



I. R. IRAN

ISSN: 1728-1431

e-ISSN: 1735-9244



International Journal of Engineering

Journal Homepage: www.ije.ir



TRANSACTIONS A: BASICS

Volume 37, Number 07, July 2024

Materials and Energy Research Center

INTERNATIONAL JOURNAL OF ENGINEERING

Transactions A: Basics

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

EDITORIAL ADVISORY BOARD

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

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.

CONTENTS

Transactions A: Basics

M. R. Gavadimoghaddam, N. Hadiani, S. M. A. Sadreddini, A. H. Eqbali	Performance of Pile Group in Soils Prone to Liquefaction Affected by Far and Near Earthquakes	1208-1220
E. A. Smirnova, L. A. Saychenko	Hydrodynamic Modeling and Evaluation of Partial Substitution of Cushion Gas During Creation of Temporary Underground Gas Storage in an Aquifer	1221-1230
A. V. Mikhailov, D. A. Shibarov, A. E. Bessonov, C. Bouguebrine	Comprehensive Assessment Production Efficiency of Electric Rope Shovel through Operator Qualification Criteria	1231-1238
M. Zoghi, H. Yaghobi	Synchronous Generator Dual Estimation Using Sigma Points Kalman Filter	1239-1251
E. Yogafanny, R. Triatmadja, F. Nurrochmad, I. Supraba	Leaching Potential and Effectiveness of Pervious Mortar Filters in Bacteria and Turbidity Removal from Surface Water	1252-1262
J. Selvan, S. Manavalla	An Innovative and Reliable Hybrid Cooling Method for Electric Vehicle Motors	1263-1273
T. Truong Cong, T. Nguyen Vu, D. Bui Minh, H. Vo Thanh, V. Dang Quoc	Analytical Modelling of a Six-Phase Surface Mounted Permanent Magnet Synchronous Motor	1274-1283
B. J. Rameshbhai, K. Rana	Investigating Hostile Post Detection in Gujarati: A Machine Learning Approach	1284-1295
M. Matin, M. Azadi	Effect of Training Data Ratio and Normalizing on Fatigue Lifetime Prediction of Aluminum Alloys with Machine Learning	1296-1305
A. Mazari, H. Ait Abbas, K. Laroussi, B. Naceri	Enhancing Wind Power Conversion System Control Under Wind Constraints Using Single Hidden Layer Neural Network	1306-1316
R. Fitriadi Kurnia, A. Anwar, Y. Latief, L. B. Sihombing	Critical Success Factor Implementation of Land Value Capture on a Toll Road Assignment Scheme: SEM-PLS Approach	1317-1330
F. Peyravi, S. E. Hosseini	Accurate Analytical Modeling of Drain Current of Heterojunction Tunneling Field Effect Transistor	1331-1342

G. Y. Korobov, A. A. Vorontsov, G. V. Buslaev, V. T. Nguyen	Analysis of Nucleation Time of Gas Hydrates in Presence of Paraffin During Mechanized Oil Production	1343-1356
F. Farrokh, A. Vahedi, H. Torkaman, M. Banejad	Design and Optimization of the Dual-Stator Axial-Field Flux-Switching Permanent Magnet Motor with High-Torque Density and Low-cost	1357-1368
M. Vatandoost, M. Golabchi, A. Ekhlassi, M. Rahbar	Topology and Thickness Optimization of Concrete Thin Shell Structures Based on Weight, Deflection, and Strain Energy	1369-1383
H. Hamidi, M. Mohammadi	A Model for Analysis of Social Media in Adoption of Mobile Banking	1384-1394
D. S. Tananykhin	Scientific and Methodological Support of Sand Management During Operation of Horizontal Wells	1395-1407
M. Najafi, A. Ghodratnama, S. H. R. Pasandideh, R. Tavakkoli-Moghaddam	Bi-objective Economic Production Quantity with Partial Backordering under Uncertainty	1408-1421
F. Shamshiri, P. Shahnazari-Shahrezaei, M. Fallah, H. Kazemipour	Fuzzy Dynamic Modeling for Export Consortia in Small and Medium-Sized Enterprises	1422-1442
H. Ataei, F. Ahmadizar, J. Arkat	Energy-Conscious Common Operation Scheduling in an Identical Parallel Machine Environment	1443-1465



Performance of Pile Group in Soils Prone to Liquefaction Affected by Far and Near Earthquakes

M. R. Gavadimoghaddam, N. Hadiani*, S. M. A. Sadreddini, A. H. Eqbali

Department of Civil Engineering, Islamshahr Branch, Islamic Azad University, Tehran, Iran

PAPER INFO

Paper history:

Received 07 August 2023

Received in revised form 16 December 2023

Accepted 30 December 2023

Keywords:

Liquefaction

Soil Liquefaction

Building Pile

System Dynamics

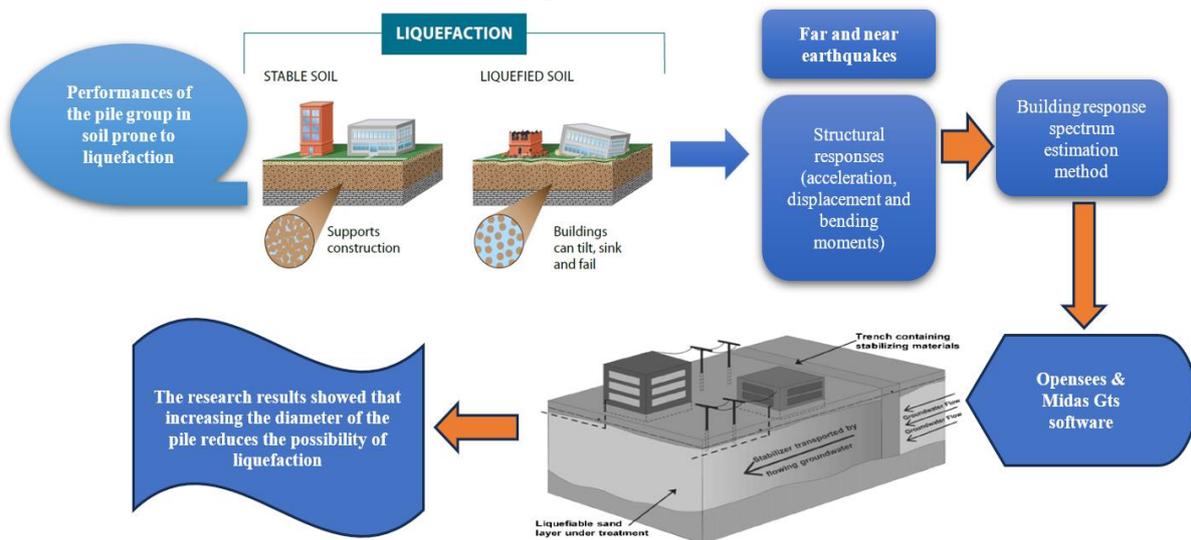
Numerical Modeling

ABSTRACT

The aim of the current research is to investigate the pile group functions in liquefaction prone soil under the influence of far and near earthquakes and structural responses such as acceleration history, displacement and bending moments. The method of estimating the response spectrum of the building was used according to the distance, size, structural conditions and fault mechanism and measuring the response spectrum of a selected recorded or artificial earthquake. By using opensees software, Midas Gts, different modes of research were formed. In the selection of earthquake accelerometers for the far and near areas, Peer Seismography Center, Iran Accelerography Center and Road and Housing Research Center were used to select and modify the accelerometers. The results of the research showed that increasing the diameter of the pile reduces the possibility of liquefaction, so that by increasing the diameter of the pile from 0.5 m to 1 m, the possibility of liquefaction and collapse of the pile group decreases by about 33%. In addition, with an increase in the distance between the piles, the probability of liquefaction increases, so that the probability of liquefaction of the group of piles and its collapse increases by about 14% when the distance between the piles increases from 1.5 to 5 meters. Finally, using the results of numerical modeling and with coding software, a suitable model for the pile group against liquefaction was presented for earthquakes in the far and near domains.

doi: 10.5829/ije.2024.37.07a.01

Graphical Abstract



* Corresponding Author Institutional Email: n.hadiani@gmail.com (N Hadiani)

1. INTRODUCTION

The damages caused by the earthquake may be directly attributed to the effects of earthquake or may be the indirect result of its direct damages. In addition, direct damages may be caused by the primary effects of shaking (namely, earthquakes and fault displacements) or secondary effects such as landslides and soil liquefaction, which themselves arise from the primary effects (1). Liquefaction is a term used to describe phenomena in which the production of excess pore water pressure leads to significant softening or weakening of soil deposits (2). This softening can be caused by dynamic loading or static loading. The phenomenon of changing the behavior of sand from solid state to liquid state was recognized in the first steps of expanding the knowledge of soil mechanics (3). Observing the destruction and failure of surface foundations due to liquefaction and lateral expansion in past earthquakes has revealed the need to fully understand the effects of this phenomenon on surface foundations. In such a way that the seismic response studies of foundations in liquefied soil are now one of the main topics of research in seismic geotechnical engineering. Significant research efforts have been made in this field over the past few years.

The review of past studies shows good information about parameters affecting liquefaction, but still there is a lack of a comprehensive parametric study considering earthquakes in the near area and the effects related to the type of faulting and extrawall and subwall effects. The lateral expansion caused by liquefaction has caused severe damage to many important structures during earthquakes (4). As a result, the evaluation of mass response to lateral expansion is an important step towards safe and resistant design against this destructive phenomenon (5). The purpose of the current research is to study the response of the group of piles that are exposed to lateral expansion caused by liquefaction, the responses of the structure such as the history of acceleration, displacement and bending moments should be investigated. In addition, the distribution of lateral soil pressure and movement patterns around the piles should be investigated and based on the results of the analysis of the responses of the pile group for different earthquake states (far field and near field), a suitable model for the design and arrangement of the pile group for different states should be made. According to the research conducted in non-flowing soils, the appropriate relative distance for arranging the piles is 3 to 5 times the diameter of the pile.

Fluidization has a significant effect on the dynamic response of the piles and increases the displacement of the pile compared to the non-fluidized state (6). Now, the question arises that in liquefied soils, can the same arrangement be used, or do other factors that result from

the effect of liquefaction on the dynamic response of the pile group have an effect on the choice of arrangement? In this research, an attempt is made to answer this question by studying this problem with the help of numerical modeling with the finite element method, and based on that, the optimal distance in flowable soils for the pile group is obtained. In this research, the assumption of a continuous environment and the assumption of large deformations are used.

At first, for an earthquake record, parametric studies will be done on the distance between the piles. Considering the importance of earthquakes in the nearby area, the effect of these earthquakes on liquefaction will also be investigated. Earthquakes in the near-fault zone are characterized by a shock movement with a short duration and a long period, which exposes the structure to a large input energy at the beginning of the record (7). This pulsed motion is particularly common in the forward direction ("progressive directionality"), where the fault rupture propagates toward the structure at a speed close to that of the shear wave. The "permanent displacement effect" which is the result of the displacement of the residual earth due to the tectonic deformations related to the failure mechanism, is described by a one-way velocity pulse with a long amplitude and with a uniform step in the time history of the displacement (8). Another feature of these earthquakes is the "overwall effect" which is created due to the proximity of the buildings on the overwall to the fault plane compared to the buildings with the same distance on the downwall (9). Investigating this factor (underwall or extrawall effect) on the structure and if the structure is in the area of the extrawall or underwall, what acceleration will it receive and how will the response of the structure be in front of these accelerations. Another important feature is "vertical component effect". The reason for the importance of this characteristic is the greater ratio of the maximum vertical to horizontal acceleration of the earth in records near the fault compared to far from the fault. Considering the importance of knowing the characteristics of these earthquakes, the characteristics of earthquakes near the fault and far from the fault and the effect of these on the acceleration on the structure and the type of failure and the design mechanism of the pile group for each of the acceleration pulses and providing a suitable model and taking into account the effect of types. The fault movements and liquefaction conditions are from the Shama group's plan using parametric analysis in the field of finite elements and dynamic analysis, i.e. Midas and Opensees, which will be done using the outputs of these two softwares and checking, comparing and relating the results of these two softwares. The purpose of this research is to provide a suitable model of the pile group design against lateral expansion resulting from liquefaction for near and far earthquakes.

2. MATHEMATICAL MODEL

In short, the research steps are as follows. The first step provides an overview of earthquakes in far and near areas. In the next step, numerical modeling for different research situations is created using OpenSees, Midas Gts engineering software. The number of research models is based on the number of variables. Another research step covers the use of the specific constitutive model or liquefaction. In the selection of earthquake accelerometers for far and near areas, the Peer Seismography Center, the National Accelerography Center, and the Road and Housing Research Center are used to select and modify the accelerometers. Finally, a suitable model is presented for the pile group against liquefaction for far and near-field earthquakes using the results of numerical modeling and coding software. To check the effect of the overwall and underwall, the pile group is once modeled above the fault (the pile group on the side where the fault moves upwards) and again below the fault (the pile group on the side where the fault moves downwards). Then, the behavior of Shama group is investigated against earthquakes in far and near areas.

In this article, all the simulations were done in three stages using OpenSees software. After validating the numerical model using centrifuge test, analysis was done for different conditions. By comparing the numerical results with the centrifuge test, it can be concluded that the use of P-Y curves with different degradation factors in flowable sand brings reasonable results. In addition, the non-parallel surface layer plays a key role in ground displacement, especially in sloping terrain.

