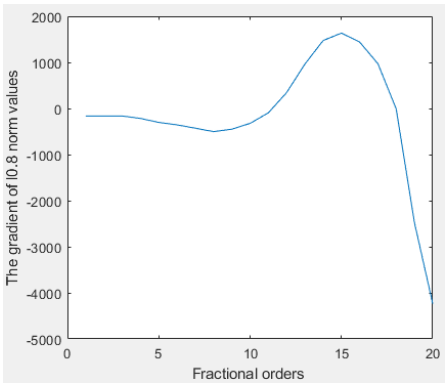
The analysis of the $l_{0.8}$ norm of gradient information for 20 fractional orders on one instance image is displayed in Figure 4. The values of $l_{0.8}$ norm of gradient information exist in Figure 4(a) that displays results across the 20 evaluated fractional orders. The gradient of these 20 obtained $l_{0.8}$ norm values appears in Figure 4(b) while using similar axis representations.

The initial 11 fractional orders appear to cause image blurring based on Figures 4(a) and 4(b) as they decrease both $l_{0.8}$ norm values and their gradients according to the data. The application of such orders leads to deterioration of image details while producing overall image blurring. The 12th through 18th fractional orders generate rising trends for both $l_{0.8}$ norm values and their gradients. The image enhancement process of these intermediate orders works better because they enhance contrast simultaneously with detail retention. The $l_{0.8}$ norm values and their gradients both decrease when used with the 19th and 20th fractional orders thus potentially causing deterioration of image quality. Elevated fractional orders past specified values can cause over-sharpening problems in digital images.

Our proposed approach will choose the best fractional order from the 12th to 18th fractional orders because of these discovered results. The targeted fractional order interval includes various values which without significant impact on image quality alongside those that trigger undesirable sharpening effects. Hence, we use the gradient of $l_{0.8}$ norm values to select the optimal fractional order which results in the maximum increase of $l_{0.8}$ norm values in our proposed adaptive selection method. We use this fractional order for enhancing contrast because it delivers optimal results without creating over-sharpening artifacts. Note that these observations and analysis pertain to an instance input image. Hence, in our proposed approach, a similar manner is generalized across various input images. This approach allows our method to adaptively adjust the degree of enhancement based on the inherent characteristics of each image, potentially leading to more consistent and desirable results across various types of visual content.



(a) The values of $l_{0.8}$ norm of gradient information for 20 different fractional orders



(b) The gradient of 20 obtained $l_{0.8}$ norm values

Figure 4. The values of $l_{0.8}$ norm of gradient information for different fractional orders and their gradient

4. EXPRIMENTAL RESULTS

We subjectively and objectively evaluate the effectiveness of our adaptive fractional-order differentiation method in this section. Three objective metrics serve this analysis: Pratt's Figure of Merit (FOM) (57, 58), the Structural Similarity Index Metric (SSIM) (59, 60), and the Peak Signal-to-Noise Ratio (PSNR) (61). The outcomes for both FOM and SSIM are expressed as values ranging from 0 to 1. Higher values in these three metrics indicate greater similarity between the two images.

The experiment relies on the CSIQ database (62), which contains 30 reference images with different degrees of Gaussian blur. The reference image collection comprises five progressively blurred versions, classified from level 1 to level 5. Blur intensity increases as the levels progress from 1 to 5, while high-frequency elements experience progressively worse degradation, particularly involving edge components. Images become unreadable in terms of edge definition at levels 4 and 5 because edge degradation reaches high levels. Our proposed approach shows reduced effectiveness when combating blurriness at higher reference image levels because it primarily focuses on edge extraction tasks. Therefore, our proposed adaptive fractional-order differentiation method was applied to all 90 potential blurred images at levels 1, 2, and 3 for evaluation.

We compare our approach with results derived from the Caputo fractional differential mask, using three distinct fractional order values: 0.01, 0.6, and 0.9. The mean values and variances for FOM, SSIM, and PSNR obtained from these approaches are summarized in Table 1. As indicated in the table, significant differences are observed in the mean values of FOM, SSIM, and PSNR among the various approaches. Specifically, the imposed blurriness caused by a low fractional order (0.01) leads to insufficient detail, while excessive sharpening with a high fractional order (0.9) and the use of a constant fractional order value (0.6) across all input images—without adapting to specific enhancement needs—results in reduced mean values for FOM, SSIM, and PSNR. It is evident that the mean FOM, mean SSIM, and mean PSNR for our adaptive fractional-order differentiation approach outperform those of the comparative approaches.

In Table 2, we compare our proposed approach with image enhancement approaches proposed by Huang et al. (9) and Nchama et al. (42), which are based on fractional calculus, as well as with integer-order differentiation approaches, namely, classic un-sharp masking (5) and the Laplacian filter (6). The image enhancement techniques described by Huang et al. (9) and Nchama et al. (42) used constant fractional order values for their fractional differential masks without considering specific enhancement requirements. Classic un-sharp masking and the Laplacian filter operate on image enhancement without adapting to the particular level of enhancement needed. In classic un-sharp masking, first, the edge information of the input image is extracted by subtracting a blurred version of the image from the original image.

TABLE 1. Performance Comparison of the Proposed Method with the Caputo fractional differential mask using different alpha

	Blur images	Our adaptive fractional-order differentiation approach	Caputo fractional differential mask using different alpha		
			alpha = 0.01	alpha = 0.6	alpha = 0.9
FOM Mean	0.8742	**0.9010**	0.7255	0.8447	0.8839
SSIM Mean	0.9346	**0.9472**	0.8449	0.9362	0.8051
PSNR Mean	32.6389	**32.4978**	26.9026	31.6691	23.7734

TABLE 2. Comparison of the Proposed Method with Fractional and Integer-Order Image Enhancement Techniques.

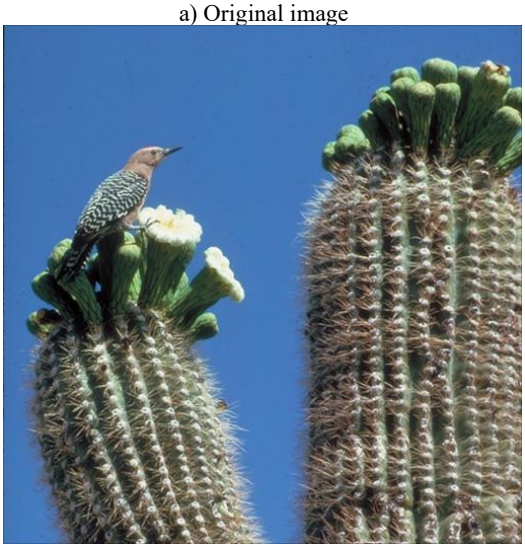
	Blur images	Our proposed approach	The method proposed in (9)		The method proposed in (42)		Un-sharp masking (5)		Laplacian filter (6)	
FOM Mean	0.8742	**0.9010**	alpha = 1.5	0.8257	alpha = 0.5	0.8630	landa = 0.1	0.8619	alpha = 1	0.8501
SSIM Mean	0.9346	**0.9472**		0.8488		0.7353		0.7117		0.9192
PSNR Mean	32.6389	**32.4978**		27.6965		20.4225		20.6360		30.7600

Then, a parameter named gain factor (lambda) is used to scale the extracted edges before adding them to the original image. Additionally, the Laplacian filter highlights image edge information using a second-order derivative, with a scaling coefficient (alpha) controlling the level of edge enhancement. The performance of both methods depends on the proper tuning of these coefficients to avoid over-sharpening and noise amplification. Considering these points, our proposed method achieves superior mean FOM, along with higher mean SSIM and mean PSNR values, when compared to other methods based on Table 2.

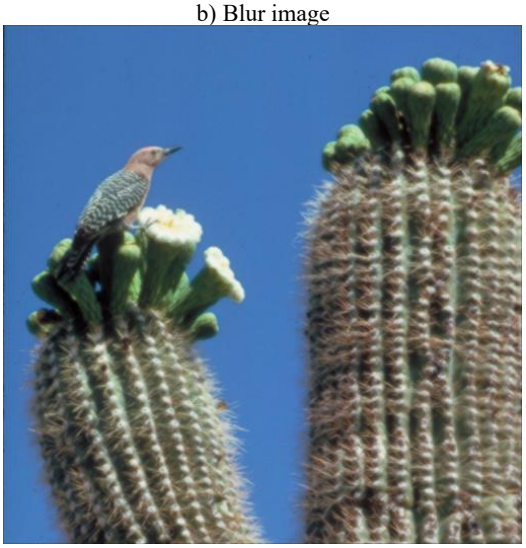
Figures 5 to 8 present four instances demonstrating the effectiveness of our proposed approach when compared to results obtained from the Caputo fractional differential mask, utilizing three different fractional order values: 0.01, 0.6, and 0.9, as well as classic un-sharp masking (5), the Laplacian filter (6), and the methods

a) Original image

b) Blur image



c) Our adaptive fractional-order differentiation approach



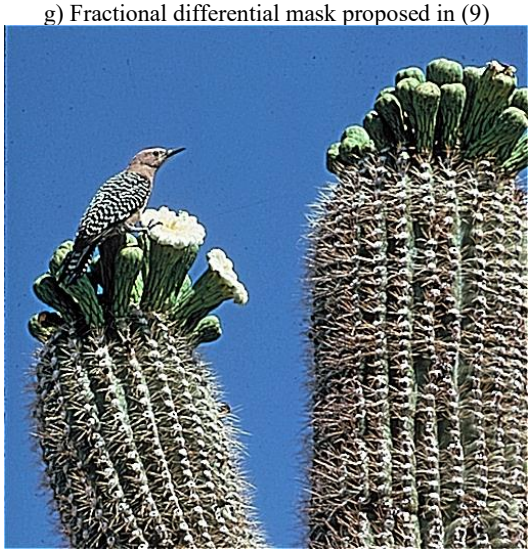
d) Caputo fractional differential mask using alpha = 0.01

e) Caputo fractional differential mask using alpha = 0.6

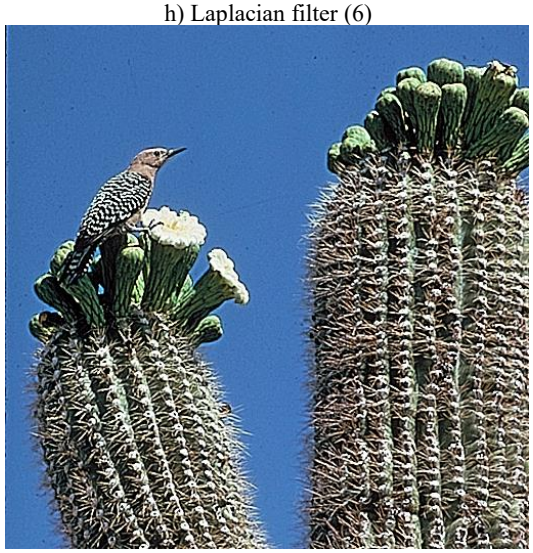
f) Caputo fractional differential mask using alpha = 0.9

g) Fractional differential mask proposed in (9)

h) Laplacian filter (6)



i) Fractional differential mask proposed in (42)

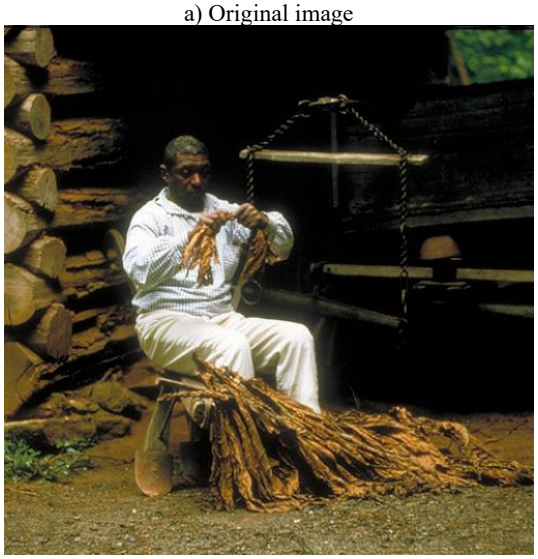


j) Un-sharp masking (5)

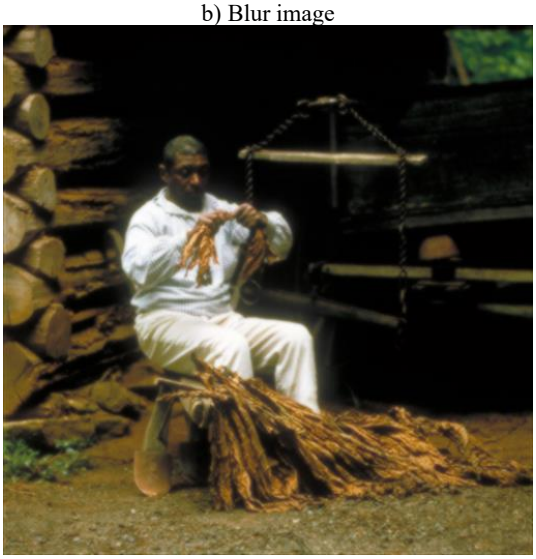
Figure 5. Results obtained from our adaptive fractional-order differentiation approach, the Caputo fractional differential mask using three distinct values for fractional order as 0.01, 0.6, and 0.9, classic un-sharp masking and the laplacian filter, as well as the methods proposed by Huang et al. (9) and Nchama et al. (42).