2. 1. The Record Selection Method Selecting a number of earthquake records is the first step of incremental dynamic analysis. Here, a set of 40 three-component ground motion records (in three directions) without scale, each in the direction of the normal multiplication component, is used as earthquake excitation. This group of records has a large and almost constant variance around the average spectral values in different periods. The last feature ensures that earthquake records with different specification ranges are available for analysis and the aleatory ground motion uncertainty can be considered in situations where researchers are interested in evaluating the effect of ground motion variance on structural response estimation.

2. 2. Introduction of the UBCSAND Constitutive Model In this research, the UBCSAND constitutive model was used to investigate the behavior of liquefacted sand under dynamic loading. UBCSAND is a two-dimensional effective stress plasticity model that is used in advanced stress-deformation analysis of geotechnical structures (10-19). This model was developed to investigate the behavior of sandy soils with liquefaction

potential under dynamic loading (e.g. sands and silty sands with a relative density of approximately less than 80%). In this model, the stress-shear behavior of the soil is forecasted using a hypothetical hyperbolic relationship and the volumetric strain of the soil skeleton with a flow rule that is a function of the instant stress ratio. This model can perform a fully-coupled analysis in which the mechanical calculations and water flow are considered simultaneously and parallelly. In the UBCSAND constitutive model, the elastic components of the soil response with shear modulus G^e and bulk modulus B^e are specified as follows:

$$G^e = K_G^e \cdot P_a \cdot \left(\frac{\sigma'}{P_a}\right)^{n_e} \quad (1)$$

$$B^e = \alpha \cdot G^e \quad (2)$$

where K_G^e is the shear modulus number that depends on the relative density and ranges from 500 for loose sand to 2000 for dense sand. P_a is the atmospheric pressure, σ' is the average all-round stress in the loading plane, n_e ranges from 0.4 to 0.6 and is almost equal to 0.5, and α refers to Poisson's ratio. Plastic strains are controlled by yield surface and flow rule. The yield surface is expressed by a radial line passing through the origin of the stress space as shown in the figure. For the first shear loading, the yield surface is controlled by the current stress state (point A in Figure 1). When the shear stress increases, the stress ratio ($\eta = \frac{\tau}{\sigma'}$) increases and causes the stress point to move to point B. τ and σ' are the effective shear and normal stresses in the maximum effective shear stress plane. The yield surface is moved to a new location that passes through point B and the origin, leading to the creation of shear and volumetric plastic strains. The development of shear plastic strain is obtained by $d\gamma^p = \frac{1}{G^p \sigma'} \cdot d\eta$ where G^p is the plastic shear modulus. Further, it is assumed that a hyperbolic relationship exists between η and G^p as follows:

$$G^p = G_p^i \cdot \left(1 - \frac{\eta}{\eta_f} \cdot R_f\right)^2 \quad (3)$$

where G_p^i is the plastic modulus at a low stress ratio. η_f is the ratio of stress and tension in rupture equal to $\sin \phi_f$ where ϕ_f is the maximum friction angle. R_f is the rupture ratio which is used to fit the best hyperbolic relationship and avoid overpredicting the resistance during the rupture. R_f varies from 0.7 to 0.98 and decreases by increasing relative density (20, 21).

$d\varepsilon_v^p$ refers to the plastic volume strain development, and p represents the flow rule. Further, $d\gamma^p$ refers to the plastic shear strain development according to the following equation.

$$d\varepsilon_v^p = \left(\sin \phi_{cv} - \frac{\tau}{\sigma'}\right) d\gamma^p \quad (4)$$

where ϕ_{cv} is the constant volume friction angle or the phase transformation angle.

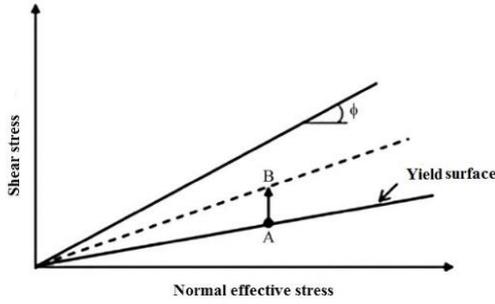


Figure 1. Yield surface in the UBCSAND model (10)

2. 3. Introduction of Moore-Coulomb's Constitutive Model

The Mohr-Coulomb model is often used for granular materials, such as MIDAS GTS concrete. In this model, the stress-strain relationship is considered as a complete elastic paste. Figure 2 presents Mohr-Coulomb yield criterion based on assumption causes a logical solution in nonlinear analysis.

The Mohr-Coulomb failure criterion is determined according to the failure envelope based on the Mohr stress circles and $F_1(\sigma)$ from laboratory tests. This criterion is a more general form of the Tresca criterion, in this way that the maximum shear stress is the yield criterion, but the maximum shear stress depends on the condition the compressive stresses are dependent. In other words, the following equation is true in the two-dimensional state:

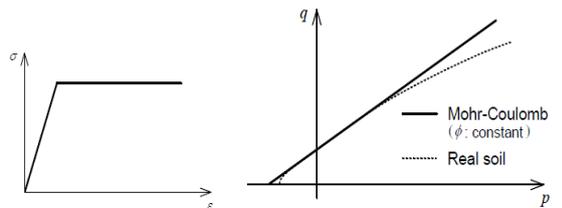
$$|\tau| = F_1(\sigma) \tag{5}$$

where τ is the maximum shear stress and σ is the compressive stress. F_1 is the function obtained from the test. The simplest form of function $F_1(\sigma)$, is a linear function. We can display this linear function as follows:

$$|\tau| = c - \sigma \tan \phi \tag{6}$$

where c and ϕ are obtained from experimental results. Physically, c is adhesion and ϕ is the angle of internal friction. This relationship was first presented by Coulomb and then independently by Mohr.

In the case that $\sigma_1 \geq \sigma_2 \geq \sigma_3$, using the principal stresses, the Mohr-Coulomb criterion is written as follows:



(a) Stress-strain relation (b) Schematic of yield function

Figure 2. Mohr-Coulomb yield criterion (22)

$$\Sigma_1 \frac{(1-\sin\phi)}{2c\cos\phi} - \Sigma_3 \frac{(1+\sin\phi)}{2c\cos\phi} = 1 \tag{7}$$

The Mohr-Coulomb failure surface is an irregular hexagon in the 3D space of principal stresses. where the bisector is a straight line. Figure 3 shows the Moore-Coulomb fracture surface in the three-dimensional apparatus of principal stresses. The failure surface of a regular octagon is located on the π plane ($\sigma_1 + \sigma_2 + \sigma_3 = 0$).

$$\rho_{to} = \frac{2\sqrt{6}c\cos\phi}{3+\sin\phi} \tag{8}$$

$$\rho_{co} = \frac{2\sqrt{6}c\cos\phi}{3-\sin\phi} \tag{9}$$

$$\frac{\rho_{to}}{\rho_{co}} = \frac{3-\sin\phi}{3+\sin\phi} \tag{10}$$

The Mohr-Coulomb yield surface on π plane and bisector plane is shown in Figure 4.

2. 4. Specifications of Numerical Model in Midas Software

In this research, a two-dimensional model was used based on the type of analysis and the time-consuming nature of the analysis. Figure 1 shows

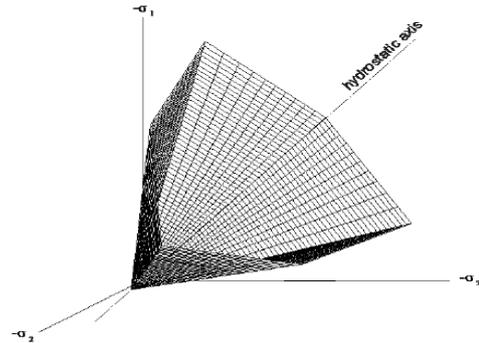
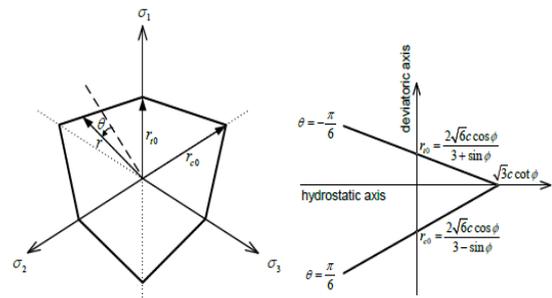


Figure 3. A view of the Moore-Coulomb fracture surface in the three-dimensional apparatus of principal stresses (6)



(a) Yield surface on π -plane (b) Yield surface on meridian plane ($\theta = -\frac{\pi}{6}$)

Figure 4. Mohr-Coulomb yield surface on π plane and bisector plane (12)

components in the numerical model. The model in the software includes three parts: Soil, pile, and foundation. The model consists of 4 layers of soil, respectively. The first layer is 2 meters thick, the second layer is liquefiable soil with a thickness of 5 meters, the third layer is 2 meters thick granular soil, and the last layer is a 15-meter thick weak rock. There is a foundation with a thickness of 1 meter and a length of 5 meters above the soil. Under the foundation, 4 piles with different distances and different lengths are modeled according to the type of analysis. Figure 5 shows the numerical model built in Midas.

2. 5. Boundary Conditions Elastic boundaries are needed in the boundaries due to the dynamic analysis of the model. For a better and more optimal performance of the process of applying the acceleration time history during seismic dynamic analysis, if boundaries are simultaneously elastic, the state with viscose, false, and intensified vibrations disappear and the model will be more closer to reality (23). The necessary tools are available in the Midas GTS software. In this regard, the supports are perpendicular to the plane with the spring capability of the bed itself. According to the following equation, they have damping coefficients to create a viscose mode. The equations for the coefficients of damping and damping are as follows:

$$k_{V0} = k_{h0} = \frac{1}{30} \cdot \alpha \cdot E_0 \tag{11}$$

It ranges from 1 to 4 according to the sample and laboratory conditions. The ground reaction coefficients (K_V, K_h) can be obtained by the following equation (in cm):

$$K_V = k_{V0} = \left(\frac{B_V}{30}\right)^{-3}, B_V = \sqrt{A_V} \tag{12}$$

Damping coefficients C_p, C_s are also calculated using Figure 6 as follows:

$$C_p = \rho \cdot A \cdot \sqrt{\frac{\lambda+2G}{\rho}} = W \cdot A \cdot \sqrt{\frac{\lambda+2G}{W \times 9.81}} = c_p \cdot A \tag{13}$$

$$C_s = \rho \cdot A \cdot \sqrt{\frac{G}{\rho}} = W \cdot A \cdot \sqrt{\frac{G}{W \times 9.81}} = c_s \cdot A \tag{14}$$

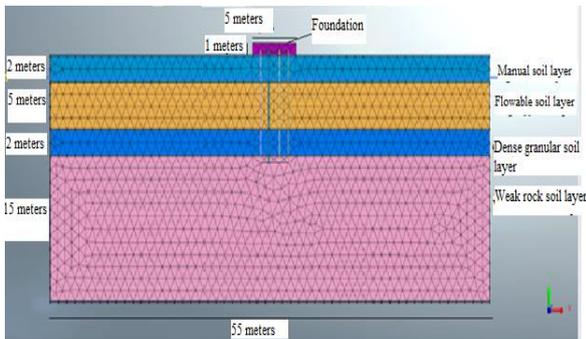


Figure 5. Numerical model built in Midas

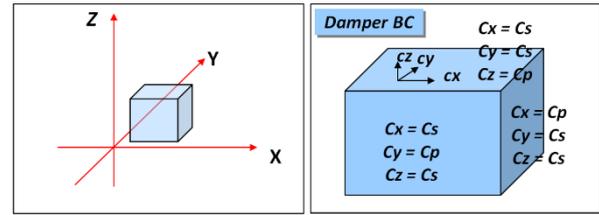


Figure 6. Components of the damping coefficient in the general coordinate system (10)

$$\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}, G = \frac{E}{2(1+\nu)} \tag{15}$$

The seismic risk estimation and design criteria for x, y, z directions are presented based on the acceleration results of artificial maps in the region, as well as the history of accelerations that are proposed as seismic design criteria. Design engineers should understand how loads are carried from the to piles and soil media so they can predict performance like settlement, bearing pressure enhancement, and borrowing capacity rate, as well as behavior like displacement and load sharing across (10). (Experimental Study of Lateral Loading on Piled Raft Foundations on Sandy Soil). Therefore, the acceleration history is selected based on this critical seismic profile in the software. It is worth noting that the vibration period is reduced to 10 seconds by ignoring the low-amplitude vibrations to shorten the analysis period or run the numerical model. In this way, the model is ready to be analyzed under the influence of the benchmark acceleration history in the main coordinate directions. It should be noted that a simple static analysis is necessary for the mesh control. Further, there is a need for eigenvalues and natural vibration periods of different modes of the model, which is also obtained by taking an Eigenvalue analysis for dynamic analysis settings. Additionally, the Rayleigh damping value of 0.05 is used in the dynamic analysis (11).

3. NUMERICAL MODELING AND ANALYSIS RESULTS

This research introduces a set of records of earthquakes in the near field with a pulse caused by the progressive directionality in the time history of the speed of earth's movement. Assuming the problem that the location of accelerograms is on the bedrock, 10 accelerograms (5 near-field accelerograms and 5 far-field accelerograms), recorded on hard rock, are used for analysis. Table 1 presents the specifications of these accelerometers.

3. 1. Analysis Parameters In this research, 4 parameters were subjected to sensitivity analysis. The analysis parameters included the pile length, pile diameter, pile distance, and pile load. For sensitivity

TABLE 1. Records of earthquakes in the near area

Earthquake	Year	Station	Tp	PGV	Mw
San Fernando	1971	Pacoima Dam	1.6	28	6.6
Coyote Lake	1979	Gilroy Array	1.2	4	5.7
Morgan Hill	1984	Coyote Lake Dam	1	27	6.2
Cape Mendocino	1992	Petrolia	3	47	7
Northridge -01	1994	Pacoima Dam	0.5	34	6.7

analysis, each of the four parameters are studied three times. Table 2 presents the number of far-field analyses. The analysis parameters of the pile groupn are summarized in Table 3.

3. 2. Geotechnical Specifications of the Model

The Moore-Coulomb model was used for non-liquefiable layers, and the UBCSAND constitutive model was used for liquefiable soil. The following tables present the parameters for each layer.

The model used in the free response analysis of the soil in this article is a finite element model in which the soil layers are modeled by the FourNodeQuadUP element and in two modes of pressure-dependent materials and pressure-independent materials with multifaceted yield surfaces. These materials can be used in two-dimensional and three-dimensional elements. Explanations related to these two items are provided below.

nDMaterial PressureDependMultiYield \$stag \$nd \$rho \$refShearModul \$refBulkModul \$frictionAng \$peakShearStra \$refPress \$pressDependCoe \$PTAng \$contrac \$dilat1 \$dilat2 \$liquefac1 \$liquefac2 \$liquefac3 where in:

\$stag number of material used
 \$nd the number of dimensions that can be 2 for plane strain and 3 for three-dimensional analysis.
 \$rho is the saturated mass density of the substance
 \$refShearModul soil shear module
 \$refBulkModul is the bulk modulus of the soil obtained from equation 25-4

\$frictionAng internal soil friction angle

\$peakShearStra octahedral shear strain where the maximum shear strength occurred and from the relation

$$E_b = \frac{2}{3} G \frac{1+\nu}{1-2\nu} \tag{16}$$

$$\gamma = \frac{2}{3} \left[(\epsilon_{xx} - \epsilon_{yy})^2 + (\epsilon_{yy} - \epsilon_{zz})^2 + (\epsilon_{xx} - \epsilon_{zz})^2 + 6\epsilon_{xy}^2 + 6\epsilon_{yz}^2 + 6\epsilon_{xz}^2 \right]^{1/2} \tag{17}$$

\$refPress Confinement average pressure where , and are defined.

\$pressDependCoe is a non-negative coefficient that represents the G and B variables as a function of the effective confinement coefficient, which is based on the following relationship.

$$G = G_r \left(\frac{P'}{P_r} \right)^d \quad B = B_r \left(\frac{P'}{P_r} \right)^d \tag{18}$$

\$PTAng phase change angle in degrees

\$contrac is a non-negative coefficient that represents the volume reduction or pore water pressure production.

\$dilat1, 2 coefficients that are used as indicators of volume increase or expansion.

\$liquefac1 parameters that control the liquefaction mechanism.

3. 3. Calibration of Materials

The Morcoulomb pattern behavior was used to calibrate the model materials, using the three-axis experiemntasl results of

TABLE 2. Records of earthquakes in far area

Earthquake	Year	Station	Tp	PGV	Mw
San Fernando	1971	Lake Hughes	-	10	6.6
San Fernando	1971	Santa Anita Dam	-	11	6.6
Manjil Iran	1990	Abbar	-	29	7.4
Taiwan SMARTI	1986	SMART1 E02	-	36.7	7.3
Nahanni	1985	Site 1	-	15	6.76

TABLE 3. Parameters used for non-liquefacted and liquefacted materials

	(kN/m ²)	(kN/m ³)	Sat (kN/m ²)	C (kN/m ²)	Φ	K(cm/s)	v	Ψ
EMBANKEMENT	4000	18	20	15	25	0.1	0.35	1
WEATHER ROCK	1000000	20	51	35	33	0.1	0.3	7
SOFT ROCK	900000	24	25	150	37	0.01	0.01	7

Material	E(kN/m ²)	Y(kN/m ³)	Y Sat (kN/m ²)	P0 (kN/m ²)	ELASTIC SHEAR MODULUS	ELASTIC SHEAR MODULUS
LIQUFACTION LAYER	60000	20	21	100	1069	0.5

the calibration. To calibrate, an axially symmetric two-dimensional model was created in the software with all experimntasl conditions for performing three-axis tests on model materials. In this model, a triaxial cylinder is used with a diameter of 15 cm, a height of 30 cm, with 48 nodes, and 33 elements. Figure 6 shows the axial symmetry section of the triaxial test in MIDAS GTS. The calibration curves of materials and product materials in the CD test under 6 kg/cm² are shown in Figures 7 and 8, respectively.

3. 4. Length Change Results This section examines the results of changes in pore water pressure

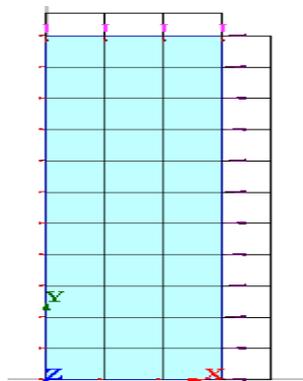


Figure 6. Triaxial geometry pattern in MIDAS GTS

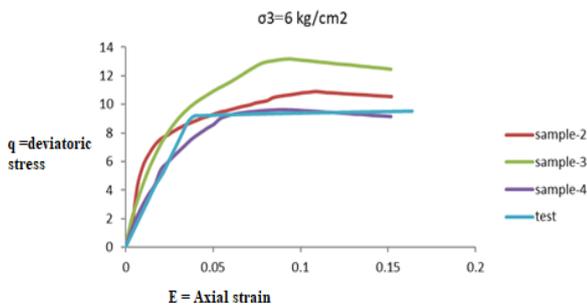


Figure 7. Calibration of materials in the CD test under 6 kg/cm² all-round stress

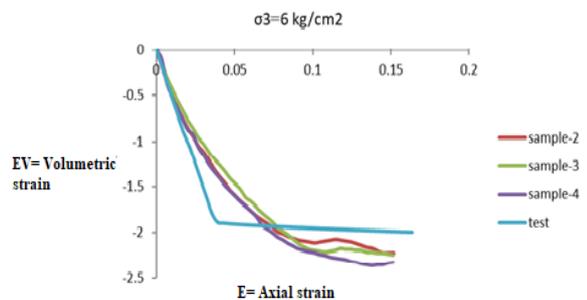


Figure 8. Calibrated product of materials in CD test under 6kg/cm² all-round stress

ratio, as well as the liquefaction criterion ratio for different acceleration maps. Figures 9 and 10 show the acceleration maps of 1 and 5, respectively.

3. 4. 1. Bore Water Pressure Results With changes in the length of the pile (5, 8 and 10) meters for 5 distant earthquake records, the pore water pressure decreases with an increase in the pile length. Further, the pressure changes start from the 4th second onwards for most of the acceleration maps. The rate of changes decreases with increasing the candle length.

3. 4. 2. The Results of the Ratio of Shear to Vertical Stress (Liquefaction Criterion) With changes in the pile length (5, 8, and 10 meters) for 5 distant earthquake records, the ratio of the maximum stress decreases with an increase in the pile length. Further, the stress changes start from the 4th second onwards for most acceleration maps. Additionally, the rate of changes decreases with an increase in the pile length. Based on the diagrams for all deflection maps, the probability that liquefaction will occur is close to zero for a pile with a length of 10 meters. The acceleration map 1 normalized maximum stress ratio is shown in Figure 11.

3. 5. The Results of Changing the Candle Diameter This section examines the results of changes in the pore water pressure ratio and the liquefaction criterion ratio

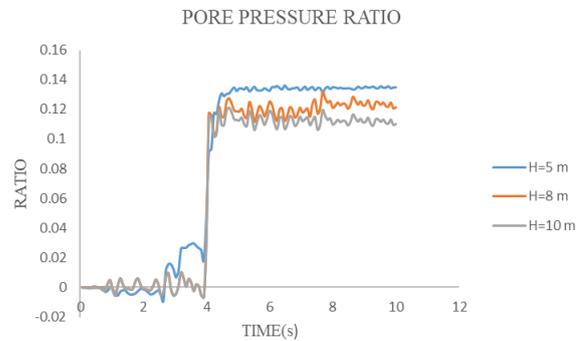


Figure 9. Acceleration map 1

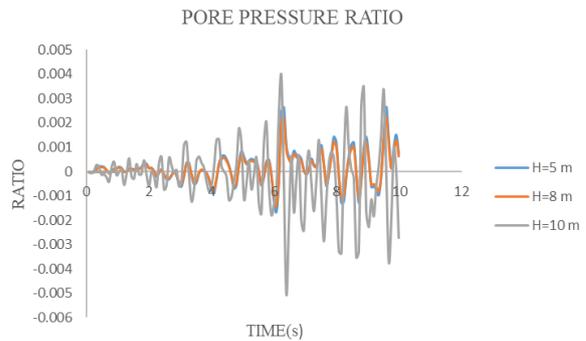


Figure 10. Acceleration map 5

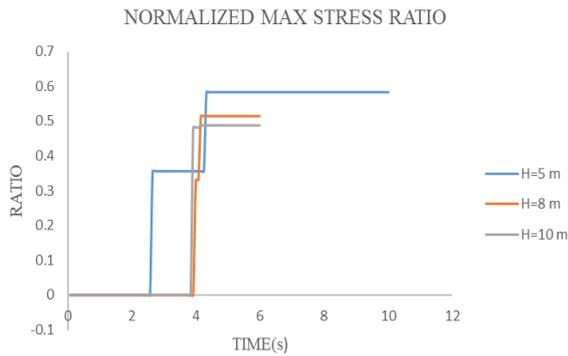


Figure 11. Acceleration map 1

for the acceleration of different mappings with changes in the pile diameter.

3. 5. 1. Results of Pore Water Pressure With the changes in pile diameter (0.5, 0.8, and 1 meters) for 5 earthquake records in the far area, an increase in the pile diameter raises the pore water pressure, as well as the pressure changes from 4th second onwards for most of the acceleration maps. Additionally, the rate of change increases with an increase in candle diameter. An increase in the ratio of pore water pressure is due to the greater confinement of soil between the piles with an increase in the pile diameter. The acceleration map 1 pore pressure ratio is shown in Figure 12.

3. 5. 2. The Results of the Ratio of Shear to Vertical Stress (Liquefaction Criterion) With the changes in pile diameter (0.5, 0.8 and 1) meters for 5 remote area earthquake records, it can be seen that with the increase in pile diameter, the maximum stress ratio has a decreasing trend, and also the stress changes from the 4th second onwards for most of the acceleration maps start Also, the rate of change decreases with the increase of the diameter of the pile. Also, by checking the graphs for all acceleration maps, it can be seen that the probability of liquefaction is less for piles with a length of 10 meters.

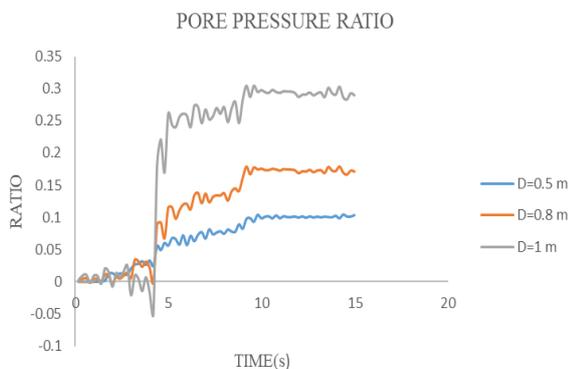


Figure 12. Acceleration map 1

For the acceleration mapping number 4, diameter 0.8 has given a lower stress ratio than the pile with a diameter of 1 meter. This inconsistency in the results is due to the convergence operator of the analysis, which was difficult for the navigation map number 4 due to the acceleration of the analysis convergence mapping, and this contradiction is due to the difference in the convergence of the analysis. The acceleration map 1 normalized maximum stress ratio for 0.5, 0.8 and 1m is shown in Figure 13.

3. 6. The Results of Changing the Candle Distance

This section examines the results of changes in pore water pressure ratio, as well as the liquefaction criterion ratios for different acceleration mappings under the effect of pile distance changes.

3. 6. 1. Bore Water Pressure Results With the changes in pile distance (1.5, 2.5, and 5 meters) for 5 distant earthquake records, an increase in pile distance decreases the pore water pressure. Further, the pressure changes from 4th second onwards for most of the acceleration maps. The rate of change is constant with increasing candle distance. The reduction of the pore water pressure ratio is due to the reduction of soil confinement between the piles by increasing the pile distance. The acceleration maps 1 and 5 pore pressure ratio for 1.5, 2.5 and 5m are shown in Figures 14 and 15, respectively.