a) Original image

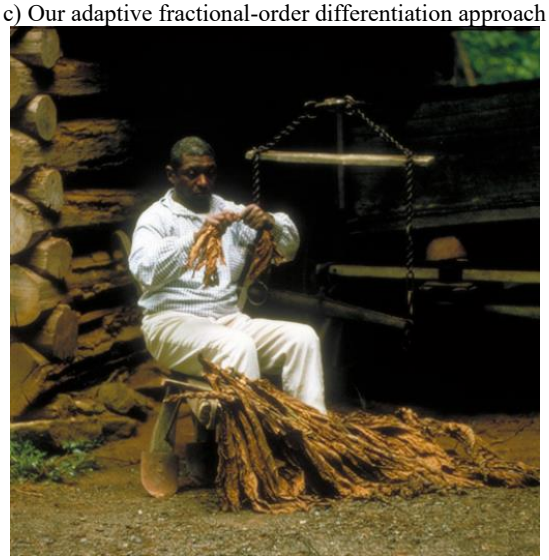
b) Blur image



c) Our adaptive fractional-order differentiation approach



d) Caputo fractional differential mask using alpha = 0.01

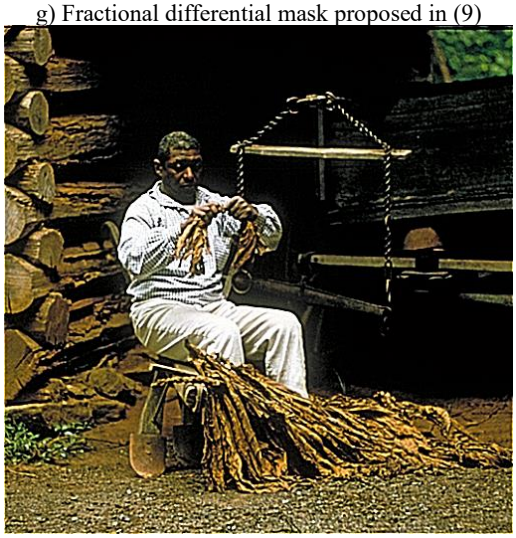


e) Caputo fractional differential mask using alpha = 0.6

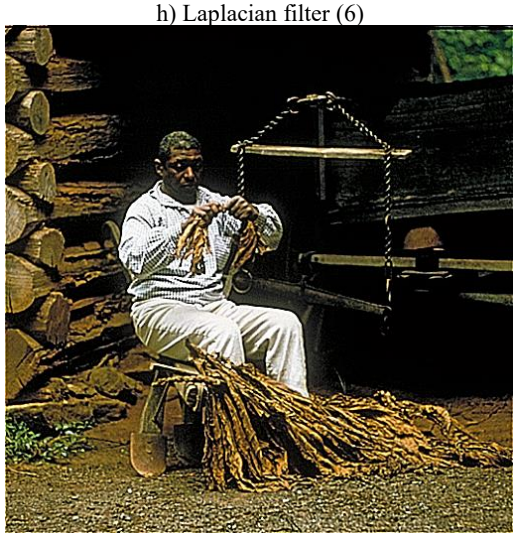
f) Caputo fractional differential mask using alpha = 0.9

g) Fractional differential mask proposed in (9)

h) Laplacian filter (6)



i) Fractional differential mask proposed in (42)



j) Un-sharp masking (5)

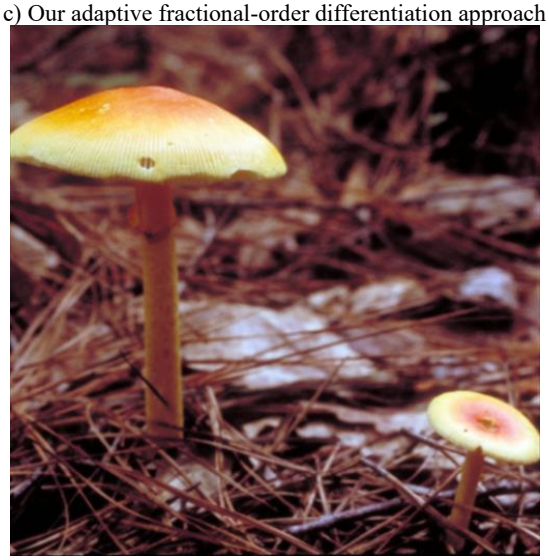
Figure 6. Results obtained from our adaptive fractional-order differentiation approach, the Caputo fractional differential mask using three distinct values for fractional order as 0.01, 0.6, and 0.9, classic un-sharp masking and the laplacian filter, as well as the methods proposed proposed by Huang et al. (9) and Nchama et al. (42).

a) Original image

b) Blur image

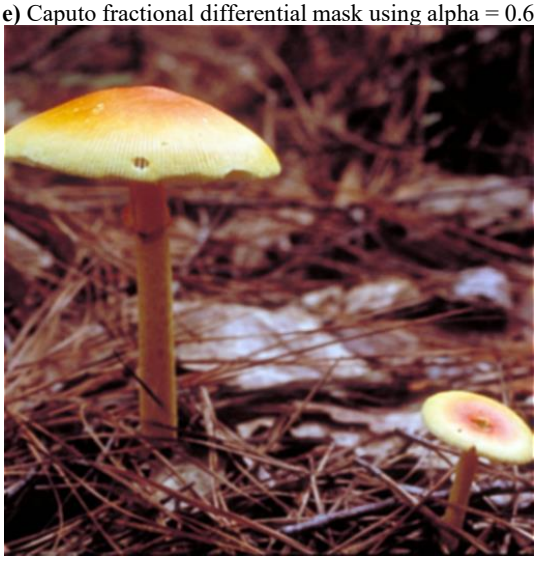
c) Our adaptive fractional-order differentiation approach

d) Caputo fractional differential mask using alpha = 0.01

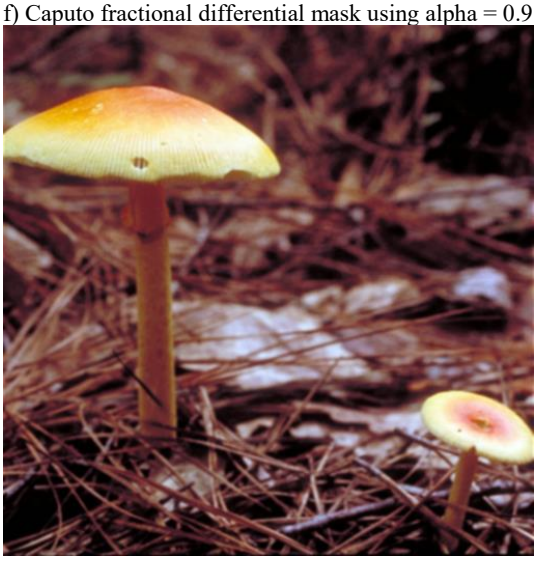


e) Caputo fractional differential mask using alpha = 0.6

f) Caputo fractional differential mask using alpha = 0.9



g) Fractional differential mask proposed in (9)



h) Laplacian filter (6)

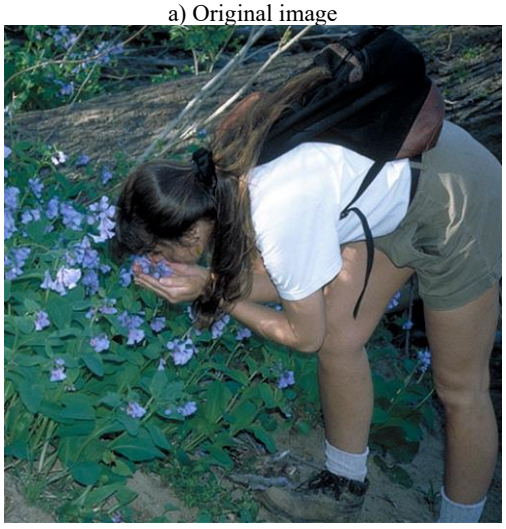
i) Fractional differential mask proposed in (42)

j) Un-sharp masking (5)

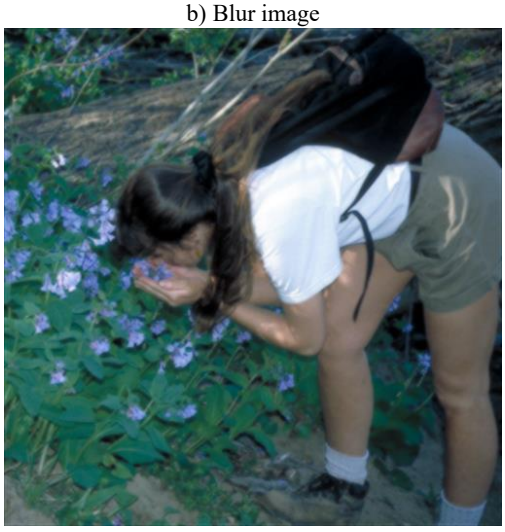
Figure 7. Results obtained from our adaptive fractional-order differentiation approach, the Caputo fractional differential mask using three distinct values for fractional order as 0.01, 0.6, and 0.9, classic un-sharp masking and the laplacian filter, as well as the methods proposed by Huang et al. (9) and Nchama et al. (42).

a) Original image

b) Blur image



c) Our adaptive fractional-order differentiation approach

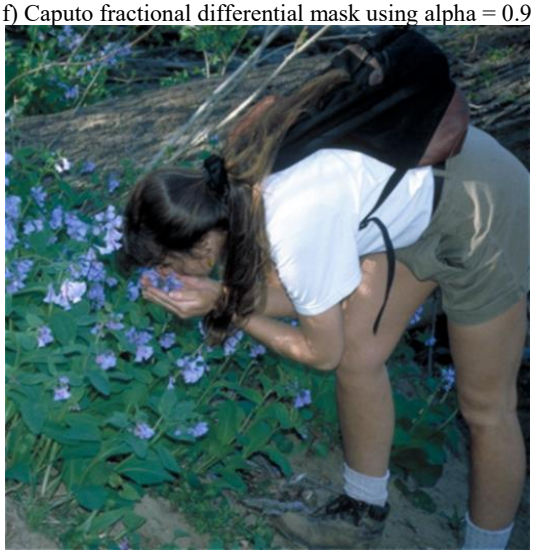


d) Caputo fractional differential mask using alpha = 0.01

e) Caputo fractional differential mask using alpha = 0.6

f) Caputo fractional differential mask using alpha = 0.9

g) Fractional differential mask proposed in (9)



h) Laplacian filter (6)

i) Fractional differential mask proposed in (42)

j) Un-sharp masking (5)

Figure 8. Results obtained from our adaptive fractional-order differentiation approach, the Caputo fractional differential mask using three distinct values for fractional order as 0.01, 0.6, and 0.9, classic un-sharp masking and the laplacian filter, as well as the methods proposed by Huang et al. (9) and Nchama et al. (42).

proposed by Huang et al. (9) and Nchama et al. (42). The images enhanced by our adaptive fractional-order differentiation approach are noticeably more visually appealing. In contrast, using a low fractional order of 0.01 in the Caputo fractional differential mask results in significant degradation of the original images, introducing unwanted blurriness. Additionally, fixing the fractional order at 0.6 does not generate sufficient quality improvement in the processed images. A high fractional order value of 0.9 produces image degradations through excessive image sharpening, which leads to reduced quality standards. Traditional integer-order differentiation techniques such as classic un-sharp masking and the Laplacian filter, along with the methods introduced by Huang et al. (9) and Nchama et al. (42) based on fractional calculus, fail to yield appropriate image improvements since they do not consider individual image enhancement needs.

The research details the FOM, SSIM, PSNR values, together with fractional order (alpha) parameters for individual instance images in Table 3. Our method adaptively selects varying fractional order values to enhance each individual image according to its particular requirements as shown in the provided table. Our adaptive fractional-order differentiation method obtains superior FOM, SSIM, and PSNR metrics than alternative approaches in objective performance measurements.

TABLE 3. Performance Analysis of FOM, SSIM, and PSNR with Adaptive Fractional Order Selection across Samples.

		Blur images	Our proposed approach		Caputo fractional differential mask using			The method proposed in [9]	The method proposed in [42]	Un-sharp masking [5]	Laplacian filter [6]
					alpha = 0.01	alpha = 0.6	alpha = 0.9				
Sample 1	FOM	0.9439	**alpha = 0.66**	**0.9507**	0.7958	0.9209	0.9246	0.8898	0.9017	0.8968	0.9208
	SSIM	0.9693		**0.9879**	0.8665	0.9753	0.8100	0.8890	0.7271	0.7054	0.9523
	PSNR	32.2237		**34.9588**	25.9124	32.1919	20.0057	26.9347	17.1096	16.9060	30.3252
Sample 2	FOM	0.9130	**alpha = 0.71**	**0.9548**	0.7404	0.8835	0.9355	0.8636	0.9093	0.9077	0.8870
	SSIM	0.9532		**0.9861**	0.8494	0.9558	0.8618	0.8842	0.7603	0.7569	0.9357
	PSNR	32.3221		**35.9164**	27.5238	32.0002	23.9633	28.9682	19.6990	19.9957	31.0927
Sample 3	FOM	0.8978	**alpha = 0.81**	**0.9746**	0.7877	0.8696	0.9454	0.8715	0.8996	0.9044	0.8809
	SSIM	0.9425		**0.9766**	0.8863	0.9355	0.9425	0.9206	0.8457	0.8611	0.9354
	PSNR	32.2531		**36.2001**	28.6382	31.2041	29.7402	30.8415	23.1100	24.4652	31.5638
Sample 4	FOM	0.8893	**alpha = 0.76**	**0.9846**	0.7149	0.8582	0.9224	0.8308	0.8943	0.8897	0.8809
	SSIM	0.9294		**0.9827**	0.8156	0.9341	0.8381	0.8519	0.7082	0.7005	0.9354
	PSNR	32.2852		**37.3249**	27.8072	31.8246	24.7817	29.3481	20.0606	20.5089	31.5638

Our proposed adaptive fractional-order differentiation approach can be applied to various computer vision applications. In sports scene analytics, it can be used as a powerful pre-processing step for enhancing the visual quality of frames in cricket-related tasks, such as batting shot classification (63) and umpire signal recognition (64), by improving edge and contrast information. Furthermore, in traffic scene analysis (65), the method can be used to enhance the visibility of vehicles, road signs, and lane markings in images. Beyond these applications, the proposed approach is also applicable in medical imaging (66), remote sensing (41), and surveillance systems (67), where adaptive enhancement can help reveal subtle features.