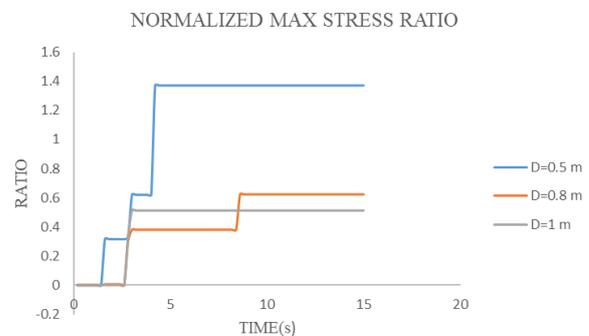


Figure 13. Acceleration map 1

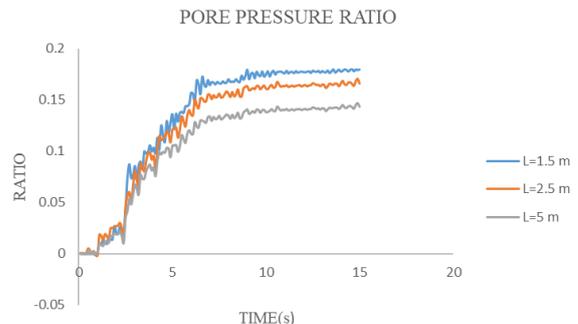


Figure 14. Acceleration map 1

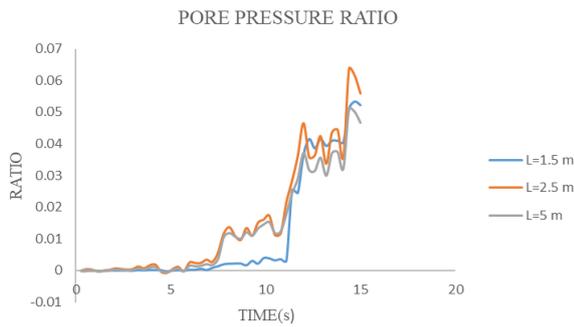


Figure 15. Acceleration map 5

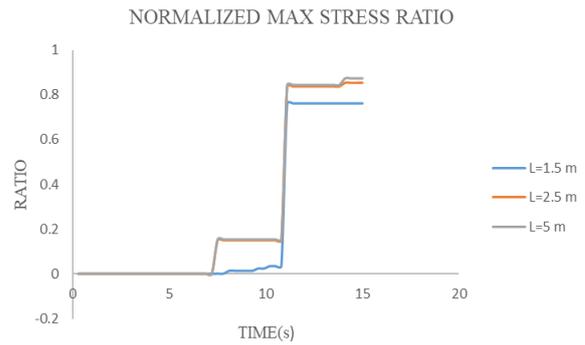


Figure 17. Acceleration map 5

3. 6. 2. The Results of the Ratio of Shear to Vertical Stress (Liquefaction Criterion)

Based on changes in pile spacing (1.5, 2.5, and 5 meters) for 5 earthquake records in the far area, an increase in pile spacing raised the ratio of the maximum stress, and also the stress changes from the 4th second onwards for most of the acceleration maps. Further, the rate of changes with increasing pile distance is almost constant. Based on the diagrams for all acceleration maps, short pile distance reduces the probability of liquefaction for almost 5 earthquakes, but the rate of decrease in the probability of liquefaction is lower with increasing distance from 2 to 5 meters, indicating that increasing the pile distance beyond a certain limit has no effect on liquefaction. The acceleration maps 1 and 5 normalized maximum stress ratio are shown in Figures 16 and 17, respectively.

3. 7. Results of Changing the Pile Load

This section presents the results of changes in pore water pressure ratio, as well as liquefaction criterion ratio for different acceleration mappings under the effect of load changes on the piles.

3. 7. 1. Results of Bore Water Pressure

Based on changes in the load on the pile foundation (1, 5, and 10 tons) for 5 records of far-field earthquakes, an increase in the pile load decreases the pore water pressure, as well as the pressure changes from the 4th second onwards for

most of the acceleration maps. Additionally, the rate of change is constant with an increase in entering the candle. The reduction of the pore water pressure ratio is due to the increase in effective soil stress between the piles with an increase in the pile load. The acceleration map 5 pore pressure ratio is shown in Figures 18.

3. 7. 2. The Results of the Ratio of Shear to Vertical Stress (Liquefaction Criterion)

Based on changes in the load on the pile foundation (1, 5, and 10 tons) for 5 earthquake records in the far area, an increase in the pile load decreases the ratio of the maximum stress, and the stress changes start from the 4th second onwards for most of the acceleration maps. Additionally, the rate of change increases with an increase in the pile load. Based on the graphs for all acceleration maps, an increase in the pile load reduces the probability of liquefaction for almost 5 earthquakes, and these changes increase with the load. The acceleration map 1 normalized maximum stress ratio is shown in Figure 19.

3. 8. Incremental Dynamic Analysis

The incremental dynamic analysis method is a non-linear dynamic analysis that can determine the level of damage by the earthquake intensity. In this method, the scaling of the acceleration maps of past earthquakes is used in such a way that it covers the structure behavior from a linear

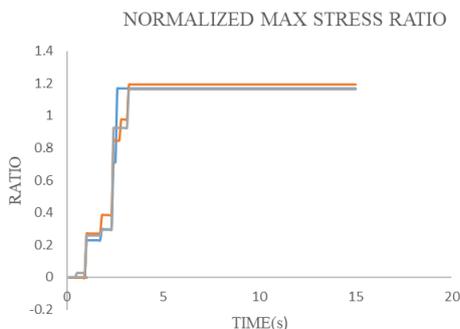


Figure 16. Acceleration map 1

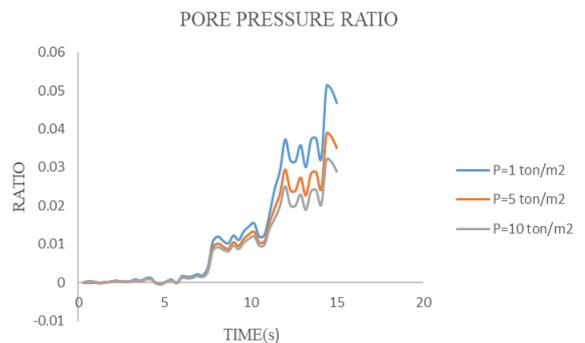


Figure 18. Acceleration map 5

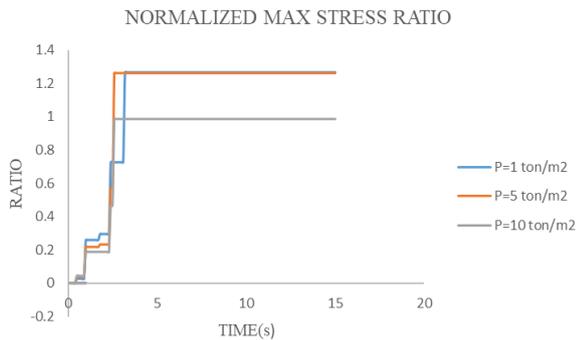


Figure 19. Acceleration map 1

elastic to a collapse stage of the structure. In this method, which is also developed to consider the effect of inherent uncertainty in earthquakes, the seismic response of structures is evaluated to take into account the uncertainty of frequency and spectral shape of earthquakes. Further, an appropriate number of earthquake records should be used. Each earthquake record is scaled in such a way to cover a suitable range of earthquake intensities and structure behavior from elastic limit to failure (24, 25).

To perform the analysis, first, the seismic intensity parameter is considered, for example, the maximum ground acceleration or spectral acceleration corresponding to the first mode of the structure, from a very small value for the occurrence of elastic behavior in the structural model under dynamic analysis to a certain level of seismic intensity to achieve the desired failure limit, with an appropriate algorithm. Each time, this scale factor is applied to the earthquake record, and the structure under the influence of that record becomes a dynamic time history analysis. At the end of each analysis stage, the level of damage intensity corresponding to the earthquake intensity level, in which the analysis is performed, is recorded. Finally, under each scaled record, a response curve is obtained against the intensity that is referred as the IDA single record curve.

The structure behavior under an earthquake record cannot be generalized for all records because the result is not general and cannot be considered as the general behavior of the structure under all conditions at different times and earthquakes. Several single-record curves are needed to conduct probabilistic and statistical studies to achieve the correct result of the structure's behavior; hence, this operation is performed for all desired records, and the resulting single record curves are all drawn in a seismic intensity-damage intensity coordinate system, called IDA curves. Figure 20 shows the IDA curves for a five-story structure.

3. 8. 1. Incremental Dynamic Analysis (IDA) Results for Different Pile Diameters

Even though the structure's behavior under an earthquake

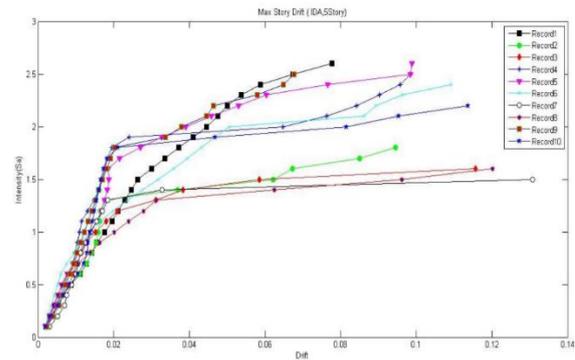


Figure 20. A set of IDA curves for a five-story structure

record is important but cannot be generalized for all records because the result does not have a general state and cannot be considered the general behavior of the structure under all conditions at different times and earthquakes, in other words, for evaluating the seismic behavior of the structure. Several single curves are necessary to achieve the correct result of behavior. Based on IDA curves, there are 15 records for the candles in the figures below. Based on the curves, all stages of pile group behavior under earthquake can be fully observed from elastic to collapse limit, and general instability. At the beginning of all curves, there is an elastic region as a straight line and a part of the IDA curve as a common section of all curves. After this area, the nonlinear behavior is visible in the curves when the curve leaves the a straight line. Further, a decrease in hardness or its increase show the nonlinear behavior. Therefore, this set of curves can get a general view of soil and pile behavior from full elastic limit to full failure. A comparison of the behavior of piles at different diameters and on the same type of ground incates that an increase in the diameter of piles leads to the entry of the pile group into the nonlinear region and a higher capacity. Figure 21 illustrates the IDA diagram for a pile in 1 m diameter.

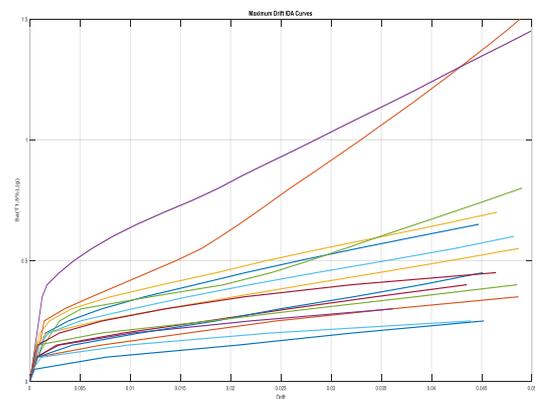


Figure 21. IDA diagram for a pile in a diameter of 1 m

3. 8. 2. The probability of Liquefaction for Different Pile Diameters under Far-field Records

Fragility curves are used to extract the probability of limit states from the output of IDA. To draw these curves, IM values corresponding to the occurrence of the desired limit state are arranged in descending order for all records. Then, the probability of occurrence of the limit state in the pile is calculated using the sorted values, and its curve is drawn against IM. Based on the curve, an increase in the diameter of piles at a fixed level of seismic intensity decreases the probability of collapse or lack of a performance level. The fragility curves for different diameters is shown in Figure 22.

Based on the results, an increase in pile diameter decreases the probability of liquefaction so that the probability of pile group liquefaction and collapse decreases by about 33% when the diameter of piles increases from 0.5 to 1 meters. It is also significant. When the diameter increases from 0.5 to 0.8 meters, the probability of liquefaction decreases by about 12%. Table 4 summarized numerical analysis of liquefaction probability.

3. 8. 3. Results of IDA for Different Pile Distances

The IDA method greatly depends on the acceleration map and the desired structure. Therefore, the same models sometimes show different behaviors under the influence of different acceleration maps. The occurrence of a specific and linear region in seismic intensities is common among all IDA curves. This linear behavior ends with the appearance of the first submission in the candlestick. The slope of this linear area is an indicator of elastic stiffness of the structure and is different under the effects of different records on the structure. After

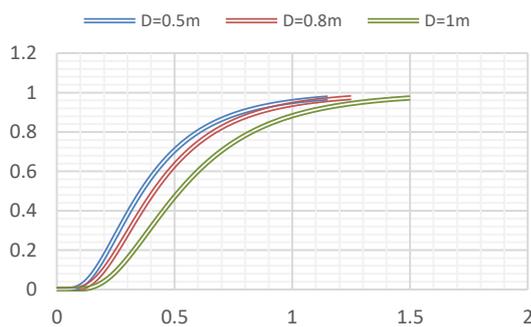


Figure 22. Fragility curve for different diameters

TABLE 4. Numerical study of liquefaction probability

	D=0.5m	D=0.8m	D=1m
Sa(T1,g)	0.5	0.5	0.5
Probability%	70	62	47

crossing the linear region, the behavior of the IDA curve is different and varied. After the first yield in the members in some cases, a sharp drop in the curve occurs and the graph quickly moves towards larger values of the failure criterion. In other cases, the curve becomes steeper due to hardening. In some cases, the curve oscillates around its initial slope in a way that expresses successive hardening and softening. This phenomenon occurs due to the irregular nature of the record so that the acceleration or the force resulting from the earthquake causes damage at a certain time of the acceleration map. Based on the occurrence of yielding in some members at other times, this acceleration decreases so much and turns the curve into a lower failure criterion or even reverses the direction of the curve. Comparing the behavior of piles at different distances (3d, 5d, 10d) equivalent to 1.5, 2.5, and 5 meters on the same type of ground indicates that an increase in distance between the piles leads to the quick entry of the pile group into the nonlinear region at a lower capacity. The IDA diagram for 5d distance is shown in Figure 23.