As mentioned before, in the proposed method, the appropriate fractional order is adaptively selected for each image based on its gradient information. Consequently, the method shows reduced effectiveness when processing images with extremely low-gradient information. To address this limitation, future work may involve combining gradient information with other image cues, such as entropy, or developing contrast-aware boosting strategies to stabilize performance under such conditions. Additionally, the proposed adaptive approach can be extended to other fractional-order differentiation methods such as Grünwald–Letnikov (G-L) and Riemann–Liouville (R-L), potentially improving their performance in image enhancement tasks.

5. CONCLUSION

This paper presents the key advantages of adaptive fractional-order differentiation techniques for image

enhancement, addressing the limitations of conventional integer-order methods. Our approach employs a Caputo fractional differential mask while dynamically adjusting the fractional order for each image, thereby enhancing visual quality, preserving critical details, and reducing noise amplification. Experimental results demonstrate that our method outperforms traditional techniques, confirming the significant potential of fractional-order differentiation in image processing applications.

6. REFERENCES

1. Dyke RM, Hormann K. Histogram equalization using a selective filter. The visual computer. 2023;39(12):6221-35. https://doi.org/10.1007/s00371-022-02723-8
2. Archana R, Jeevaraj PE. Deep learning models for digital image processing: a review. Artificial Intelligence Review. 2024;57(1):11. https://doi.org/10.1007/s10462-023-10631-z
3. J. Y. Advancements in Spatial Domain Image Steganography: Techniques, Applications, and Future Outlook. Applied and Computational Engineering. 2024;94:6-19. https://doi.org/10.54254/2755-2721/94/2024MELB0058
4. Chaudhary G. Optimizing fast fourier transform (FFT) image compression using intelligent water drop (IWD) algorithm. International Journal of Interactive Multimedia and Artificial Intelligence. 2022. https://doi.org/10.9781/ijimai.2022.01.004
5. Gonzalez R, Woods R. Digital image processing, addison-wesle y longman publishing co. Inc, Boston, MA, USA. 2001.
6. Khidse S, Nagori M. A comparative study of image enhancement techniques. Int J Comput Appl. 2013;81(15):28-32. https://doi.org/10.5120/14201-2421
7. Ding L, Goshtasby A. On the Canny edge detector. Pattern recognition. 2001;34(3):721-5. https://doi.org/10.1016/S0031-3203(00)00023-6
8. Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on pattern analysis and machine intelligence. 2002;12(7):629-39. https://doi.org/10.1109/34.56205
9. Huang G, Qin H-y, Chen Q, Shi Z, Jiang S, Huang C. Research on application of fractional calculus operator in image underlying processing. Fractal and Fractional. 2024;8(1):37. https://doi.org/10.3390/fractalfract8010037
10. Muresan CI, Birs IR, Dulf EH, Copot D, Miclea L. A review of recent advances in fractional-order sensing and filtering techniques. Sensors. 2021;21(17):5920. https://doi.org/10.3390/s21175920
11. Sánchez-Rivero M, Duarte-Mermoud MA, Travieso-Torres JC, Orchard ME, Ceballos-Benavides G. Analysis of Fractional Order-Adaptive Systems Represented by Error Model 1 Using a Fractional-Order Gradient Approach. Mathematics. 2024;12(20):3212. https://doi.org/10.3390/math12203212
12. Almeida R. On the variable-order fractional derivatives with respect to another function. Aequationes mathematicae. 2024:1-18. https://doi.org/10.1007/s00010-024-01082-0
13. Motłoch S, Sarwas G, Dzieliński A. Fractional derivatives application to image fusion problems. Sensors. 2022;22(3):1049. https://doi.org/10.3390/s22031049
14. Leonenko N, Podlubny I. Monte Carlo method for fractional-order differentiation. Fractional Calculus and Applied Analysis. 2022;25(2):346-61. https://doi.org/10.1007/s13540-022-00017-3
15. Nchama GAM, Alfonso LDL, Morales RR, Aneke EN. Asumu Fractional Derivative Applied to Edge Detection on SARS-COV2 Images. Journal of Applied Mathematics. 2022;2022(1):1131831. https://doi.org/10.1155/2022/1131831
16. Jalab HA, Ibrahim RW. Texture Enhancement Based on the Savitzky-Golay Fractional Differential Operator. Mathematical Problems in Engineering. 2013;2013(1):149289. https://doi.org/10.1155/2013/149289
17. Zakaria M, Moujahid A, Ikhouba M. A new fractional derivative operator and applications. International Journal of Nonlinear Analysis and Applications. 2023;14(1):1277-82. https://doi.org/10.22075/ijnaa.2022.26841.3423
18. Ben Makhlouf A, Benjemaa M, Boucenna D, Mchiri L, Rhaima M. On generalized proportional fractional order derivatives and Darboux problem for partial differential equations. Discrete Dynamics in Nature and Society. 2023;2023(1):6648524. https://doi.org/10.1155/2023/6648524
19. Caputo M. Linear models of dissipation whose Q is almost frequency independent—II. Geophysical journal international. 1967;13(5):529-39. https://doi.org/10.1111/j.1365-246X.1967.tb02303.x
20. Mishra SK, Singh KK, Dixit R, Bajpai MK. Design of Fractional Calculus based differentiator for edge detection in color images. Multimedia Tools and Applications. 2021;80(19):29965-83. https://doi.org/10.1007/s11042-021-11187-2
21. Gonzalez-Lee M, Vazquez-Leal H, Garcia-Martinez JR, Pale-Ramon EG, Morales-Mendoza LJ, Nakano-Miyatake M, et al. A new class of edge filter based on a cross-correlation-like equation derived from fractional calculus principles. Applied Sciences. 2024;14(13):5428. https://doi.org/10.3390/app14135428
22. Ananthi V, Thangaraj C, Easwaramoorthy D. Multifractal dimensions and fractional differentiation in automated edge detection on intuitionistic fuzzy enhanced image. Frontiers of Fractal Analysis: CRC Press; 2022. p. 153-71.
23. Mortazavi M, Gachpazan M, Amintoosi M. Improving Canny edge detection algorithm using fractional-order derivatives. Journal of Mathematical Modeling. 2022;10(4):495-514. https://doi.org/10.22124/jmm.2022.21875.1921
24. Chen M. Fractional-Order Adaptive P-Laplace Equation-Based Art Image Edge Detection. Advances in Mathematical Physics. 2021;2021(1):2337712. https://doi.org/10.1155/2021/2337712
25. Elgezouli DDE, Abdoon M, Belhaouari SB, Almutairi DK. A Novel Fractional Edge Detector Based on Generalized Fractional Operator. European Journal of Pure and Applied Mathematics. 2024;17(2):1009-28. https://doi.org/10.29020/nybg.ejpam.v17i2.5141
26. Ismail SM, Said LA, Madian AH, Radwan AG. Fractional-order edge detection masks for diabetic retinopathy diagnosis as a case study. Computers. 2021;10(3):30. https://doi.org/10.3390/computers10030030
27. Balochian S, Baloochian H. Edge detection on noisy images using Prewitt operator and fractional order differentiation. Multimedia Tools and Applications. 2022;81(7):9759-70. https://doi.org/10.1007/s11042-022-12011-1
28. Babu NR, Sanjay K, Balasubramaniam P. EED: enhanced edge detection algorithm via generalized integer and fractional-order operators. Circuits, Systems, and Signal Processing. 2022;41(10):5492-534. https://doi.org/10.1007/s00034-022-02028-0
29. Wang W, Jia Y, Wang Q, Xu P. An Image Enhancement Algorithm Based on Fractional-Order Phase Stretch Transform and Relative Total Variation. Computational Intelligence and Neuroscience. 2021;2021(1):8818331. https://doi.org/10.1155/2021/8818331

30. Asghari MH, Jalali B. Edge detection in digital images using dispersive phase stretch transform. International journal of biomedical imaging. 2015;2015(1):687819. https://doi.org/10.1155/2015/687819
31. Xu L, Yan Q, Xia Y, Jia J. Structure extraction from texture via relative total variation. ACM transactions on graphics (TOG). 2012;31(6):1-10. https://doi.org/10.1145/2366145.2366158
32. Huang T, Wang X, Xie D, Wang C, Liu X. Depth image enhancement algorithm based on fractional differentiation. Fractal and Fractional. 2023;7(5):394. https://doi.org/10.3390/fractalfract7050394
33. Ruiyin T, Bo L. Application of Fractional Differential Model in Image Enhancement of Strong Reflection Surface. Mathematics. 2023;11(2):444. https://doi.org/10.3390/math11020444
34. Land EH. The retinex theory of color vision. Scientific american. 1977;237(6):108-29.
35. Musa P, Al Rafi F, Lamsani M, editors. A Review: Contrast-Limited Adaptive Histogram Equalization (CLAHE) methods to help the application of face recognition. 2018 third international conference on informatics and computing (ICIC); 2018: IEEE.
36. Chen D, Chen Y, Xue D. Digital fractional order Savitzky-Golay differentiator. IEEE Transactions on Circuits and Systems II: Express Briefs. 2011;58(11):758-62. https://doi.org/10.1109/TCSII.2011.2168022
37. Mcbride AC. Univalent functions, fractional calculus, and their applications, edited by HM Srivastava and S. Owa. Pp 404.£ 39· 95. 1989. ISBN 0-7458-0701-1 (Ellis Horwood). The Mathematical Gazette. 1990;74(469):326-7. https://doi.org/10.2307/3619871
38. Kaur K, Jindal N, Singh K. Fractional Fourier Transform based Riesz fractional derivative approach for edge detection and its application in image enhancement. Signal Processing. 2021;180:107852. https://doi.org/10.1016/j.sigpro.2020.107852
39. Bouzeffour F. The Orthogonal Riesz Fractional Derivative. Axioms. 2024;13(10):715. https://doi.org/10.3390/axioms13100715
40. Mortazavi M, Gachpazan M, Amintoosi M, Salahshour S. Fractional derivative approach to sparse super-resolution. The Visual Computer. 2023;39(7):3011-28. https://doi.org/10.1007/s00371-022-02509-y
41. Azarang A, Ghassemian H. Application of fractional-order differentiation in multispectral image fusion. Remote sensing letters. 2018;9(1):91-100. https://doi.org/10.1080/2150704X.2017.1395963
42. Nchama GM, Alfonso L, Mecıas A, Richard M. Construction of Caputo-Fabrizio fractional differential mask for image enhancement. Progress in Fractional Differentiation and Application. 2020. https://doi.org/10.18576/pfda/070203
43. Caputo M, Fabrizio M. A new definition of fractional derivative without singular kernel. Progress in fractional differentiation & applications. 2015;1(2):73-85. http://dx.doi.org/10.12785/pfda/010201
44. Jalab HA, Ibrahim RW, Hasan AM, Karim FK, Al-Shamasneh AaR, Baleanu D. A new medical image enhancement algorithm based on fractional calculus. 2021. https://doi.org/10.32604/cmc.2021.016047
45. Gamini S, Kumar SS. Homomorphic filtering for the image enhancement based on fractional-order derivative and genetic algorithm. Computers and Electrical Engineering. 2023;106:108566. https://doi.org/10.1016/j.compeleceng.2022.108566
46. Karim FK, Jalab HA, Ibrahim RW, Al-Shamasneh AaR. Mathematical model based on fractional trace operator for COVID-19 image enhancement. Journal of King Saud University-Science. 2022;34(7):102254. https://doi.org/10.1016/j.jksus.2022.102254
47. Al-Shamasneh AaR, Ibrahim RW. Image denoising based on quantum calculus of local fractional entropy. Symmetry. 2023;15(2):396. https://doi.org/10.3390/sym15020396
48. Huang T, Wang C, Liu X. Depth image denoising algorithm based on fractional calculus. Electronics. 2022;11(12):1910. https://doi.org/10.3390/electronics11121910
49. Tanriover E, Kiris A, Tunga B, Tunga MA. A novel image denoising technique with Caputo type space–time fractional operators. Nonlinear Dynamics. 2024;112(21):19487-513. https://doi.org/10.1007/s11071-024-10087-y
50. Chen H, Qiao H, Wei W, Li J. Time fractional diffusion equation based on caputo fractional derivative for image denoising. Optics & Laser Technology. 2024;168:109855. https://doi.org/10.1016/j.optlastec.2023.109855
51. Diwakar M, Singh P, Garg D. Edge-guided filtering based CT image denoising using fractional order total variation. Biomedical Signal Processing and Control. 2024;92:106072. https://doi.org/10.1016/j.bspc.2024.106072
52. Ben-Loghfyry A, Charkaoui A. Regularized Perona & Malik model involving Caputo time-fractional derivative with application to image denoising. Chaos, Solitons & Fractals. 2023;175:113925. https://doi.org/10.1016/j.chaos.2023.113925
53. Atlas A, Bendahmane M, Karami F, Meskine D, Oubbih O. A nonlinear fractional reaction-diffusion system applied to image denoising and decomposition. Discrete and Continuous Dynamical Systems-B. 2021;26(9):4963-98. https://doi.org/10.3934/dcdsb.2020321
54. Khan A, Gaur S, Suthar D. Generalized Caputo–Fabrizio fractional operator: an application in image denoising. Applied Mathematics in Science and Engineering. 2024;32(1):2434002. https://doi.org/10.1080/27690911.2024.2434002
55. Scherer R, Kalla SL, Tang Y, Huang J. The Grünwald–Letnikov method for fractional differential equations. Computers & Mathematics with Applications. 2011;62(3):902-17. https://doi.org/10.1016/j.camwa.2011.03.054
56. Heymans N, Podlubny I. Physical interpretation of initial conditions for fractional differential equations with Riemann-Liouville fractional derivatives. Rheologica Acta. 2006;45(5):765-71. https://doi.org/10.1007/s00397-005-0043-5
57. Pratt WK. Digital image processing: PIKS Scientific inside: Wiley Online Library; 2007.
58. Mortezaie Z, Hassanpour H, Asadi Amiri S. An adaptive block based un-sharp masking for image quality enhancement. Multimedia Tools and Applications. 2019;78(16):23521-34. https://doi.org/10.1007/s11042-019-7594-4
59. Wang Z, Bovik AC. A universal image quality index. IEEE signal processing letters. 2002;9(3):81-4. https://doi.org/10.1109/97.995823
60. Mortezaie Z, Hassanpour H, Asadi Amiri S. Image enhancement using an adaptive un-sharp masking method considering the gradient variation. International Journal of Engineering Transactions B: Applications. 2017;30(8):1118-25. https://doi.org/10.5829/idosi.ije.2017.30.08b.02
61. Poobathy D, Chezian RM. Edge detection operators: Peak signal to noise ratio based comparison. IJ Image, Graphics and Signal Processing. 2014;10:55-61. https://doi.org/10.5815/ijigsp.2014.10.07
62. Larson EC, Chandler DM. Most apparent distortion: full-reference image quality assessment and the role of strategy. Journal of electronic imaging. 2010;19(1):011006--21. https://doi.org/10.1117/1.3267105

63. Dey A, Biswas S. Shot-ViT: cricket batting shots classification with vision transformer network. International Journal of Engineering Transactions C: Aspects. 2024;37(12):10.5829. https://doi.org/10.5829/ije.2024.37.12c.04

64. Dey A, Biswas S, Abualigah L. Umpire's signal recognition in cricket using an attention based DC-GRU network. International Journal of Engineering Transactions A: Basics. 2024;37(4):662-74. https://doi.org/10.5829/ije.2024.37.04a.08

65. Dorrani Z. Traffic scene analysis and classification using deep learning. International Journal of Engineering, Transactions C: Aspects. 2024;37(3):496-502. https://doi.org/10.5829/IJE.2024.37.03C.06

66. Farsi H NM, SM, Barati A, Mohamadzadeh, S. . Development of a Deep Learning Model Inspired by Transformer Networks for Multi-class Skin Lesion Classification. International Journal of Engineering Transactions A: Basics. 2026;39(1):135-47. https://doi.org/10.5829/ije.2026.39.01a.11

67. Mortezaie Z, Hassanpour H, Beghdadi A. People re-identification under occlusion and crowded background. Multimedia Tools and Applications. 2022;81(16):22549-69. https://doi.org/10.1007/s11042-021-11868-y

COPYRIGHTS

©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

ایجاد بهبود در کیفیت تصاویر یکی از ارکان‌های مهم پردازش تصویر دیجیتال محسوب می‌شود که هدف آن افزایش وضوح بصری تصویر از طریق روش‌های عملیاتی مختلف است. در حوزه مکان، از دو تکنیک محبوب برای بهبود کیفیت تصاویر استفاده می‌شود که شامل روش‌های مشتق‌گیری با مرتبه صحیح و مشتق‌گیری با مرتبه کسری است. تکنیک‌های سنتی مشتق‌گیری با مرتبه‌ی صحیح، با وجود کاربرد گسترده خود، از محدودیت‌هایی نظیر عدم تمایز در رفتار فرکانس‌های فضایی و افزایش نویز رنج می‌برند که منجر به کاهش کیفیت تصویر می‌شود. این مقاله رویکردی بر اساس مشتق‌گیری با مرتبه کسری معرفی می‌کند که از ماسک‌های مشتق کسری کاپوتو برای تقویت جزئیات تصویر استفاده می‌کند. در این رویکرد، اطلاعات گرادیان تصویر برای تعیین مرتبه کسری مناسب استفاده می‌شود. با تنظیم پویای مرتبه کسری بر اساس میزان بهبود کیفیت مورد نیاز در هر تصویر، ضمن حفظ جزئیات تصویر و کاهش نویز، کیفیت تصویر بهبود می‌یابد. نتایج آزمایشگاهی با استفاده از معیارهایی ارزیابی مختلف نشان می‌دهد که رویکرد پیشنهادی این مقاله نسبت به سایر روش‌های مورد مقایسه عملکرد بهتری در بهبود کیفیت تصاویر دارد.

IJE TRANSACTIONS B: Applications Vol. 39, No. 05, (May 2026) 1293-1304

International Journal of Engineering

Journal Homepage: www.ije.ir

Contribution of Perforated Pipe Roughness to Improve Sediment Transport in Hybrid Fluidization Method as an Alternative Dredging Technique

R. Azis*[a], M. Arsyad Thaha[b]

[a] *Civil Engineering Department, Engineering Faculty, Lakidende University, Indonesia*
[b] *Civil Engineering Department, Engineering Faculty, Hasanuddin University, Indonesia*

PAPER INFO

Paper history:
Received 26 April 2025
Received in revised form 15 June 2025
Accepted 18 July 2018

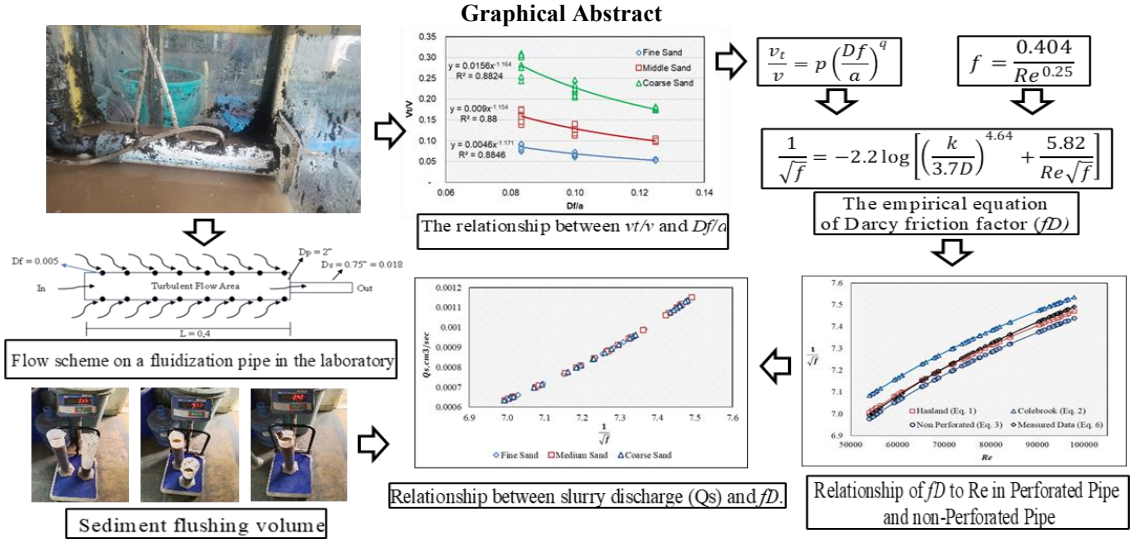
Keywords:
The Hybrid Fluidization
Darcy Friction Factor
Perforated Pipe Parameter
Pipe Roughness
Reynold Number

ABSTRACT

The slurry discharge (Qs) through the pump is made possible by the roughness of the perforation pipe during the dredging technique with the hybrid fluidization method. This study purposes to find the relationship between pipe roughness formed based on the physical parameters of the hole (*Df/a*) expressed through the equation of Darcy friction factor (*fD*) to the value of Reynold (Re) and Qs. Physical simulations in the laboratory with sediment variations in characteristics (ρ, d50), sediment thickness over the fluidization pipe (*db*), and variations in perforation hole spacing (*a*), and constant hole diameter (*Df*) were the experimental sets used. The dimensionless analysis method was used to analyze the correlation between flow parameters (*vt/v*) and the physical parameters of the pipe (*Df/a*). The *fD* equation was utilized by Colebrooke and Haaland to analyse the pipe's roughness in relation to the flow condition, which resulted in the formation of the friction relationship (*f*), *k/D*, and Re. The results of the analysis found the relationship between the ratio of vt/v and *Df/a* had *a* positive impact on the sediment transport process around the perforation pipe which resulted in a friction that was expressed as *fD* equation. The *fD* equation has been developed in an experimental and gives an empirical equation ($1/\sqrt{f} = -2.2\log[(k/3.7D)^{4.64} + 5.82/Re\sqrt{f}]$) is a theoretical contribution to the developed simple dredging technology. The empirical equation *fD* finding the relationship of the slurry discharge and the darcy friction factor by meeting the criteria of the physical parameters of the perforation hole (*Df/a*).

***doi**: 10.5829/ije.2026.39.05b.20*

Graphical Abstract



*Corresponding Author Email: rudiazis.civil@gmail.com (R. Azis)

Please cite this article as: Azis R, Arsyad Thaha M. Contribution of Perforated Pipe Roughness to Improve Sediment Transport in Hybrid Fluidization Method as an Alternative Dredging Technique. International Journal of Engineering, Transactions B: Applications. 2026;39(05):1293-304.

1. INTRODUCTION

The Hybrid fluidization represents a promising alternative dredging technique for trench maintenance, integrating the principles of fluidization with sediment management to enhance efficiency and reduce environmental impacts. This technique combines fluidization and sediment flushing within a single system, utilizing a perforated pipe that facilitates both processes simultaneously. Azis highlights the effectiveness of this hybrid system, demonstrating through experimental modeling that it simplifies dredging operations while maintaining performance standards (1). The study emphasizes the potential for hybrid fluidization to streamline maintenance dredging, particularly in challenging environments where traditional methods may be less effective.

The dynamics of fluidized beds, particularly in hybrid configurations, are crucial for optimizing dredging operations. Research by Ali et al. (2) indicated that the behavior of fluidized beds, including pulsed fluidization, can significantly influence sediment separation and management techniques. This is particularly relevant in dredging applications where the physical properties of sediments, such as bulk density and particle size distribution, affect the efficiency of sediment removal and transport. The integration of pulsed fluidization techniques can enhance the fluidization behavior, thereby improving the overall dredging process (3).

Moreover, the environmental implications of dredging practices cannot be overlooked. Dredging operations often lead to sediment resuspension, which can adversely affect water quality and aquatic ecosystems. Luo et al. (4) discussed the role of dredging in ecological restoration, noting that while it is essential for maintaining navigable waterways, it also poses risks of nutrient release from disturbed sediments. Therefore, hybrid fluidization techniques that minimize sediment disturbance and enhance sediment management are particularly valuable in mitigating these environmental impacts.

In addition to environmental considerations, the economic aspects of dredging operations are critical. Pellegrini et al. (5) presented a techno-economic assessment of innovative dredging methods, including jet pump systems, which are designed to reduce costs while minimizing environmental footprints. The hybrid fluidization approach can similarly contribute to cost-effective dredging solutions by enhancing operational efficiency and reducing the need for extensive sediment handling.

Furthermore, the application of advanced modelling techniques, such as hybrid 1D/2D hydrodynamic models, can provide insights into the impacts of dredging on sediment transport and flood mitigation (6). This modelling approach allows for a more nuanced understanding of how hybrid fluidization can be optimized for specific trench maintenance scenarios, ensuring that dredging operations are both effective and environmentally responsible.

The hybrid fluidization technique provides a viable alternative to the dredging approach, which is characterized by significant operational costs and requires the use of a dredge machine. The experimentation of hybrid fluidization technology in 2D is currently underway, but there's another fascinating aspect of the operating system. The combination of the fluidization method and the flushing method operated in a hybrid manner requires the design of the fluidization hole in the form of a perforation hole that must be optimal in terms of size in both methods. The perforation hole was found to have an impact on sediment transport due to friction in the hybrid fluidization experiment. In the numerical simulation of fluid flow through the perforated pipe, it was found that the suction flow in the perforation hole is radial inflow that moves perpendicular to the flow in the pipe (main inflow), so that the collision of the flow in both creates a lubrication effect that can reduce friction resistance (7). The hybrid fluidization experiment suggests that the lubrication effect will increase the flow in the perforation hole, thus improving its suction power.

Lubrication within perforation holes can reduce frictional resistance, which is critical in enhancing flow rates. For instance, Kravchuk et al. (8) conducted experiments on a steel pipeline with perforated sections, measuring fluid flow and pressure loss. Similarly, Abdullah et al. (9) highlighted that increasing the diameter of perforation holes and optimizing slope conditions can enhance exfiltration rates, further implying that lubrication could facilitate smoother flow through larger orifices.