3. 8. 4. The Probability of Liquefaction Occurrence for Different Pile Distances under Far Field Records

The fragility curves are used to extract the probability of occurrence of limit states from the output of IDA. Based on the curves, an increase in the distance of piles at a fixed level of seismic intensity increases the probability of collapse or failure to meet the performance level. Figure 24 shows analytical analysis of the fragility curve at different distances.

The results indicate that the possibility of liquefaction increases with an increase in the pile distance so that the probability of liquefaction and collapse of the pile group increases by about 14% when the pile distance increases from 1.5 to 5 meters. The probability of liquefaction increases by 7% when the distance increases from 1.5 to 2.5 meters.

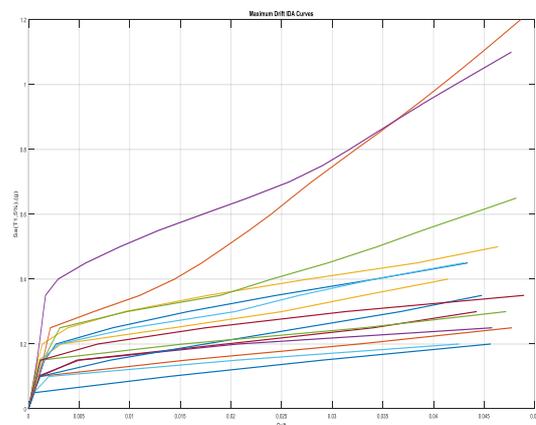


Figure 23. IDA diagram for 5d distance

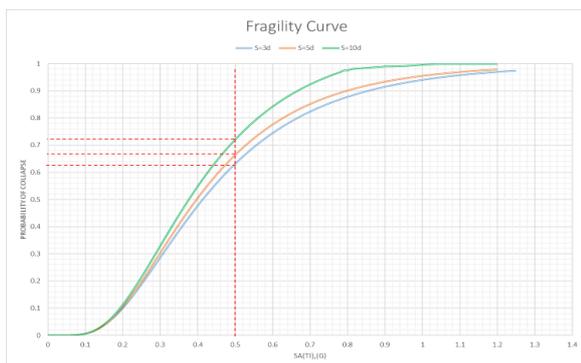


Figure 24. Analytical investigation of the fragility curve at different distances

4. CONCLUSION

With changes in the length of the pile (5, 8 and 10) meters for 5 distant earthquake records, it can be seen that with the increase in the length of the pile, the pore water pressure has a decreasing trend, and also the pressure changes start from the 4th second onwards for most of the acceleration maps. The rate of change decreases with increasing candle length. Also, the stress changes start from the 4th second onwards for most of the acceleration maps, and the rate of change decreases with the increase in the length of the pile. In addition, by checking the graphs for all the swing maps, it can be seen that for a pile with a length of 10 meters, the probability that Liquefaction occurs close to zero. By checking the graphs for all acceleration maps, it can be seen that the probability of liquefaction is less for piles with a length of 10 meters. For the acceleration mapping number 4, diameter 0.8 has given a lower stress ratio than the pile with a diameter of 1 meter. This inconsistency in the results is due to the convergence of the analysis, which for the number 4 floating map, according to the type of acceleration, the convergence of the analysis was difficult, and this contradiction is due to the difference in the convergence of the analysis. Also, the pressure changes from the 5th second onwards for most of the acceleration maps. Likewise, the rate of change decreases with increasing candle length. The rate of change decreases with the increase in the length of the pile. Also, by checking the graphs for all deflections, it can be seen that for a pile with a length of 10 meters, the probability that liquefaction will occur is very low. Comparing the behavior of piles with different diameters and on the same type of ground, it can be said that by increasing the diameter of the piles, the pile group enters the nonlinear region later and has more capacity. As the diameter of the pile increases, the possibility of liquefaction decreases, so that the possibility of liquefaction and collapse of the pile group decreases by about 33% when the diameter of the piles increases from 0.5 to 1 meter, which is a significant number. Also, when the diameter increases

from 0.5 meters to 0.8 meters, the probability of liquefaction decreases by about 12%. Comparing the behavior of piles with different distances (3d, 5d, 10d) is equivalent to (1.5, 2.5 and 5) meters on the same type of ground, it can be said that by increasing the distance between the piles, the pile group enters the nonlinear region earlier and has a lower capacity.

5. REFERENCES

- Zomorodian R, Soltani F, Sivandi-Pour A, Noroozinejad Farsangi E. Effect of foundation flexibility on the seismic performance of a high-rise structure under far-field and near-field earthquakes. *International Journal of Engineering, Transactions A: Basics*. 2021;34(7):1611-22. [10.5829/ije.2021.34.07a.06](https://doi.org/10.5829/ije.2021.34.07a.06)
- Kamrani Moghaddam P, Manafpour A. Effects of Far-and Near-Field Multiple Earthquakes on the RC SDOF Fragility Curves Using Different First Shock Scaling Methods. *International Journal of Engineering*. 2018;31(9):1505-13.
- Pan R, Xu C, Jia K, Dou P, Kamran I. Large-scale shaking table test and three-dimensional numerical simulation research on earthquake seismic failure response of laterally spreading site-pile group-superstructure system. *Bulletin of Earthquake Engineering*. 2023;21(10):4789-819. <https://doi.org/10.1007/s10518-023-01713-y>
- Sahare A, Ueda K, Uzuoka R. Influence of the sloping ground conditions and the subsequent shaking events on the pile group response subjected to kinematic interactions for a liquefiable sloping ground. *Soil Dynamics and Earthquake Engineering*. 2022;152:107036. <https://doi.org/10.1016/j.soildyn.2021.107036>
- Chandiwala A, Vasanwala S. Experimental study of lateral loading on piled raft foundations on sandy soil. *International Journal of Engineering*. 2023;36(1):28-34. [10.5829/ije.2023.36.01a.04](https://doi.org/10.5829/ije.2023.36.01a.04)
- Mohammadizadeh M, Bostani M, Farkat E, Mohammadzadeh N. Soil-Structure Interaction Effect on seismic response of Low-and mid-rise steel moment frames equipped with Pall friction damper. *Amirkabir Journal of Civil Engineering*. 2023;55(7):1337-62. [10.22060/ceej.2023.20880.7559](https://doi.org/10.22060/ceej.2023.20880.7559)
- Motamed R, Towhata I, Honda T, Tabata K, Abe A. Pile group response to liquefaction-induced lateral spreading: E-Defense large shake table test. *Soil Dynamics and Earthquake Engineering*. 2013;51:35-46. <https://doi.org/10.1016/j.soildyn.2013.04.007>
- Loli M, Knappett JA, Brown MJ, Anastasopoulos I, Gazetas G. Centrifuge modeling of rocking-isolated inelastic RC bridge piers. *Earthquake engineering & structural dynamics*. 2014;43(15):2341-59. <https://doi.org/10.1002/eqe.2451>
- Haeri SM, Rajabigol M, Salaripour S, Kavand A, Sayyaf H, Afzalsoltani S, et al. Effects of physical modeling boundary conditions on the responses of 3x3 pile groups to liquefaction induced lateral spreading *Bulletin of Earthquake Engineering*. 2023;21(5):2469-502. <https://doi.org/10.1016/j.soildyn.2023.107915>
- Li W, Chen Y, Stuedlein AW, Liu H, Zhang X, Yang Y. Performance of X-shaped and circular pile-improved ground subject to liquefaction-induced lateral spreading. *Soil Dynamics and Earthquake Engineering*. 2018;109:273-81. <https://doi.org/10.1016/j.soildyn.2018.03.022>
- Casagrande A, editor *The determination of the pre-consolidation load and its practical significance*. Proc 1st Int Conf Soil Mech; 1936.

12. Casagrande A. Characteristics of cohesionless soils affecting the stability of slopes and earth fills. *J Boston Society of Civil Engineers*. 1936;23(1):13-32.
13. Wang Y, Orense RP. Numerical analysis of seismic performance of inclined piles in liquefiable sands. *Soil Dynamics and Earthquake Engineering*. 2020;139:106274. <https://doi.org/10.1016/j.soildyn.2020.106274>
14. Wang X, Ye A, Shang Y, Zhou L. Shake-table investigation of scoured RC pile-group-supported bridges in liquefiable and nonliquefiable soils. *Earthquake Engineering & Structural Dynamics*. 2019;48(11):1217-37. <https://doi.org/10.1002/eqe.3186>
15. Jia K, Xu C, El Naggar MH, Zhang X, Du X, Dou P, et al. Large-scale shake table testing of pile group-bridge model in inclined liquefiable soils with overlying crusts. *Soil Dynamics and Earthquake Engineering*. 2022;163:107555. <https://doi.org/10.1016/j.soildyn.2022.107555>
16. Li W, Stuedlein AW, Chen Y, Liu H, Cheng Z. Response of pile groups with X and circular cross-sections subject to lateral spreading: 3D numerical simulations. *Soil Dynamics and Earthquake Engineering*. 2019;126:105774. <https://doi.org/10.1016/j.soildyn.2019.105774>
17. Ghorbani A, Hasanzadehshooiili H, Somti Foumani MA, Medzvieckas J, Kliukas R. Liquefaction Potential of Saturated Sand Reinforced by Cement-Grouted Micropiles: An Evolutionary Approach Based on Shaking Table Tests. *Materials*. 2023;16(6):2194. [10.3390/ma16062194](https://doi.org/10.3390/ma16062194)
18. Lombardi D, Bhattacharya S. Evaluation of seismic performance of pile-supported models in liquefiable soils. *Earthquake Engineering & Structural Dynamics*. 2016;45(6):1019-38. [10.1002/eqe.2716](https://doi.org/10.1002/eqe.2716)
19. Wang Y, Orense RP. Numerical Investigation of Inclined Piles under Liquefaction-Induced Lateral Spreading. *Geotechnics*. 2023;3(2):320-46. <https://doi.org/10.3390/geotechnics3020019>
20. Knappett J, Madabhushi S. Liquefaction-induced settlement of pile groups in liquefiable and laterally spreading soils. *Journal of geotechnical and geoenvironmental engineering*. 2008;134(11):1609-18. [https://doi.org/10.1061/\(ASCE\)1090-0241\(2008\)134:11\(1609\)](https://doi.org/10.1061/(ASCE)1090-0241(2008)134:11(1609))
21. Mao W, Liu B, Rasouli R, Aoyama S, Towhata I. Performance of piles with different configurations subjected to slope deformation induced by seismic liquefaction. *Engineering geology*. 2019;263:105355. <https://doi.org/10.1016/j.enggeo.2019.105355>
22. Saha P, Horikoshi K, Takahashi A. Performance of sheet pile to mitigate liquefaction-induced lateral spreading of loose soil layer under the embankment. *Soil Dynamics and Earthquake Engineering*. 2020;139:106410. <https://doi.org/10.1016/j.soildyn.2020.106410>
23. Moghaddam PK, Manafpour A. Effects of Far-and Near-Field Multiple Earthquakes on the RC Single Degree of Freedom Fragility Curves Using Different First Shock Scaling Methods. *International Journal of Engineering, Transactions C: Aspects*. 2018;31(9):1505-13. [10.5829/ije.2018.31.09c.05](https://doi.org/10.5829/ije.2018.31.09c.05)
24. Iraj A. Evaluating Finn-Byrne Model in Liquefaction Analysis of Quay Wall and Cantilevered Retaining Wall Models. *International Journal of Engineering, Transactions C: Aspects*. 2023;36(6):1075-91. [10.5829/ije.2023.36.06c.06](https://doi.org/10.5829/ije.2023.36.06c.06)
25. Razavi SA, Siahpolo N. A New Optimal Correlation for Behavior factor of EBFs under Near-fault Earthquakes using Artificial Intelligence Models. *International Journal of Advanced Structural Engineering*. 2022;12(2):631-45. [10.1007/ijase.2022.698628](https://doi.org/10.1007/ijase.2022.698628)

COPYRIGHTS

©2024 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 از پژوهش حاضر بررسی عملکردهای گروه شمع در خاک مستعد روانگرایی تحت تأثیر زلزله‌های حوزه دور و نزدیک و پاسخ‌های سازه از قبیل تاریخچه شتاب، جابجایی و گشتاورهای خمشی می‌باشد. از روش تخمین طیف پاسخ‌ساختگاه با توجه به فاصله، بزرگا، شرایط ساختگاهی و مکانیزم گسلش و مقیاس کردن طیف پاسخ یک زلزله منتخب ثبت شده یا مصنوعی، متناسب با آن استفاده گردید. با استفاده از نرم‌افزارهای *opensees*، *Midas Gts*، *Peer* در انتخاب شتاب‌نگاشت‌های زلزله برای حوزه دور و حوزه نزدیک از مرکز لرزه‌نگاری *Peer* مرکز شتاب‌نگاری کشور و مرکز تحقیقات راه و مسکن برای انتخاب و اصلاح شتاب‌نگاشت‌ها استفاده شد. نتایج تحقیق نشان داد که افزایش قطر شمع احتمال روانگرایی را کاهش می‌دهد به‌گونه‌ای که با افزایش قطر شمع از ۰.۵ متر به ۱ متر، احتمال روانگرایی و فروپاشی گروه شمع حدود ۳۳ درصد کاهش می‌یابد. علاوه بر این، با افزایش فاصله شمع‌ها، احتمال روانگرایی افزایش می‌یابد به طوری که احتمال روانگرایی گروه شمع‌ها و فروپاشی آن با افزایش فاصله شمع‌ها از ۱.۵ به ۵ متر، حدود ۱۴ درصد افزایش می‌یابد. در آخر با استفاده از نتایج حاصله از مدل‌سازی عددی و با نرم‌افزارهای کد نویسی الگویی مناسب برای گروه شمع در برابر روانگرایی برای زلزله‌های حوزه دور و حوزه نزدیک ارائه گردید.