The pressure drop across perforation holes is a significant factor in flow dynamics. The pressure drop generated in perforation holes is often negligible compared to the overall flow pressure difference, particularly in high-permeability reservoirs (10). This suggests that while lubrication may not drastically alter pressure drops, it can still contribute to maintaining optimal flow conditions by minimizing friction losses (11, 12). Furthermore, Mustafa et al. (13) discussed the components of total pressure drop in a horizontal wellbore, emphasizing that effective lubrication can mitigate friction and enhance fluid mixing, thereby improving flow efficiency.

In hybrid fluidization experiments, a factor that influences successful sediment discharge is the roughness effect of perforation holes on the fluidization pipe. The roughness caused by perforation holes has a direct impact on fluid dynamics within the pipe (14, 15). The ratio of hole diameter to hole spacing (*Df/a*) is a

parameter that forms the surface roughness effect of the fluidization pipe designed as a hole that agitates and sucks sediment in the hybrid fluidization method (1). The studies by Kravchuk et al. (8) indicate that the diameter variation of perforation holes can created measurable differences in pressure loss across the perforated section, suggesting that larger perforations may exacerbate disruptions in fluid flow and lead to greater overall resistance in hydraulic systems. This disruption of fluid flow benefits the pouring process, as sediment that enters through the perforation hole does not experience sedimentation around the perforation hole.

The suction flow must reduce the pressure along the pipe during discharge to prevent pressure changes. Research by Kravchuk et al. (8) showed that the diameter of the perforation hole significantly affects the pressure loss and flow speed of the fluid inside the pipe. In this case, the pressure drop due to the perforation hole does not have an impact on the loss of power to pour because in a hybrid fluidization experiment the role of the suction pump and the fluidization pump working simultaneously can create a balance in the flow. The fluidization pump still supplies the suction flow along the fluidization pipe with the perforation hole in this case.

The turbulent effect on the flow after the perforation hole is due to the surface roughness created by the slurry flowing through it. Various studies have provided a comprehensive explanation of this condition. For instance, Gong et al. (16) noted that a reduction in effective diameter due to increased roughness corresponds with higher friction factors, which consequently elevates losses in the system. This finding highlights the necessity of carefully considering perforation sizes and their resultant roughness to optimize fluid flow efficiency. However, in a hybrid fluidization system, the role of perforation holes both in terms of hole diameter and hole distance (*Df/a*), only serves to optimize the fluidization function and the sediment suction function in a hybrid manner (Figure 1). In order for the mixture flow to maintain turbulent conditions, it is necessary to meet the pressure factor that results in energy loss around the perforation hole.

The magnitude of fluid friction in the perforation hole is determined by calculating the friction factor provided by perforation roughness through a formula.

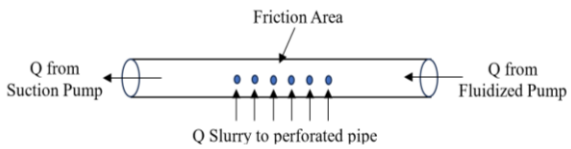


Figure 1. Perforation hole scheme in a hybrid fluidization system

This experiment uses the Darcy friction formula (*fD*) to measure the degree of roughness of the duct walls. The Darcy friction formula is a fundamental equation used to analyze the friction factor in various types of pipes, including perforated pipes. The friction factor is crucial for understanding pressure drops and flow characteristics in fluid systems. In perforated pipes, the flow dynamics are influenced by the presence of holes, which can alter the effective roughness and flow regime, thus affecting the friction factor.

The friction factor in perforated pipes can be significantly influenced by the geometry and arrangement of the perforations. For instance, studies have shown that the pressure drop in perforated pipes is affected by multiple factors, including wall friction, perforation roughness, inflow acceleration, and mixing effects (17). The friction factor can be decomposed into contributions from turbulent stress and other flow characteristics, which are modified by the presence of secondary currents in partially filled pipes (18). This highlights the complexity of flow in perforated systems, where traditional models may need adjustments to account for these unique flow patterns.

The relationship between the Reynolds number and the friction factor is well established. As the Reynolds number increases, the friction factor typically decreases, which is attributed to the inversely proportional relationship between the friction factor and average flow velocity (19, 20). This relationship is critical when applying the Darcy-Weisbach equation, which is often used in conjunction with empirical correlations like the Colebrook-White equation to estimate the friction factor in turbulent flow conditions (21). The Darcy-Weisbach equation is particularly relevant in practical applications, as it provides a more reasonable basis for calculating head loss compared to other formulas like Hazen-Williams, especially in critical flow scenarios (22).

In the context of perforated pipes, specific studies have focused on the hydraulic friction factor, revealing that the hydraulic characteristics can vary significantly based on the design of the perforations and the flow regime (8). For example, experimental investigations have shown that the hydraulic friction factor in perforated pipelines is crucial for developing reliable hydraulic calculations, which are essential for effective engineering design (23). Additionally, numerical studies have demonstrated that the friction factor can be accurately estimated using computational fluid dynamics (CFD) models, which allow for detailed analysis of flow characteristics in complex geometries (24).

Estimating the Darcy friction factor (*fD*) in the hybrid fluidization method is like estimating the energy loss that could occur because of the roughness effect in the perforation hole. The author strongly believes that estimating the magnitude of energy loss in the hybrid

fluidization method will be the same as estimating energy loss in other hydraulic case studies as in hydraulic jumping (25). Therefore, in research that is still 2-Dimensional, hydraulic behavior is produced, especially at the magnitude that will determine the success of the application of dredging technology that only uses simple equipment, namely fluidization pipes (perforation pipes) and fluidization pumps and suction pumps.

1. 1. The Hybrid Fluidization in Dredging Technique The Hybrid fluidization emphasizes the combination process of fluidization and flushing techniques in solving sedimentation problems in estuarine grooves, especially in grooves that are predominantly influenced by sea waves, resulting in sediment movement that cannot be forwarded to the sea (1). Despite the limited implementation of hybrid fluidization research in 2-dimensional testing, several hydraulic phenomena can be considered for field scale applications in the future. One of the hydraulic behaviors that tends to affect the volume of sediment that can be sucked through the fluidization hole is the occurrence of friction forces due to the interaction of the perforation hole with the fluidization flow throughout the flushing process after the sediment is fluidized.

In the physical experiments of hybrid fluidization, various phenomena were found that can be attributed to several interdisciplinary research approaches. Hybrid fluidization research is the development of a fluidization system used in handling sediment problems in river mouths and channels in general (1, 26). One of the distinctive features of the hybrid fluidization experiment is the use of hydraulic science to analyze sediment movement in closed channels. This is a separate study where for several years previous fluidization research emphasized the hydraulic mechanism used to produce jet bursts in the sediment layer (27-29).

The results of the hybrid fluidization experiment found a relationship between the ratio of the perforation hole (*Df/a*) to the ratio of sediment flow velocity (*vt/v*) where the perforation hole exerts a roughness effect on the perforation pipe (1). The perforation hole area's roughness causes disturbance in the flow due to its lubricating effect. These conditions have a beneficial impact on the hybrid fluidization system, particularly on the relationship between flow velocity and sediment particle precipitation rate (*vt/v*).

1. 2. Surface Roughness in Fluidization Pipes The perforation hole's design is established using data from previous research. From the initial experiment, the number of holes represents the best quantification result. Lubrication has an effect on the fluid flow through the perforation hole, which causes the pipe wall to be rough. As a result of the pressure difference between the fluid flow through the pipe and the sediment layer above, the sediment layer will flow into the fluidization pipe.

The flow velocity of a channel is determined by friction, whether it's closed or open (30). The sediment suction process is affected by flow disturbances, which cause friction between the fluid and the pipe wall in the basic conceptual model of flow through perforated pipes. In perforation pipes, flow disturbances cause changes in flow velocity along the pipeline formed through contact between the radial inflow near the perforation hole and the main inflow at the center of the flow of the pipe (7).

1. 3. Darcy Friction Factor (*fD*) The Reynold value (Re) describes the flow conditions produced at the level of roughness of the perforated pipe wall where the shear stress (wall shear stress) in the perforated pipe causes the flow velocity to change. The friction factor is a function of the number Re and the relative roughness of the pipe *k/D* where K is the equivalent of the roughness of the pipe and D is the diameter of the pipe. The equation developed from the von Karman formula for turbulent flow with a comparison of the Colebrook-White formula was developed by Haaland who used the Darcy friction factor (f) to describe the flow conditions in perforation pipes.

$$\frac{1}{\sqrt{f}} = -1.8 \log\left[\left(\frac{\varepsilon/D}{3.7}\right)^{1.11} + \frac{6.9}{Re}\right] \quad (1)$$

The Colebrook-White formula is essential for calculating the Darcy friction factor (f) in turbulent flow conditions (31).

$$\frac{1}{\sqrt{f}} = -2 \log\left[\frac{k}{3.7D} + \frac{2.51}{Re\sqrt{f}}\right] \quad (2)$$

The application of this formula in practical scenarios can be complicated because it requires numerical methods for its solution, which is implicit. However, Haaland proposed an explicit approximation of the Colebrook equation that simplifies the calculation of the friction factor, making it more accessible for engineers and practitioners in fluid dynamics (32, 33).

As a validation of the equation for perforated pipes, it is necessary to consider f (darcy friction factor) for the Reynolds value in porous pipes which ranges from 3.545 to 105.000, where the flow occurs at the roughness of the finest pipe under turbulence conditions (34). The exact formula for turbulent flow in pipes without perforations is:.

$$\frac{1}{\sqrt{f}} = -2 \log \frac{5.28}{R^{0.89}} \quad (3)$$

The Darcy-Weisbach equation is characterized by its use of the Darcy friction factor, which is crucial in determining pressure losses in pipe systems. The relationship of friction factor and flow characteristics, like Reynolds number and relative roughness, is crucial

in hydraulic engineering. The Haaland equation provides a reliable means to estimate the friction factor for turbulent flow, particularly when the Reynolds number exceeds 2000 (35, 36). This approximation has been validated against various empirical data, demonstrating its effectiveness in predicting friction losses in different pipe configurations (13, 37).

The Haaland approximation has been studied extensively to improve the accuracy of friction factor calculations, in addition to its practical applications. For instance, researchers have explored the limitations of traditional methods, such as the Moody diagram, and have proposed alternative approaches that leverage computational fluid dynamics (CFD) to refine the estimation of the Darcy friction factor (38, 39). The explicit nature of Haaland's correlation allows for quicker computations, which is particularly beneficial in large-scale simulations where iterative methods would be computationally expensive (33, 40).

In addition, the Colebrook-White formula and its approximations are important beyond simply calculating pipe flow. They are also applicable in complex systems involving multiphase flows and varying fluid properties, where accurate friction factor estimation is crucial for system design and optimization (41, 42). The ongoing research in this area continues to enhance our understanding of fluid dynamics and improve the tools available for engineers (43, 44).

2. MATERIAL AND METHOD

2. 1. Material and Equipment In this study, the material used consists of sedimentary material in the form of sand (ρ = 2.598 – 2.673) and perforation pipes 40 cm long with the diameter determined based on the results of the fluidization research with a hole distance consisting of 4, 5 and 6 cm (1, 45).

2. 2. Experimental Test Method The research methodology involves conducting experiments with two techniques that have been extensively researched, which are fluidization and flushing techniques. The two techniques were combined in this study, starting with the process of injecting water through a fluidization pump to generate fluid pressure in the sediment layer and agitate the sediment. The discharge given continues to increase in the range of 450 cm^3/s – 1099.03 cm^3/s in the sediment layer 40 cm above the fluidization/perforation pipe, until a condition where the sediment layer is at a critical vortex which means that the fluidization pressure has been maximized, the flushing process can continue (29). The pressure is modified due to the flow of fluid in the pipe when the suction pump is turned on and the valve at the end is opened. At this stage, both the fluidization pump and the suction pump work in a hybrid manner to produce an optimal flow rate in the fluidization pipeline so that pressure changes occur. Sediment particles that had originally turned into post-fluidized slurry were sucked into the pipe through the perforation hole. Flow velocity measurements during the flushing stage are carried out to obtain flow conditions (Figure 2).

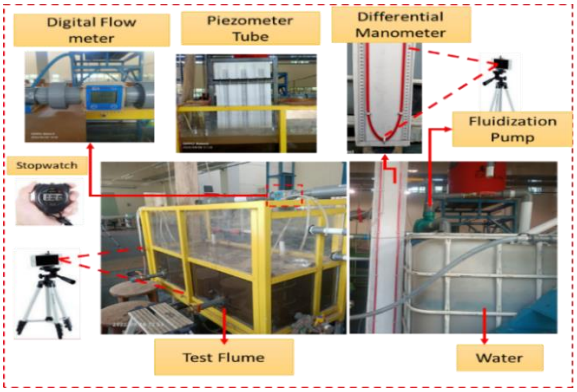


Figure 2. Research materials in the Laboratory

The volume of sediment that can be flowed through the porous fluidization pipe (perforation) was quantified by several experimental variations where the variables of suction pump height (*Hs*), perforation hole spacing (*a*), and variation in sediment type and thickness (*db*). The dimensionless analysis method was used to analyze the correlation between flow parameters (*vt/v*) and the physical parameters of the perforation hole (*Df/a*).