Hydrodynamic Modeling and Evaluation of Partial Substitution of Cushion Gas During Creation of Temporary Underground Gas Storage in an Aquifer

E. A. Smirnova*, L. A. Saychenko

Department of Oil and Gas Engineering, Empress Catherine II Saint-Petersburg Mining University, Saint-Petersburg, Russia

PAPER INFO

Paper history:

Received 17 December 2023

Received in revised form 24 January 2024

Accepted January 2024

Keywords:

Associated Gas
Underground Gas Storage
Cushion Gas
Aquifer

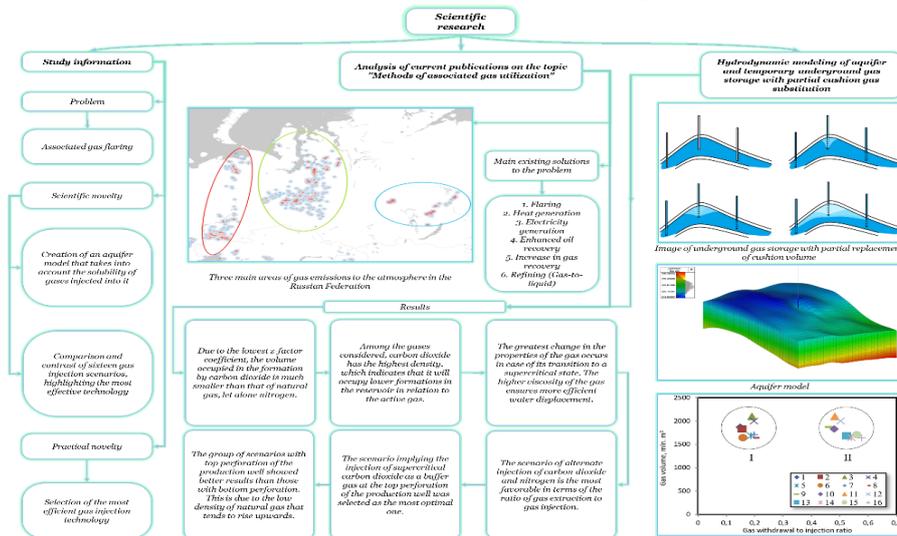
ABSTRACT

The flaring of associated gas remains a problem for oil and gas fields that are difficult to access and remote from the infrastructure. Active development of oil and gas production in Eastern Siberia has led to the fact that transportation capacities cannot keep up with field development. Increased flaring of associated gas leads to a significant increase in greenhouse gases such as carbon dioxide and methane. A possible solution to this problem is to store gas in the aquifer of the field for its future sale and monetization through the main gas pipeline. This paper analyzes the main technologies of associated gas utilization and reveals the problem of remoteness from gas transportation infrastructure of hard-to-reach fields. An effective technology to solve this problem is the creation of temporary underground storage of associated gas in the aquifer of the field. The results of hydrodynamic modeling of realization of this technology with partial replacement of cushion gas showed that joint injection of carbon dioxide and nitrogen before hydrocarbon gas allows to increase the ratio between produced and injected gas, which indicates its greater efficiency. It is recommended that in order to implement the technology, when selecting a geological injection site, to focus on aquifers with a temperature above 31.2°C, which will allow carbon dioxide to remain in a supercritical state in reservoir conditions.

doi: 10.5829/ije.2024.37.07a.02

Graphical Abstract

Hydrodynamic modeling and evaluation of partial substitution of cushion gas during creation of temporary underground gas storage in an aquifer



*Corresponding Author Email: ilindog98@mail.ru (E.A. Smirnova)

1. INTRODUCTION

Energy, raw materials-related, and environmental issues are currently acquiring a global scale, which determines the relevance of the search for technical, organizational, and economic solutions for the rational use of all kinds of natural energy resources, including associated gas (1).

Associated gas (AG) is a gaseous mixture consisting of methane, ethane, propane, butane, and other hydrocarbon compounds. Associated gas flaring and emissions into the atmosphere result in a marked increase in greenhouse gases (2). According to the International Energy Agency (IEA), emissions from the world's energy sector are steadily high. In 2022, 37 Gt of carbon dioxide was recorded, which is 1% above the 2019 level (3). It is also worth noting that a fast and sustained reduction in methane emissions is key to combating global warming (4). Gas flaring generates more than 500 million tons of CO₂-equivalent greenhouse gases per year, including both CO₂ and methane emissions.

Over the past nine years, the top seven gas-flaring countries have been Russia, Iraq, Iran, the United States, Algeria, Venezuela, and Nigeria (5). These countries produce 40% of the world's oil annually but account for about two-thirds of the world's associated gas flaring (5).

Russia topped the world ranking of associated gas flaring countries in 2020. At the same time, different trends were observed in different regions of the country: a significant increase in associated gas flaring utilization in Eastern Siberia along with a significant flaring reduction in the Khanty-Mansy and Yamalo-Nenets Autonomous Areas in Western Siberia. The Global Methane Tracker study used satellites to record active outbreaks of gas emissions worldwide from January through August 2022. By superimposing the obtained maps on each other, one can identify three main gas emission zones in the Russian Federation: the coal industry zone and oil and gas zones in Western and Eastern Siberia (Figure 1). Although satellite coverage of Russia is low, it is enough to give an estimate of major emission sources.

Although over 54% of all flared associated gas is located within a radius of 20 km from the nearest pipeline, at least 46% remains, which includes most of the remote and hard-to-reach fields. Active development of oil and gas production in remote regions contributes to an increase in the amount of gas flared (6).

Currently, when natural gas prices are near their all-time highs (7), gas flaring not only exerts a negative impact on climate change and human health but also is an enormous waste of money (8). Sun et al. (9) highlighted two main directions for reducing methane emissions: controlling natural gas demand for future emissions and introducing cost-effective reduction technologies for current emissions.

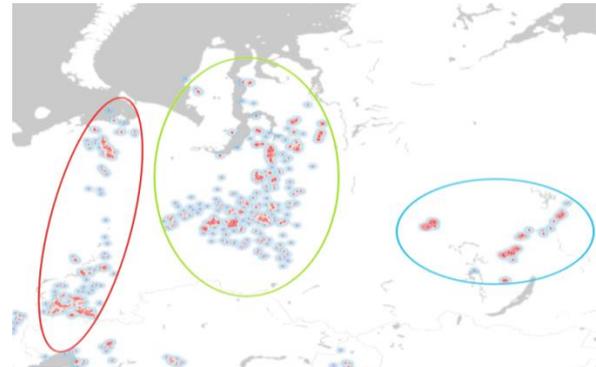


Figure 1. Red zone - coal industry in the Urals, green zone - Western Siberia, blue zone - Eastern Siberia

The main solutions currently implemented at gas fields include:

1. Gas flaring: large energy losses with significant environmental impacts. This practice is still common due to the large volumes of natural gas in the subsoil and its relatively low price (10). According to Federal Law No. 296 "On limiting Greenhouse Gas Emissions" of December 30, 2021, conditions must be created to promote sustainable and balanced development of the country's economy in the context of reducing greenhouse gas emissions.

2. Heat generation: providing heat supply to production buildings and structures of the field, heating of auxiliary and social facilities. Heat-generating units can also be used for heating oil emulsions and formation water.

3. Electricity generation: using gas-turbine or gas-piston power plants, associated gas can be used to cover the field's own needs for electricity. However, any failures in the operation of the gas turbine power plant may result in associated gas flaring (11). Besides, electricity consumption is limited by the needs of the field. If there are no consumers other than the subsoil user company itself, there may be a situation of energy oversupply (12). The payback of this technology is most often achieved by selling electricity to a third-party consumer (13).

4. Increasing oil recovery: application of waterflooding technology often becomes impossible within oil and gas condensate fields due to possible entrapment of gas by water; therefore, it becomes relevant to apply gas methods, more specifically, to enhance oil recovery by injecting carbon dioxide or its mixtures with other gases (for example, as part of associated gas). Water-gas stimulation technologies are also becoming more and more widespread. Nevertheless, it should be noted that in the process of enhancing oil recovery, carbon dioxide accumulation occurs, which requires the introduction of a monitoring and accounting system (14).

5. Increasing gas recovery: hydrodynamic modeling of injection of a mixture of carbon dioxide and hydrogen sulfide, as one of the most harmful associated gas components, in the work (15) showed the realizability of increasing gas production. Gas, due to its high viscosity, density, and solubility, re-pressurizes the formation when injected into the lower part of the formation to displace gas. At the same time, the mixing process is minimized.

6. Gas-to-liquid conversion: converting associated gas into liquid is a complicated and expensive technology, highly susceptible to changes not only in oil prices but also in gas prices. The implementation of this technology is cost-effective not under all market conditions (16). GTL technology makes it possible to convert natural gas into various products such as motor fuel, liquefied gas, lubricating oils, and other chemical products. It is potentially possible to implement GTL technology in the Russian Federation, since the country has significant natural gas reserves. However, specific projects have not been implemented yet. The main reasons for the delay are the high cost and complexity of the technology. Nevertheless, in light of the growing interest in environmentally friendly and energy-efficient fuel production, future development of GTL projects in Russia is possible.

7. Storage: in case of large associated gas volumes and the absence of field connection to the Unified Gas Supply System, the method of temporary underground gas storage is used. This method is based on the theory of creating underground gas storage facilities. Geological objects can be hydrocarbon deposits or aquifers (17). Using this technology within a field has a number of essential advantages: the availability of development history, the state of geological and geophysical exploration, and the confirmed containment capacity of already developed reservoirs.

The choice of one or another technology is based on the reserves of the field, its remoteness from the gas transportation system, the composition of the formation fluid, the stage of field development, already applied associated gas utilization technologies, and many other factors (18).

Currently, a rather new and promising direction is the temporary underground storage of associated gas. As in the case of creating underground gas storages, the objects for future injection can be oil and gas deposits and aquifers. To date, temporary underground gas storage facilities in the gas cap of the developed deposit have been created at the Yurubcheno-Tokhomskoye (19) and Novoportovskoye (18) fields. They also were created in the gas reservoir of the neighboring Vostochno-Messoyakhskoye field, which has not yet been put into development (19), and in the undeveloped gas horizon at the Verkhnechonskoye oil and gas condensate field. Figure 2 shows the classification of temporary

underground gas storage facilities created on the basis of the classification of underground gas storage facilities.

The oil and gas industry of Eastern Siberia has been developing rapidly over the last decade. Owing to the remoteness of the main production centers from the existing infrastructure of the region, a significant increase in the volume of gas flared has become noticeable. The existing gas transportation capacities do not yet cover all fields under development or planned for development.

On the example of oil and gas fields in Eastern Siberia we can realize a preliminary algorithm of object selection, which will allow us to distinguish groups of fields, where it is possible to implement the technology of temporary underground gas storage (Figure 3). The fields that already have a connection to the Power of Siberia gas trunkline (Chayanda field) and that will receive it in the near future (Kovykta field) are not included in this algorithm.

The creation of temporary underground gas storage facilities at Eastern Siberian fields will not only help avoid penalties, but also make it possible to generate future profits from gas sales through the Power of Siberia gas pipeline. The experience of realization of such type of facilities at Verkhnechonskoye, Yurubcheno-Tokhomskoye and other fields will help to find the most suitable solutions for such fields as Alinskoye, Tas-Yuryakhskoye and Srednebotuobinskoye.

In this paper aimed to solve the problem of utilization of huge volumes of associated gas from developed oil and gas condensate fields in Eastern Siberia. The technology of temporary underground storage of gas in the aquifer will be considered. Although the topic of this research is relevant to Eastern Siberia, the analysis and conclusions may find application in future research on natural gas storage in aquifers.

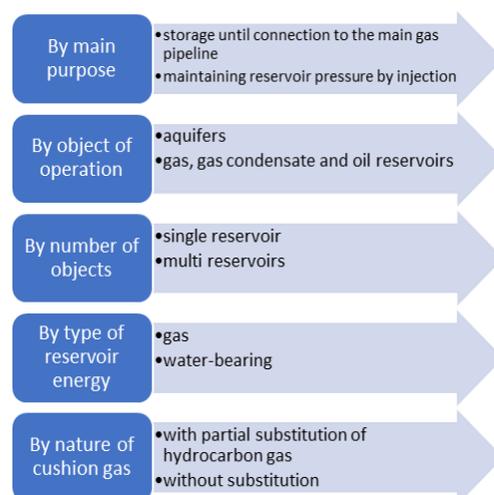


Figure 2. Classification of temporary underground gas storage facilities

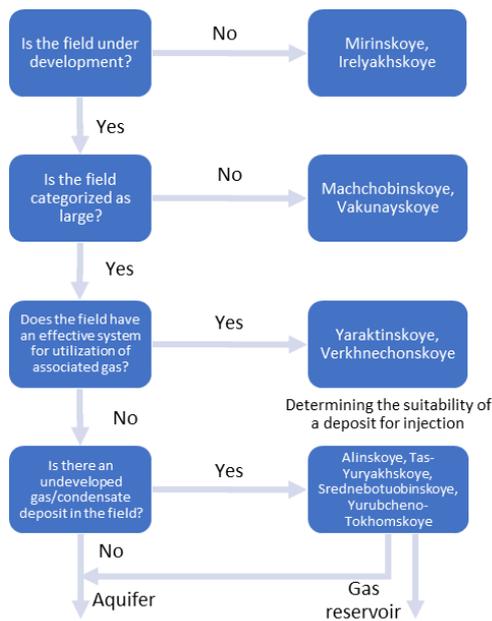


Figure 3. Algorithm for selecting a facility for temporary underground gas storage

2. MATERIALS AND METHOD

Unlike underground gas storage facilities created in depleted oil and gas reservoirs, aquifers do not have a sufficient amount of wells, and there is not enough data to evaluate the reservoir and fluid seal characteristics. The key indicators for evaluating the pressure integrity of the future storage include permeability, breakthrough pressure, mineral and mechanical brittleness, and the thickness and area of caverns (20).

One of the largest capital expenditure items when creating an underground gas storage facility is the cost of cushion gas, which accounts for about 25–30% of the overall storage facility cost (21). It is also worth noting that cushion gas can account for 15–75% of the total gas volume of the storage facility. Therefore, minimizing its volume makes it possible to significantly reduce capital expenditures.

To replace part of the cushion gas in underground storage facilities, it is possible to use inert gases such as nitrogen and helium or carbon dioxide, as well as various mixtures of these gases. However, then a new problem arises: high miscibility of non-hydrocarbon gases with hydrocarbon gases, resulting in a significant deterioration of the quality and calorific value of the gas extracted from the storage facility. It is also necessary to work out the possibility of using part of hydrocarbon components for the own needs of the field: boiler houses at the facilities, heating of well products, operation of compressor drives, and many other things making up the energy complex and technological needs. In particular, when injecting CO₂,

one should take into account the corrosion hazard for equipment due to such factors as the high salinity of formation water and pressure.

Figure 4 presents a typical scheme for an underground gas storage facility in an aquifer combined with partial replacement of cushion gas. The effectiveness of this storage scheme is ensured by various factors. The aquifer should possess both the ability to accumulate gas, i.e. sufficient permeability and porosity, and a reliable fluid seal to prevent gas migration and crossflows. Other mechanisms that determine the efficiency of underground gas storage include rock heterogeneity, relative permeability hysteresis, gas dissolution, and gas miscibility (22).

Two main research methods were chosen for the study:

- Analysis of actual publications on the topic "Methods of associated gas utilization";
- hydrodynamic modeling of aquifer and temporary underground gas storage with partial replacement of cushion gas.

A detailed analysis of existing scientific research in the chosen direction was carried out using the allocation of key topics, taking into account the specific features of the found problem. To achieve the most accurate result, hydrodynamic modeling was carried out in a specialized program using modern computational algorithms that increase the accuracy and speed of calculations.

The results of the study were analyzed using the method of comparing gas and methane withdrawal-to-injection ratios.

2. 1. Associated Gas In hard-to-reach fields, there is no possibility to deliver gas from the outside; therefore, when choosing cushion gas, it is necessary to take into account the composition of associated gas produced at the field, since the source will be flue gases obtained by combustion. Table 1 presents associated gas composition used in this research.

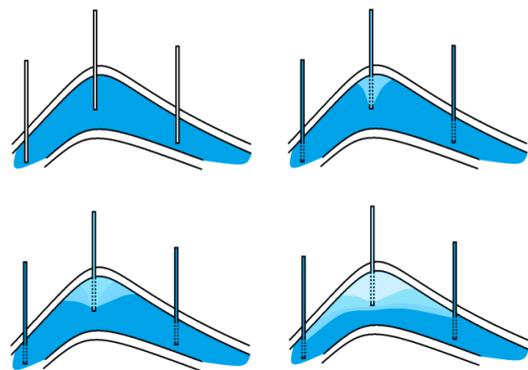


Figure 4. Schematic representation of underground gas storage with partial replacement of cushion volume

TABLE 1. Associated gas composition

N ₂	CO ₂	H ₂ S	CH ₄	C ₂ H ₆	C ₃ H ₈	iC ₄ H ₁₀	nC ₄ H ₁₀	C ₅₊
0,002	0,035	0,003	0,693	0,132	0,065	0,010	0,028	0,032
0,008	0,025	0,000	0,700	0,124	0,076	0,014	0,035	0,018
0,006	0,025	0,002	0,709	0,163	0,054	0,008	0,016	0,016

2. 2. Aquifer The thermobaric conditions of the selected geological object and the component composition determine such characteristics of the injected gas as the supercompressibility coefficient, density, and viscosity. They also exert a significant impact on the completeness

of displacement, the degree of gas mixing, and the volume of cushion gas (23). Calculation was performed for reservoir temperature 298,15 K.

For Eastern Siberia, and in particular, for the Yurubcheno-Tokhomszkaya oil and gas accumulation area, three water-bearing complexes are typical: suprasalt, saline, and subsalt. Since for gas storage, a good fluid seal is required, this paper will consider the Osinsky horizon, which is located at the junction of the saline and subsalt formations and overlain by a regional cover – the Usolye suite. Waters of this complex are of the calcium chloride type (24). Table 2 presents the component composition of waters (in mg-eq/dm³) with a mineralization of 299 g/dm³ (depths 2090–2100).

Temperatures in the Osinsky horizon vary within 20–30 °C; the pressure is within 21–23 MPa (25). When gas is injected into the horizon, a significant increase in pressure occurs, resulting in changes in the properties of both gas and formation water. To calculate the supercompressibility coefficient, density, and viscosity, it is necessary to assume that the formation temperature is constant.

2. 3. Cushion Gas Properties In this work, the pseudocritical properties of the gas mixture are calculated by the Sutton method for associated gas (26). To correct the dependencies on the content of carbon dioxide and hydrogen sulfide in associated gas, corrections for pseudocritical pressure and temperature obtained by Sutton (26) can be used. In the considered case, the presence of nitrogen should also be taken into account (27). As early as 1942, Standing and Katz (28) developed a diagram of the dependence of the z-factor on pseudo-reduced pressures and temperatures. After that, various researchers proposed mathematical correlation dependencies most closely reproducing the original

TABLE 2. Component composition of the Osinsk aquifer

Cl ⁻	Br ⁻	SO ₄ ²⁻	HCO ₃ ⁻	Ca ²⁺	Na ⁺	Mg ²⁺	Fe ²⁺
99,1	0,6	0,2	0,1	45,3	40,6	13,8	0,3

diagram (27). Standing and Katz (28) evaluated the convergence of most of the above-presented models. The best results were shown by the Hall and Yarborough model. The best results were demonstrated by Hemmati-Sarapardeh et al. (29). Taking this into consideration, the results of the calculation of the supercompressibility coefficient based on these dependencies for associated gas and temperature and pressure of the Osinsky horizon are presented in Figure 5.

Density and viscosity are found according to Lee’s dependencies. The dependency graphs are shown in Figures 6 and 7.

In aquifer conditions, CO₂ has the lowest supercompressibility coefficient, which implies that the volume it will occupy in the formation will be significantly less than that of natural gas, let alone nitrogen. Besides, since the density of CO₂ is significantly higher than that of hydrocarbon gas and nitrogen, carbon dioxide, when injected into the formation, under the influence of the gravity gradient will accumulate in the lower part of the formation. This will

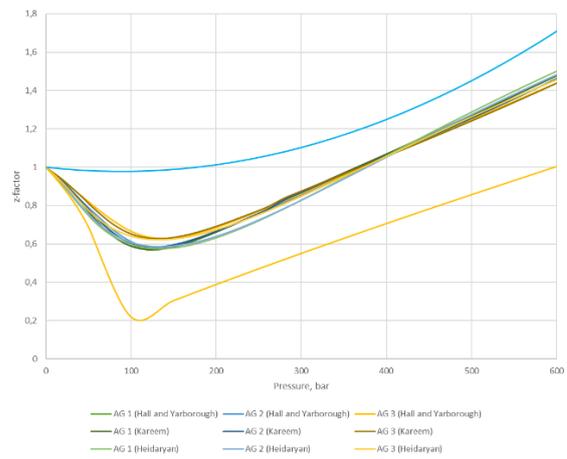


Figure 5. Z-factor dependence on pressure

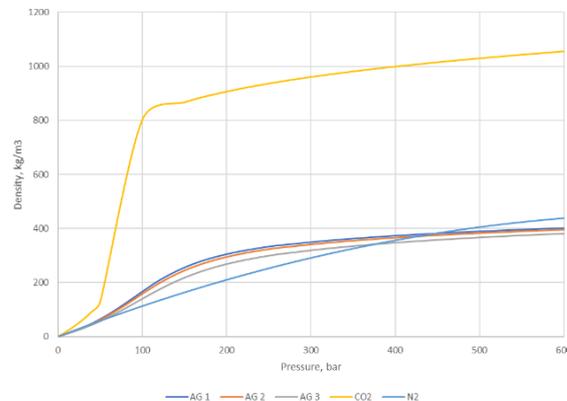


Figure 6. Dependence of density to pressure

make it possible to reduce the degree of mixing of hydrocarbon and non-hydrocarbon gases. When injecting CO₂, it is expedient to perforate the bottom of the production string closer to the formation bottom.

The higher the viscosity of the injected fluid, the better its displacement properties due to the higher ratio of gas viscosity to formation water viscosity.

One can conclude that carbon dioxide has the best displacement ability among gases proposed for use as cushion gases. Nitrogen, in turn, has the worst displacement ability, but due to its significant expansion in the formation, it can provide an effective cushion between CO₂ and associated gas.

Kareem et al. (30) considered the possibility of sequential injection of N₂ and CO₂. The results of the sequential injection experiment showed that this technology made it possible not only to increase the CO₂ storage capacity but also to improve CH₄ recovery. Nuhu et al. (31) studied the displacement of stored CH₄ using a CO₂-N₂ mixture. The results showed that this technology reduced pore swelling and provided high hydrocarbon gas recovery and CO₂ storage. Nitrogen acts as a kind of cushion preventing the mixing of carbon dioxide and methane. It can be concluded that the combined use of nitrogen and carbon dioxide provides better results in terms of both reducing miscibility and displacement of hydrocarbon gas. Zhang et al. (32) investigated adsorption at the gas-rock interface. The highest convergence was achieved using the Cuckoo algorithm.

2. 4. Case Study

The present work studies the possibility of minimizing cushion gas by partially replacing it with CO₂, N₂, or their mixture. Using the aquifer model (parameters are presented in Table 3) built in the RFD tNavigator software, an experimental study of 24 storage scenarios was conducted (Figure 8).

Although the presence of water and low temperature catalyze hydrate formation, they were not considered in this model (33, 34). Any deposits in both the borehole and bottomhole area of production and injection wells reduce the productivity and throughput of the well (35, 36). This aspect will also be considered in future studies.

TABLE 3. Aquifer model

Property	Value
Depth, m	2000
Temperature, K	298,15 308,15
Formation pressure, bar	220
Salinity, g/l	290
Porosity, %	20
Permeability, mD	300
Capacity, m	100

The base option was associated gas injection without partial replacement.

The calculation time was taken based on the assumption that in 10 years, the field will be connected to the main gas pipeline, and reverse gas production will start. The model took into account the solubility of injected gas components in the reservoir water.

The relative phase permeabilities in the gas-water system were adjusted based on laboratory studies by Zhu et al. (37). The results are shown in Figure 9.

In models 1–8, both production and injection wells are perforated in the lower part of the formation, and in models 9–16, the upper part of the production well is perforated. The aggregate state of carbon dioxide changed depending on the formation temperature. When the temperature reaches 31.2 °C, it passes to a supercritical state at formation pressure.

The model did not take into account the hysteresis of relative permeabilities. With successive changes in filtration directions, the relative permeability to water is reduced with each cycle (37). The cyclic process results not only in the expansion of the filtration area and enhanced rock hydrophilicity but also in enhanced gas saturation (38).