The results of the experiment were analyzed by analytical method with the friction factor formula developed by Darcy (*fD*) which was quantified from dimensionless parameters, namely the fluidization pipe parameters expressed in the perforation hole ratio (*Df/a*) expressed in the pipe roughness ratio (*ε/D, k/D*) and the mixed flow parameter (vt/v) expressed in Reynolds (Re).

3. ANALISYS RESULT

The velocity required in a hybrid fluidization system is the flow rate that can transport sediment through the fluidization pipeline where there are perforation holes as a factor that affects the flow conditions along the fluidization pipeline (Figure 3). The *vt/v* ratio is used to visualize the relationship between *Df/a* and *vt/v*, which has a certain velocity effect on the fluidization pipe

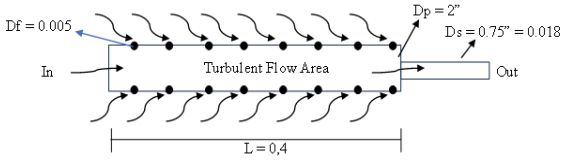


Figure 3. Flow scheme on a fluidization pipe in the laboratory

through the *vt/v* ratio. The graph's curves that show in Figure 4, medium, and coarse sediment types can be represented by equations.

$$\frac{v_t}{v} = p\left(\frac{Df}{a}\right)^q \quad (4)$$

The sediment velocity ratio (*vt/v*) ranges from 0.05 to 0.06 when the *Df/a* ratio is 0.13. The value of *vt/v* rises as the deaf/a ratio decreases. The vt/v ratio decreases when the *Df/a* ratio is 0.10 and 0.08, respectively. In the graph, it is evident that the vet/v ratio depends on the settling velocity (*vt*) of sedimentary particles. When the settling velocity is large, it leads to a high *vt/v* ratio. Fine sediments have a *vt/v* ratio of 0.05 to 0.09, medium sediments have a *vt/v* ratio of 0.10 to 0.17, and coarse sediments have a *vt/v* ratio of 0.17 to 0.31.

Comparing the magnitude of the friction coefficient (*f*) generated along the sediment flow in the perforated pipe to the Reynolds flow parameter (Re) allows for the expression of the *vt/v* ratio produced by the *Df/a* ratio. However, it has been confirmed that the Reynolds value (Re) of the flow is turbulent where the viscosity is very small, but the value of Re can explain the relationship between the flow parameters of the sediment mixture (*vt/v*) and the perforation hole parameters (*Df/a*). The formula proposed by Blassius is followed for determining the friction coefficient (*f*) during laboratory experiments. According to Blassius equation, the empirical equation for Fanning friction coefficient is:

$$f = \frac{0.404}{Re^{0.25}} \quad (5)$$

The friction factor (f) tends to have little impact on any increase in the Reynold value, this can be shown in the relationship f to any change in the Reynold value (Figure 5).

The flow relationship between Reynold's number and the friction factor is showing a positive trend in *f*. The friction factor was reduced by 0.404/Re0.25 (Equation 5) using an increased Reynold value of 1 x 10^5, which was applicable to all tested sediments. for sediments that are coarse, medium, or fine. The values of Reynold are categorized into fine sediments at 54,485– 96,600, medium sediments at 53,480 – 97,926, and coarse sediments at 53,962 – 96,635.

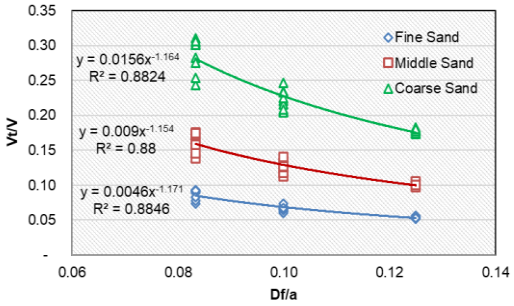


Figure 4. The relationship between *vt/v* and *Df/a*

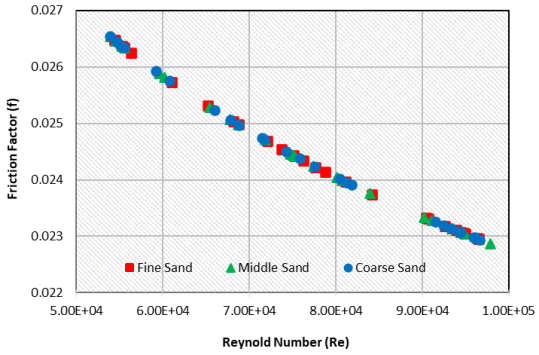


Figure 5. Reynold (Re) and Friction (f) Relationship

The friction value remained consistent within the range of 0.023–0.026 for every increase in sediment diameter, except for certain Reynolds numbers. This indicates that the range of friction numbers for increasing sediment diameter is limited to the flow of slurry around the perforation hole. This data indicates that the validation of non-perforated pipe slurry flow is different due to the increase in friction rate with increased sediment diameter (46). Furthermore, the friction coefficient (*f*) can be seen with *k/D* to produce a constant that is suitable for the development of the Darcy friction equation (*fD*)

Figure 6 of the *k/D* relationship in hybrid fluidization with friction coefficient *f* shows that the roughness caused by the lubrication effect of the perforation hole is proportional to the *k/D* ratio where the smaller the hole spacing means that the number of holes is greater than the large hole spacing the *k/D* value increases. The *k/D* values for hole spacing were 0.0017 to 4 cm, 0.0019 to 5 cm, and 0.002 to 6 cm, respectively.

The relationship between the friction coefficient, the number of Re, and *k/D* in the fluidization pipe can be obtained by determining it through the Blassius formula. For each perforation hole spacing (*a*), the *k/D* value is equal to 0.0017 for α = 4 cm, 0.0019 for α = 5 cm, and 0.002 for α = 6 cm in the moody diagram (47). The magnitude of the *k/D* value in each hole space tends to

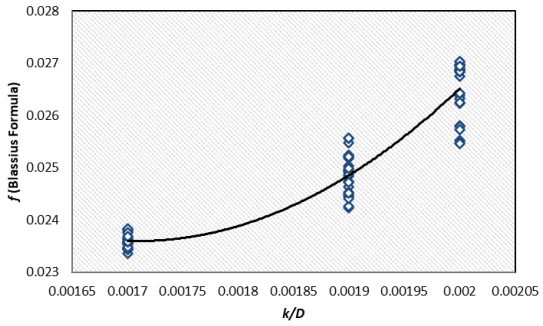


Figure 6. The interrelationship between *f* and *k/D*

affect the magnitude of the friction coefficient so that by using the Blassius formula for the magnitude of the friction coefficient (f) in the transition zone and the Colebrook equation for coarse hydraulic turbulent flow, the darcy friction factor (fD) can be written with an empirical equation:

$$\frac{1}{\sqrt{f}} = -2.2 \log\left[\left(\frac{k}{3.7D}\right)^{4.64} + \frac{5.82}{Re\sqrt{f}}\right] \quad (6)$$

Equation 6 is a development of Equation 1 proposed by Haaland and Equation 2 proposed by Colebrook, so empirically the equation of Darcy friction factor (fD) obtained from the results of the experiment by considering the roughness of the fluid flow that contracts between the sediment and the water flow from the fluidization pump (Figures 7 and 8).

The empirical equation fD can be compared to Equation 2 proposed by Haaland and Equation 1, which ranges from 7.1 to 7.5. According to the comparison in Figure 9, Equation 6's outcome is equivalent to the Colebrook equation, which represents the coarse hydraulics in the perforated pipe. By using the Haaland equation, the roughness value matches with the random distribution of the experiment's results.

A significant Darcy friction value exists in the small hole space, resulting in the main flow in the perforation pipe experiencing turbulence proportional to the k/D value. Darcy's friction coefficient (fD) at a hole distance of 4 cm ranges from 7.4 – 7.5, for a hole distance of 5 cm the fD value is between 7.4 – 7.2, and for a hole distance of 6 cm it ranges from 7.0 – 7.2. All types of sediments are subject to the same field value during flow experiments.

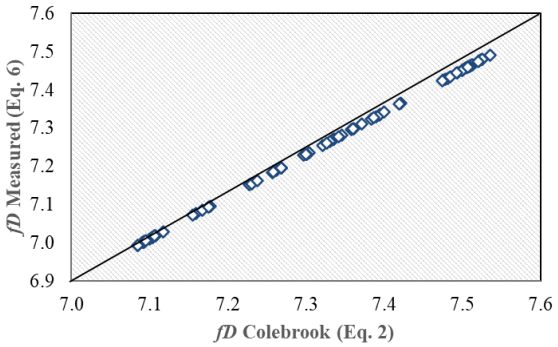


Figure 7. fD Equations 2 and 6

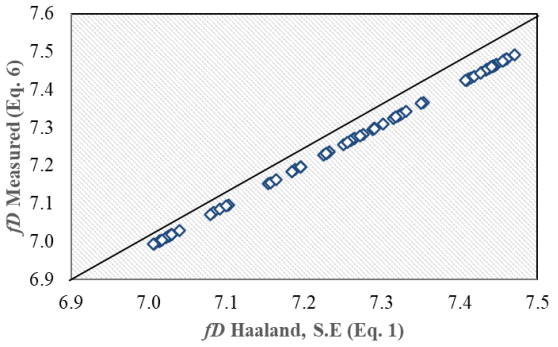


Figure 8. fD Equations 1 and 6

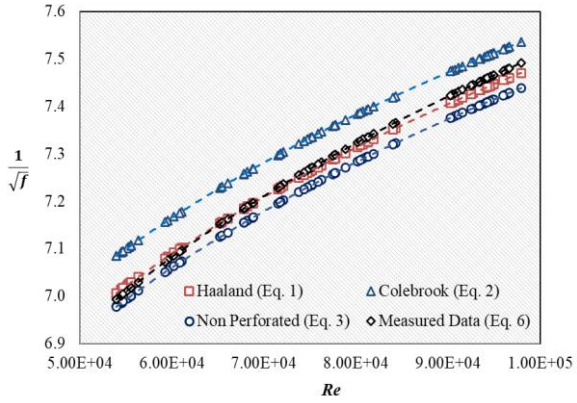


Figure 9. Relationship of fD to Re in Perforated Pipe and non-Perforated Pipe

It is based on the effect of flowing sediment particles which is considered an important factor affecting the field value, that with a typical sediment flow that has a diameter size of φ and a greater mass density ρ can minimize the effect of k/D or $\mathcal{E}/D$ values.

The field value for perforated pipes can be closely compared with that for non-perforated pipes in turbulent flow conditions. When the Reynold value is increased, the field value on the perforated pipe tends to be in close proximity to the Colebrook equation. The value of k/D or ε/d is not taken into account in non-perforated pipes due to the smaller roughness of the pipe (smoothing). The transition zone in perforated pipes in the hybrid fluidization model is where the lubrication effect is restricted to the horizontal walls of the pipe. The equations proposed by Haaland and the results of laboratory observations (Equation 6) are both prone to moving at specific k/D values until they reach the turbulent zone. Nikuradse divided turbulent zone results in the zone of Darcy's friction coefficient in Equation 6 being in the transition zone (31).

The value of Re in Figure 4 depends on the height condition of the pump (Hs) and the perforation hole (a). When the pump (Hs) increases, the Re value decreases. The same is true for the hole spacing (a). The greater the hole spacing, the less the Re value decreases. In fine sediments the Re values at Hs 40 cm range from 90.457 – 94.515 for a = 4 cm, 68.791 – 94.515 for a = 5 cm, and 54.446 – 61.045 for a = 6 cm. During Hs 30 cm, the range of values for a = 4 cm, 72.136 – 78.835 for a = 5 cm, and 54.548 – 65.225 for a = 6 cm is all over the same scale. The Re value for 4 cm in Hs 20 cm was 93,934 – 96,604, while for 5 cm in Hs 20 cm, it was 73,790 –

84,267, and 55,342 – 68,212. Medium sediments and coarse sediments with the same trend line have the same tendency line, and the Re value tends to be influenced by *Hs* conditions.

4. DISCUSSION

The *Df/a* ratio has an impact on the flow i.e. it gives a turbulence effect on the flow along the fluidization pipeline although the relationship is very small where turbulence can be seen at the Reynolds value (Re) formed on the viscosity effect of the fluid. The perforation hole's smoothing effect on the flow is influenced by the roughness of the perforation pipe wall (fluidization hole). The fluid flow that was mixed with sediment caused a collision with the fluid flow produced by the fluidization pump, resulting in a retention of the main flow around the pipe wall. As a result, the flow velocity in the fluidization pipeline decreases, which follows the results of the experiment in the graph The relationship of *vt/v* with *Df/a* is shown that the ratio of sediment flow velocity decreases with an increase in the ratio of *Df/a*.

The Re value is greater than the *k/D* ratio at the small perforation hole distance. This value is proportional to the magnitude of Darcy friction (*fD)* as shown in the figure of the *k/D* relationship with the value of *Re*. Turbulent flow conditions increase as *fD* increases, but they are not proportional to the *k/D* ratio. This is very suitable for the relationship between *Df/a* and *vt/v* where when the *Df/a* ratio is large, sediment particles in the perforation pipe are disturbed due to the lubrication effect.

Ouyang et al. (48) stated that turbulence in perforated pipes is a product of mass displacement of flow which causes the average velocity profile of the time for turbulent pipe flow to change due to the interaction of axial flow and inflow or outflow of the wall. The inflow lifts and expands the turbulent boundary layer thereby increasing the axial velocity of the flow. However, the process decreases the speed within the layers to follow the law of mass eternity. As a result, the axial velocity gradient near the pipe wall decreases, and so does the frictional tension of the wall (*fD*).

Turbulent flow in hybrid fluidization is a product of the cooperation of the fluidization pump and the suction pump in the flow fluidization pipe will experience a disturbance right at the perforation hole where the flow will experience different friction on the pipe wall that is not perforated. As a result, the flow of fluid in the fluidization pipe will experience turbulence conditions. The turbulent flow in the fluidization pipeline will be enough to help the sediment transportation process so that sedimentation does not occur during the discharge process. Darcy Friction Factor (*fD*) is used to obtain the relationship between the friction coefficient with the Reynold number (Re) and the equivalent friction coefficient (*k/D*) which describes the flow conditions along the pipe affected by the perforation hole factor. The large *k/D* ratio is directly proportional to the hole spacing (α) where the large hole spacing has a large *k/D* ratio.

The height of the Suction Pump (*Hs*) contributes positively to the change in flow conditions during the flushing process. Although this study focuses more on the hydraulic behavior of the flow based on *Re* and *k/D*, it is operationally quite influential, but in this analysis, the change is quite small so the *Hs* ratio can be ignored. Based on the relationship between the perforation hole parameter and the pump position, the *k/D* value to the Rep provides a correlation pattern that tends to correspond to the changing conditions of the hole spacing parameters (a) and the pump position (*Hs*). Figure 4 proves that *k/D* depends on the hole spacing (a) which forms a pattern of roughness in the pipe due to the lubrication effect.

As an expression of the potential disturbance of sediment deposits during the discharge process, the relationship between darcy friction (*fD*) and the value of Re can provide proof of the volume of sediment discharged just by comparing the level of roughness of the perforation holes in the fluidization pipe.

The relationship between slurry discharge and darcy coefficient (*fD*) coefficients in each type of sediment proves that the fluidization pipe density around the perforation hole has an effect on the volume of sediment that can be transported from the perforation hole (Figure 10). The lubrication effects of sediments with a moderate fraction of the d50 value of 0.321 are still greater than those of fine and coarse sediments. Each type of sediment fraction has its own tendency for increased slurry discharge (Qs) with each increase in lubrication effect. The slurry discharge in the hybrid fluidization model still depends on the *Df/a* ratio as in Equation 4, that the amount of sediment volume that can be sucked

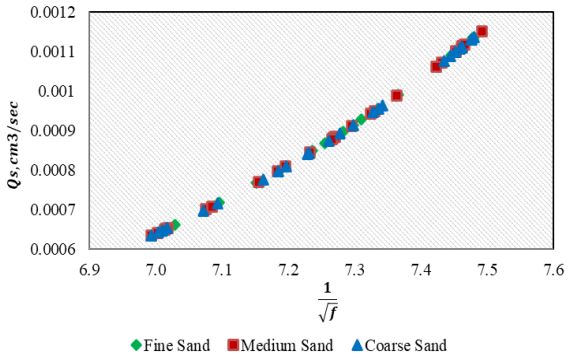


Figure 10. Relationship between slurry discharge (Qs) and *fD*

is of course influenced by the diameter of the perforation hole (*a*) and the hole space (*Df*). The non-dimensional parameter, *Df/a* ratio, can be utilized to modify the physical parameters of the hybrid fluidization design. Changing the physical parameters of the perforation hole can lead to the possibility of adjusting the value of darcy friction (*fD*) around the perforation hole.

4. 1. Theoritical Contribution The Darcy equation of friction (*fD*) can be utilized to gain a clear understanding of how mixed flow or slurry flow behaves during the passage through the perforation hole. The amount of friction around the perforation hole is expressed in Equation 6 which is a combination of the flow of slurry entering through the perforation hole (radial inflow) and the fluidization flow moving along the pipe where the speed of the radial inflow depends on the speed of the main flow (main inflow). The flow expression described in Equation 6 occurs when the perforation hole parameter (*Df/a*) is met, resulting in the slurry flow around the perforation hole not settled. The relationship of dimensionless numbers in Equation 4 represents this. Equation 6 also quantifies the magnitude of headloss due to friction in the perforation hole, but some studies state that the magnitude of headloss can be reduced by the magnitude of the speed in the main flow [7].

The perforation hole (*fD*), which has a Darcy friction value, can be represented by the expression $(1/\sqrt{f})$ which means that the friction coefficient has a quadratic effect on the slurry flow. The magnitude of the friction coefficient (*f*), in Equation 5 is expressed as a function of Reynolds (Re) where for the viscosity value is given a condition limit at a temperature of 30°C i.e. dynamic viscosity (μ) of 1.801x 10^{-3} N.d/m^2. Equation 6 is the formula for Darcy friction factor that is used to calculate energy loss in perforation holes. From an analytical perspective, Equation 6 is closer to the results achieved in Equation 1 (Haaland formula), but the equation structure resembles that of Equation 2 (Coolebrok formula) applied to experimental results as depicted in Figure 7. The Darcy friction factor (*fD*) formula can be empirically applied to perforation pipes in the suction method by using Equation 6.

4. 2. Practical Contribution The results of the darcy friction factor (*fD*) analysis in Equation 6 have provided a theoretical formulation that can be practically applied to quantify the magnitude of headloss in dredging techniques with the hybrid fluidization method. Although the hybrid fluidization method has not been applied in practical applications, the friction factor formula will be very useful in determining the ideal perforation hole (*Df/a*) ratio. Related to the limitations that must be met by the perforation hole design, the perforation pipe design for the hybrid fluidization method has been discussed in another paper where the author stipulates that the perforation hole ratio should follow Equation 4 where the constant magnitude for each type of transported slurry has been given according to Figure 3 about the relationship of the dimensionless parameter *Df/a* to *vt/v*.

However, the practical benefits of the results of this study still need to be tested in the 3D experiment to obtain the strength of the *fD* equation relationship strength to the volume of sediment that can be poured if the variation in the dimension of the perforation pipe is changed by a constant Df/a ratio as in Equation 4. In Figure 8 on the relationship between slurry volume and *fD* value, it has been obtained with a fairly good change and applies equally to all types of sediments, it shows that the limit of the *Df/a* ratio has been at the optimal point to be applied to both methods, namely hybrid fluidization and suction. The 3D experiment will be conducted to evaluate the geometric dimensions of grooves formed using the hybrid fluidization method.

5. CONCLUSION

From the results of the analysis and discussion, several conclusions can be written as follows:

1. The *vt/v* ratio and *Df/a* relationship have a beneficial effect on the sediment transport process surrounding the perforation pipe, producing a friction coefficient that is measured as Darcy friction factor (*fD*). As the *Df/a* ratio increases, the *vt/v* ratio decreases, indicating a lubrication effect that disturbs the possibility of sediment depositing in the fluidization pipe.
2. Preventing sedimentation during the pouring process is made possible by the expected friction coefficient determined by the magnitude of *k/D* and Re values.
3. The Darcy Friction Equation (*fD*) $(1/\sqrt{f} = -2.2\log[(k/3.7D)^{4.64} + 5.82/Re\sqrt{f}])$ is a theoretical contribution of the *fD* equation is an adaptation of Colebrooke and Haaland's *fD* equation that has a significant fit for this experiment.
4. The empirical findings on the *fD* equation apply to the perforation pipe type expressed in the relationship of the slurry discharge (Qs) and *fD* where for each increases Qs directly proportional to the increase in Darcy friction (*fD*) by meeting the criteria of the physical parameters of the perforation hole (*Df/a*) so that it can be used in estimating the magnitude of power loss during the operation of the hybrid fluidization method.
5. The scope of this research is limited to 2-D experiments that only measure hydraulic behavior. There is still a need for 3-D experiments that

directly obtains a geometric shape, groove, large volume of sediment that can be moved by changing (increasing) the diameter of the fluidization pipe and the power of the fluidization pump and suction pump

6. REFERENCES

1. Azis R, Maricar, F., Thaha, M. A., and Bakri, B. . The Hybrid System of Fluidization and Sediment Flushing for Maintenance Dredging Technique. Civil Engineering Journal (Iran). 2024;10(7):2275–92. https://doi.org/10.28991/CEJ-2024-010-07-013
2. Ali SS, Hossain SS, Asif M. Dynamics of partially collapsing pulsed fluidized bed. The Canadian Journal of Chemical Engineering. 2021;99(11):2333-44. https://doi.org/10.1002/cjce.24235
3. Ali SS, Basu A, Alfadul SM, Asif M. Nanopowder fluidization using the combined assisted fluidization techniques of particle mixing and flow pulsation. Applied Sciences. 2019;9(3):572. https://doi.org/10.3390/app9030572
4. Luo R, Li W, Zhong J, Dai T, Liu J, Zhang X, et al. Combining Multiple Remediation Techniques Is Effective for the Remediation of Eutrophic Flowing Water. Water. 2024;16(6):858. https://doi.org/10.3390/w16060858
5. Pellegrini M, Preda G, Saccani C. Application of an innovative jet pump system for the sediment management in a port channel: Techno-economic assessment based on experimental measurements. Journal of Marine Science and Engineering. 2020;8(9):686. https://doi.org/10.3390/JMSE8090686
6. Saad HA, Habib EH. Assessment of riverine dredging impact on flooding in low-gradient coastal rivers using a hybrid 1D/2D hydrodynamic model. Frontiers in Water. 2021;3:628829. https://doi.org/10.3389/frwa.2021.628829
7. Tangdan FM, Rohmat TA. Simulasi Numerik Aliran Fluida Melalui Pipa Berlubang. Journal of Mechanical Design and Testing. 2021;3(1):29-46. https://doi.org/10.22146/jmdt.58063
8. Kravchuk A, Cherniuk V, Kravchuk O, Airapetian T. ASSESSING THE VALUE OF THE HYDRAULIC FRICTION FACTOR IN PIPELINES WORKING WITH A FLOW CONNECTION ALONG THE PATH. Eastern-European Journal of Enterprise Technologies. 2022;119(7). https://doi.org/10.15587/1729-4061.2022.265670
9. Abdullah J, Mohd Arif Zainol MRR, Riahi A, Zakaria NA, Yusof MF, Shaharuddin S, et al. Investigating the relationship between the manning coefficients (n) of a perforated subsurface stormwater drainage pipe and the hydraulic parameters. Sustainability. 2023;15(8):6929. https://doi.org/10.3390/su15086929
10. Jiang H, Wu M, Zheng Y, Zhao Q, Gao Y. Construction and application of flow pressure drop model of perforated well considering pressure loss of perforation hole. Scientific Reports. 2023;13(1):19074. https://doi.org/10.1038/s41598-023-46443-y
11. Gajbhiye BD, Kulkarni HA, Tiwari SS, Mathpati CS. Teaching turbulent flow through pipe fittings using computational fluid dynamics approach. Engineering Reports. 2020;2(1):e12093. https://doi.org/10.1002/eng2.12093
12. Xu X, Fang L, Li A, Wang Z, Li S. 3D numerical investigation of energy transfer and loss of cavitation flow in perforated plates. Engineering Applications of Computational Fluid Mechanics. 2020;14(1):1095-105. https://doi.org/10.1080/19942060.2020.1792994
13. Mustafa MA, Rishack QA, Abdulwahid MA. Comparison Between Numerical and an Experimental Results of Pressure Drop in a Perforated Horizontal Wellbore with a 90 Perforations Phasing. 2023. https://doi.org/10.33971/bjes.23.1.15
14. Kravchuk A, Kravchuk O, Kochetov G. Improving the calculation of collecting perforated pipelines for water treatment structures. Восточно-Европейский журнал передовых технологий. 2020;6(10):23-8. https://doi.org/10.15587/1729-4061.2020.216366
15. Grzelak J, Szwaba R. Influence of Holes Manufacture Technology on Perforated Plate Aerodynamics. Materials. 2021;14(21):6624.
16. Gong J, Erkelens M, Lambert MF, Forward P. Experimental study of dynamic effects of iron bacteria–formed biofilms on pipeline head loss and roughness. Journal of Water Resources Planning and Management. 2019;145(9):04019038. https://doi.org/10.1061/(asce)wr.1943-5452.0001105
17. Tran Duc H, Thien PD. Research on Apparent Friction Factor in Perforated and Slotted Horizontal Filters to Dewater Opencast Mines Using Numerical Analysis. Inzynieria Mineralna-Journal of the Polish Mineral Engineering Society. 2020;1. https://doi.org/10.29227/im-2020-02-24
18. Liu Y, Stoesser T, Fang H. Effect of secondary currents on the flow and turbulence in partially filled pipes. Journal of Fluid Mechanics. 2022;938:A16. https://doi.org/10.1017/jfm.2022.141
19. Faraj AF, Azzawi ID, Yahya SG. Investigate the Effect of Pitch Variations on Helically Coiled Pipe for Laminar Flow Region Using CFD. International Journal of Heat and Technology. 2020;38:447-56. https://doi.org/10.18280/ijht.380221
20. Noori Y, Teymourtash AR, Zafarmand B. Use of random vortex method in simulating non-newtonian fluid flow in a t-junction for various reynolds numbers and power-law indexes. International Journal of Engineering, Transactions B: Applications. 2022;35(5):954-66. https://doi.org/10.5829/IJE.2022.35.05B.11
21. Khan WA, Arif M, Mohammed M, Farooq U, Farooq FB, Elbashir MK, et al. Numerical and Theoretical Investigation to Estimate Darcy Friction Factor in Water Network Problem Based on Modified Chun-Hui He's Algorithm and Applications. Mathematical Problems in Engineering. 2022;2022(1):8116282. https://doi.org/10.1155/2022/8116282
22. Abdulameer LS, Dzhumagulova N, Algretawee H, Zhuravleva L, Alshammari MH. Comparison between Hazen-Williams and Darcy-Weisbach equations to calculate head loss through conveyancing treated wastewater in Kerbala city, Iraq. Eastern-European Journal of Enterprise Technologies. 2022;1(1):115. https://doi.org/10.15587/1729-4061.2022.251385
23. Kravchuk A, Cherniuk V, Kochetov G, Kravchuk O, Airapetian T. DETERMINATION OF THE PARTICULARITIES OF THE HYDRAULIC FRICTION FACTOR VARIATION OF COLLECTING DRAINAGE PIPELINES. Eastern-European Journal of Enterprise Technologies. 2023;126(7). https://doi.org/10.15587/1729-4061.2023.292258
24. Csizmadia P, Dombóvári G, Till S, Minkó M. CFD-based Estimation of Friction Factor in Rough Pipes with Herschel-Bulkley Fluids. Periodica Polytechnica Chemical Engineering. 2023;67(1):74-82. https://doi.org/10.3311/ppch.20857
25. El-Baradei S, El-Abd M, Hazem N. Estimate of power output from hydraulic jumps generated downstream from barrages. J Hum Earth Future. 2024;5(1):72-84. https://doi.org/10.28991/HEF-2024-05-01-06
26. Weisman RN, Lennon GP, Roberts EW. Experiment on fluidization in unbounded domains. Journal of Hydraulic Engineering. 1988;114(5):502-15.

https://doi.org/https://doi.org/10.1061/(ASCE)0733-9429(1988)114:5(502)

27. Lennon GP, Weisman RN. Head requirement for incipient fluidization of fine sands in unbounded domains. Journal of Hydraulic Engineering. 1995;121(11):838-41. https://ascelibrary.org/doi/pdf/10.1061/%28ASCE%290733-9429%281995%29121%3A11%28838%29
28. Thaha MA, Triatmadja R, Yuwono N, Nizam N. Minimum jet velocity for unbounded domain fluidization as a new dredging methods. Engineering Journal. 2018;22(5):1-11. https://doi.org/10.4186/ej.2018.22.5.1
29. Azis R, Maricar F, Thaha MA, Bakri B, editors. Overview of Critical Vortex on Horizontal Jet Fluidization for Sediment Flushing Systems. International Conference on Civil Engineering; 2023: Springer.
30. Hasyim S, Salam, N., Pallu, M. S., and Maricar, F. . Experimental Study of Open Channel Model of Glass Fiber Reinforced Polymer (Gfrp) and Its Flowing Characteristics. International Journal of Engineering Transactions A: Basics. 2024;37(4):692–710. 10.5829/IJE.2024.37.04A.10
31. Colebrook CF, Blench T, Chatley H, Essex E, Finniecome J, Lacey G, et al. Correspondence. turbulent flow in pipes, with particular reference to the transition region between the smooth and rough pipe laws.(includes plates). Journal of the Institution of Civil engineers. 1939;12(8):393-422.
32. Siddiq HI. Energy Saving Via Accurate Computation of Crude Flow Measurement and Calculations. Journal of Petroleum Research and Studies. 2024;14(2):58-71. https://doi.org/10.52716/jprs.v14i2.845
33. Guelpa E, Sciacovelli A, Verda V. Thermo-fluid dynamic model of large district heating networks for the analysis of primary energy savings. Energy. 2019;184:34-44. https://doi.org/10.1016/j.energy.2017.07.177
34. Siwoń Z. Solutions for lateral inflow in perforated conduits. Journal of Hydraulic Engineering. 1987;113(9):1117-32.
35. Liu Y, Guo L, Wang Y, Liao Y, Zhao G. Numerical simulation and design optimization of large-scale geothermal production based on a multiwell layout in xianxian geothermal field. Lithosphere. 2022;2021(Special 5):2929551. https://doi.org/10.2113/2022/2929551
36. Cabral A, Wiggins CS, Carasik LB, Fishler J. Identification of surrogate fluids for molten salt coolants used in energy systems applications including concentrated solar and nuclear power plants. International Journal of Energy Research. 2022;46(3):3554-71. https://doi.org/10.1002/er.7405
37. Mikolajková-Alifov M, Pettersson F, Björklund-Sänkiaho M, Saxén H. A model of optimal gas supply to a set of distributed consumers. Energies. 2019;12(3):351. https://doi.org/10.3390/en12030351
38. Shaikh MM, Wagan AI. A sixteen decimal places' accurate Darcy friction factor database using non-linear Colebrook's equation with a million nodes: A way forward to the soft computing techniques. Data in brief. 2019;27:104733. https://doi.org/10.1016/j.dib.2019.104733
39. Plascencia G, Díaz–Damacillo L, Robles-Agudo M. On the estimation of the friction factor: A review of recent approaches. SN Applied Sciences. 2020;2(2):163. https://doi.org/10.1007/s42452-020-1938-6
40. Pérez-Pérez EdJ, López-Estrada FR, Valencia-Palomo G, Torres L, Puig V, Mina-Antonio JD. Leak diagnosis in pipelines using a combined artificial neural network approach. Control Engineering Practice. 2021;107:104677. https://doi.org/10.1016/j.conengprac.2020.104677
41. Qu W, Yao W, Xiong J, Cheng X. Experimental study of pressure loss in a 5× 5–rod bundle with the mixing vane spacer grid. Frontiers in Energy Research. 2021;9:675494. https://doi.org/10.3389/fenrg.2021.675494
42. Wambua CM, Joel OF, Ikiensikimama SS, Oyoo DO. Validation of Calculated Pressure Drop Using Experimental Data with Standard Methods of an Existing Software. https://doi.org/10.9734/jerr/2019/v9i317020
43. Dwivedi A, Sawant G, Karn A. Computational Solvers for Iterative Hydraulic loss Calculations in Pipe Systems. Journal of Engineering Education Transformations. 2015. https://doi.org/10.16920/jeet/2022/v35i4/22106
44. Saldivia D, Taylor RA. A Novel Dual Receiver–Storage Design for Concentrating Solar Thermal Plants Using Beam-Down Optics. Energies. 2023;16(10):4157. https://doi.org/10.3390/en16104157
45. Sucipto A, Thaha, M. A., Hatta, M. P., and Mahmuddin, F. . Sedimentation Characteristics and Sediment Transport in the Palu River Estuary. Civil Engineering Journal (Iran). 2025;11(2):437–52. https://doi.org/10.28991/CEJ-2025-011-02-03
46. Bartosik A. Validation of friction factor predictions in vertical slurry flows with coarse particles. Journal of Hydrology and Hydromechanics. 2020;68(2):119-27. https://doi.org/10.2478/johh-2020-0005
47. Moody LF. Friction factors for pipe flow. Transactions of the American Society of mechanical engineers. 1944;66(8):671-8.
48. Ouyang L-B, Arbabi S, Aziz K. A single-phase wellbore-flow model for horizontal, vertical, and slanted wells. Spe Journal. 1998;3(02):124-33. https://doi.org/10.2118/36608-pa

COPYRIGHTS
©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

Persian Abstract

چکیده

تخلیه دوغاب (Qs) از طریق پمپ به دلیل زبری لوله سوراخکاری در طول تکنیک لایروبی با روش سیال سازی ترکیبی امکان پذیر است .هدف این مطالعه یافتن رابطه بین زبری لوله تشکیل شده بر اساس پارامترهای فیزیکی سوراخ (Df/a) است که از طریق معادله Darcy Friction Factor (fD) به مقدار Reynold (Re) و Qs بیان می شود .شبیه سازی های فیزیکی در آزمایشگاه با تغییرات رسوب در مشخصات ،(d50 ، ρ) ضخامت رسوب روی لوله سیال سازی (db) و تغییرات در فاصله سوراخ سوراخ (a) و قطر سوراخ ثابت (Df) مجموعه های آزمایشی مورد استفاده بودند .از روش تحلیل بدون بعد برای تجزیه و تحلیل همبستگی بین پارامترهای جریان (vt/v) و پارامترهای فیزیکی لوله (Df/a) استفاده شد .معادله fD توسط Colebrooke و Haaland برای تجزیه و تحلیل زبری لوله در رابطه با شرایط جریان استفاده شد که منجر به تشکیل رابطه اصطکاک(f) ، k/D و Re شد .نتایج تجزیه و تحلیل نشان داد که رابطه بین نسبت vt/v و Df/a تأثیر مثبتی بر فرآیند انتقال رسوب در اطراف لوله سوراخ دار دارد که منجر به اصطکاکی می شود که به صورت معادله fD بیان می شود .معادله fD به صورت تجربی توسعه یافته و یک معادله تجربی $(1/\sqrt{f} = -2.2\log[(k/3.7D)^{4.64} + 5.82/Re\sqrt{f}])$ ارائه می دهد که یک کمک نظری به فناوری لایروبی ساده توسعه یافته است .معادله تجربی fD با برآورده کردن معیارهای پارامترهای فیزیکی سوراخ دار (Df/a) رابطه تخلیه دوغاب و Darcy Friction Factor را پیدا می کند.

AIMS AND SCOPE

The objective of the International Journal of Engineering is to provide a forum for communication of information among the world's scientific and technological community and Iranian scientists and engineers. This journal intends to be of interest and utility to researchers and practitioners in the academic, industrial and governmental sectors. All original research contributions of significant value focused on basics, applications and aspects areas of engineering discipline are welcome.

This journal is published in three quarterly transactions: Transactions A (Basics) deal with the engineering fundamentals, Transactions B (Applications) are concerned with the application of the engineering knowledge in the daily life of the human being and Transactions C (Aspects) - starting from January 2012 - emphasize on the main engineering aspects whose elaboration can yield knowledge and expertise that can equally serve all branches of engineering discipline.

This journal will publish authoritative papers on theoretical and experimental researches and advanced applications embodying the results of extensive field, plant, laboratory or theoretical investigation or new interpretations of existing problems. It may also feature - when appropriate - research notes, technical notes, state-of-the-art survey type papers, short communications, letters to the editor, meeting schedules and conference announcements. The language of publication is English. Each paper should contain an abstract both in English and in Persian. However, for the authors who are not familiar with Persian, the publisher will prepare the latter. The abstracts should not exceed 250 words.

All manuscripts will be peer-reviewed by qualified reviewers. The material should be presented clearly and concisely:

- *Full papers* must be based on completed original works of significant novelty. The papers are not strictly limited in length. However, lengthy contributions may be delayed due to limited space. It is advised to keep papers limited to 7500 words.
- *Research* notes are considered as short items that include theoretical or experimental results of immediate current interest.
- *Technical notes* are also considered as short items of enough technical acceptability with more rapid publication appeal. The length of a research or technical note is recommended not to exceed 2500 words or 4 journal pages (including figures and tables).

Review papers are only considered from highly qualified well-known authors generally assigned by the editorial board or editor in chief. Short communications and letters to the editor should contain a text of about 1000 words and whatever figures and tables that may be required to support the text. They include discussion of full papers and short items and should contribute to the original article by providing confirmation or additional interpretation. Discussion of papers will be referred to author(s) for reply and will concurrently be published with reply of author(s).

INSTRUCTIONS FOR AUTHORS

Submission of a manuscript represents that it has neither been published nor submitted for publication elsewhere and is result of research carried out by author(s). Presentation in a conference and appearance in a symposium proceeding is not considered prior publication.

Authors are required to include a list describing all the symbols and abbreviations in the paper. Use of the international system of measurement units is mandatory.

- On-line submission of manuscripts results in faster publication process and is recommended. Instructions are given in the IJE web sites: www.ije.ir-www.ijeir.info
- Hardcopy submissions must include MS Word and jpg files.
- Manuscripts should be typewritten on one side of A4 paper, double-spaced, with adequate margins.
- References should be numbered in brackets and appear in sequence through the text. List of references should be given at the end of the paper.
- Figure captions are to be indicated under the illustrations. They should sufficiently explain the figures.
- Illustrations should appear in their appropriate places in the text.
- Tables and diagrams should be submitted in a form suitable for reproduction.
- Photographs should be of high quality saved as jpg files.
- Tables, Illustrations, Figures and Diagrams will be normally printed in single column width (8cm). Exceptionally large ones may be printed across two columns (17cm).

PAGE CHARGES AND REPRINTS

The papers are strictly limited in length, maximum 8 journal pages (including figures and tables). For the additional to 8 journal pages, there will be page charges. It is advised to keep papers limited to 3500 words.

Page Charges for Papers More Than 8 Pages (Including Abstract)	
For International Author ***	**$55 / per page**
For Local Author	**100,000 Toman / per page**

AUTHOR CHECKLIST

- Author(s), bio-data including affiliation(s) and mail and e-mail addresses).
- Manuscript including abstracts, key words, illustrations, tables, figures with figure captions and list of references.
- MS Word files of the paper.

International **J**ournal of **E**ngineering, P.O. Box 13145-1659, Tehran, I.R. Iran
Materials and Energy Research Center, 5 Ahuramazda St., Alvand Ave., Argentina Sq., Tehran, Islamic Republic of Iran
www.merc.ac.ir - www.emerc.ac.ir

THOMSON REUTERS

WEB OF SCIENCE

Scopus®

CAS®
A DIVISION OF THE
AMERICAN CHEMICAL SOCIETY

Google
Scholar