3. RESULT AND DISCUSSION

Figure 10 shows a dot plot of the dependence of the injected gas volume on the CH₄ extraction-to-injection

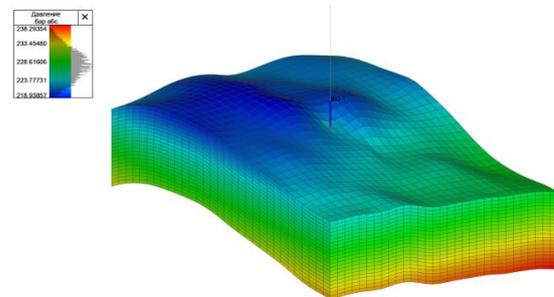


Figure 8. Aquifer model

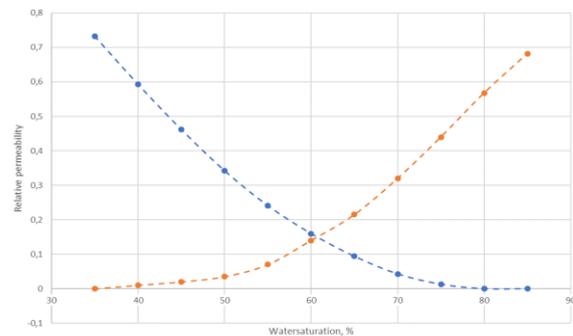


Figure 9. Gas-water relative phase permeabilities

ratio. It can be seen from the figure that the models located 1-6, included in Zone I include a production well and an injection well with perforations from 2050 to 2100 meters, and in Zone II (models 7 -16) a production well with perforations from 2000 to 2050 meters and an injection well with perforations from 2050 to 2100 meters. This implies that it is more efficient to inject cushion gas into the lower part of the aquifer and extract active gas from the upper part. Taking into account that the temporary underground gas storage facility should be able to contain as much associated gas as possible for utilization, the optimal solution is Model 12, where the cushion gas is carbon dioxide in a supercritical state.

Table 4 presents the calculation results for the simulation of partial cushion gas substitution scenarios. When evaluating the results obtained by the ratio of the injected hydrocarbon gas volume to the pumped-out volume, the best results were shown by scenario 16, in which carbon dioxide in a supercritical state is injected first, followed after some time by the injection of nitrogen and then storage gas.

The efficiency of gas extraction from the storage facility is achieved if the injection well is perforated at the bottom and the production well at the top. When calculating options with alternate injection of carbon dioxide and nitrogen, it was considered that the gases were injected for half a year before the injection of hydrocarbon gas started. To further investigate the technology of temporary underground gas storage with

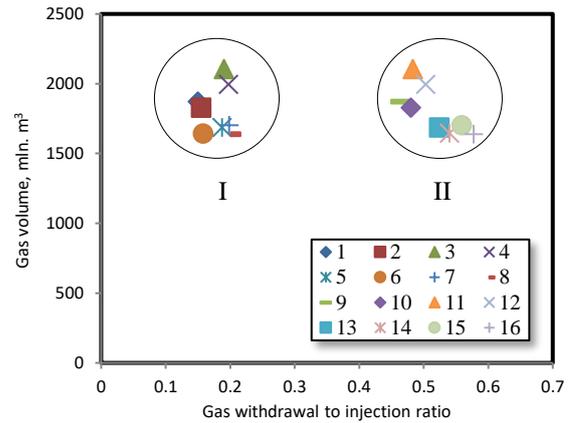


Figure 10. Graph of injected gas volume versus CH₄ withdrawal/injection ratio

partial replacement of cushion gas, it is recommended to study the ratio of carbon dioxide and nitrogen.

Modeling confirmed the effectiveness of injecting dioxide into the lower part of the reservoir: due to its higher density, it is located in the formation below nitrogen and hydrocarbon gases, which reduces its final content in the well products. Water produced together with gas from a production well can be reinjected into the lower part of the formation to increase pressure and increase gas displacement by water when pumping gas out from storage.

TABLE 4. Modeling results

Scenario	Temperature		Gas composition			Production		Gas withdrawal to injection ratio/CH ₄	
	25°C	35°C	CH ₄ , mln. m ³	CO ₂ , mln. m ³	N ₂ , mln. m ³	Gas, mln. m ³	Water, ths. m ³		
1	+	—	1872	—	—	281	2986	0,150	0,150
2	—	+	1828	—	—	283	3013	0,155	0,155
3	+	—	2105	307	—	401	2928	0,166	0,190
4	—	+	1996	322	—	393	2946	0,169	0,197
5	+	—	1687	—	150	337	3010	0,183	0,187
6	—	+	1644	—	144	288	3023	0,161	0,158
7	+	—	1702	133	216	337	2986	0,164	0,198
8	—	+	1638	136	202	330	3003	0,167	0,201
9	+	—	1872	—	—	870	1789	0,464	0,464
10	—	+	1828	—	—	877	1767	0,480	0,480
11	+	—	2105	307	—	1016	1717	0,421	0,483
12	—	+	1996	322	—	1005	1700	0,433	0,503
13	+	—	1687	—	150	885	1783	0,482	0,524
14	—	+	1645	—	144	890	1759	0,497	0,540
15	+	—	1702	134	216	953	1748	0,464	0,559
16	—	+	1638	136	202	945	1728	0,478	0,577

4. CONCLUSION

This study considered the issue of associated gas flaring at remote hard-to-reach oil and gas condensate fields, which are planned to be connected to the main gas pipelines only in the long run. For this type of facilities, most often the applied methods of associated gas utilization turn out to be unsuitable, which leads to the search for technologies of gas reserves conservation. Underground gas storage is an effective method of solving this problem, but it is also capital intensive. In this regard, it becomes urgent to search for and create a technology that allows to reduce the cost of creating a storage facility and increase its efficiency. This research consists in studying the possibility of partial substitution of hydrocarbon cushion gas by carbon dioxide and/or nitrogen. Using hydrodynamic modeling of the aquifer corresponding to the geological conditions of Eastern Siberia, sixteen scenarios of storage creation with different composition of cushion gas were investigated. As a result of analysis and comparison, the most effective gas mixture composition and injection and withdrawal method were determined.

As one of the most effective methods for solving this problem, the technology of temporary underground gas storage with partial replacement of cushion gas with carbon dioxide, nitrogen, or their combination was chosen. Based on the analysis of the dependencies of the supercompressibility coefficient, density, and viscosity on pressure at a constant temperature and modeling of the aquifer, the following conclusions can be drawn:

1. Due to the lowest z-factor coefficient, the volume occupied in the formation by carbon dioxide is much smaller than that of natural gas, let alone nitrogen.
2. Among the gases considered, carbon dioxide has the highest density in reservoir conditions ($\rho_{CO_2} = 906 \frac{kg}{m^3}$), which indicates that it will occupy lower formations in the reservoir in relation to the active gas.
3. The greatest change in the properties of the gas occurs in case of its transition to a supercritical state. The higher viscosity of the gas ensures more efficient water displacement.
4. The scenario of alternate injection of carbon dioxide and nitrogen is the most favorable in terms of the ratio of gas extraction to gas injection ($0,577 \frac{m^3}{m^3}$).
5. The scenario implying the injection of supercritical carbon dioxide as a cushion gas at the top perforation of the production well was selected as the most optimal one.
6. The group of scenarios with top perforation of the production well showed better results than those with bottom perforation. This is due to the low density of natural gas that tends to rise upwards.

7. The main limitation of the selected technology is the supercritical state of carbon dioxide, to which it passes when it reaches a temperature of 31.2 °C. It is desirable to select aquifers with temperatures above the critical temperature.

The academic contribution of this work is the creation of a new aquifer model that takes into account the solubility of the gases injected into it, as well as the allocation of the most effective gas injection technology by comparing and contrasting sixteen major existing scenarios.

Practical application of the research results is possible both in remote and hard-to-reach oil and gas condensate fields and in the creation of underground storage facilities in aquifers.

5. REFERENCES

1. Ilinova A, Kuznetsova E. CC (U) S initiatives: Prospects and economic efficiency in a circular economy. *Energy Reports*. 2022;8:1295-301. 10.1016/j.egy.2021.11.243
2. Ilinova AA, Romasheva NV, Stroykov GA. Prospects and social effects of carbon dioxide sequestration and utilization projects. *Записки Горного института*. 2020;244:493-502. 10.31897/pmi.2020.4.12
3. Roadmap NZ. A Global Pathway to keep the 1.5° C Goal in Reach. International Energy Agency. 2023. <https://www.iea.org/reports/net-zero-roadmap-a-global-pathway-to-keep-the-15-0c-goal-in-reach>
4. Zyryanova M, Snytnikov P, Amosov YI, Belyaev V, Kireenkov V, Kuzin N, et al. Upgrading of associated petroleum gas into methane-rich gas for power plant feeding applications. Technological and economic benefits. *Fuel*. 2013;108:282-91. 10.1016/j.fuel.2013.02.047
5. Amuda YJ. Impact of COVID-19 on Oil and Gas Sector in Nigeria: A Condition for Diversification of Economic Resources. *Emerging Science Journal*. 2023;7:264-80. 10.28991/ESJ-2023-SPER-019
6. Petrakov D, Loseva A, Alikhanov N, Jafarpour H. Standards for selection of surfactant compositions used in completion and stimulation fluids. *International Journal of Engineering, Transactions C: Aspects*. 2023;36(9):1605-10. 10.5829/IJE.2023.36.09C.03
7. Kotek P, Selei A, Tóth BT, Felsmann B. What can the EU do to address the high natural gas prices? *Energy Policy*. 2023;173:113312. 10.1016/j.enpol.2022.113312
8. A.V. I, A.V. S, I.K. S. Ecological and economic justification of the utilization of associated petroleum gas at oil fields of Russian Federation. *Geologiya i Geofizika Yuga Rossii*. 2020;10(1):114-26. 10.23671/VNC.2020.1.59069
9. Sun S, Ma L, Li Z. Methane emission and influencing factors of China's oil and natural gas sector in 2020–2060: A source level analysis. *Science of The Total Environment*. 2023;905:167116. 10.1016/j.scitotenv.2023.167116
10. Soltanieh M, Zohrabian A, Gholipour MJ, Kalnay E. A review of global gas flaring and venting and impact on the environment: Case study of Iran. *International Journal of Greenhouse Gas Control*. 2016;49:488-509. 10.1016/j.scitotenv.2023.167116
11. Shonin OB, Salov RA. Improvement of energy efficiency, reliability and environmental safety of power plants based on

- associated petroleum gas. *Journal of Ecological Engineering*. 2017;18(3):91-6. 10.12911/22998993/69357
12. Morenov V, Leusheva E, Buslaev G, Gudmestad OT. System of comprehensive energy-efficient utilization of associated petroleum gas with reduced carbon footprint in the field conditions. *Energies*. 2020;13(18):4921. 10.3390/en13184921
 13. Drozdov AN, Gorelkina EI. Development of a pump-ejector system for SWAG injection into reservoir using associated petroleum gas from the annulus space of production wells. *Записки Горного института*. 2022;254:191-201. 10.31897/PMI.2022.34
 14. Hill B, Hovorka S, Melzer S. Geologic carbon storage through enhanced oil recovery. *Energy Procedia*. 2013;37:6808-30. 10.1016/j.egypro.2013.06.614
 15. Khan C, Amin R, Madden G. Effects of CO₂ and acid gas injection on enhanced gas recovery and storage. *Journal of Petroleum Exploration and Production Technology*. 2013;3:55-60. 10.1007/s13202-012-0044-8
 16. Wood DA, Nwaoha C, Towler BF. Gas-to-liquids (GTL): A review of an industry offering several routes for monetizing natural gas. *Journal of Natural Gas Science and Engineering*. 2012;9:196-208. 10.1016/j.jngse.2012.07.001
 17. Barrufet MA, Bacquet A, Falcone G. Analysis of the storage capacity for CO₂ sequestration of a depleted gas condensate reservoir and a saline aquifer. *Journal of Canadian Petroleum Technology*. 2010;49(08):23-31. 10.2118/2009-197
 18. Mikhailovsky A, Kornev G, Isaeva N. Associated gas rational utilization: temporary underground gas storage engineering in gas-condensate field. *GEORESUSY*. 2010;36(4):47-51. <https://geors.ru/archive/article/229/?ysclid=lq2npk2lxe17464615>
 19. Mukhametzyanov I, Vershinina M, Glavnov N, Gareev R, Popov A, Penigin A, et al., editors. Miscible displacement as an approach to increase the development efficiency of the remote oil fields. *SPE Russian Petroleum Technology Conference?*; 2018: SPE. 10.2118/191576-18rptc-ms
 20. Ignatyev N, Shvets V, Levanov A, Zakharova E, Yaschenko S, Zyryanov K, et al. Organizing, monitoring, and operating a temporary underground gas storage at Verkhnechonskoye field in Eastern Siberia (Russian). *Oil Industry Journal*. 2021;2021(08):84-8. 10.24887/0028-2448-2021-8-84-88
 21. Jia S, Xu M, Wen C, Li B, Liu T, Xi Z, et al. A quantitative approach for sealing capacity evaluation of caprock in candidate of aquifer gas storage. *Environmental Science and Pollution Research*. 2023;30(23):63678-90. 10.1007/s11356-023-26873-x
 22. Kantyukov RR, Zapevalov DN, Vagapov RK. Analysis of the application and impact of carbon dioxide media on the corrosion state of oil and gas facilities. *Записки Горного института*. 2021;250:578-86. 10.31897/PMI.2021.4.11
 23. Gasumov RA, Gasumov ER, Minchenko JS. Features of the underground storages construction in depleted oil and gas condensate fields. *Записки Горного института*. 2020;244:418-27. 10.31897/PMI.2020.4.4
 24. Al-Shafi M, Massarweh O, Abushaikha AS, Bicer Y. A review on underground gas storage systems: Natural gas, hydrogen and carbon sequestration. *Energy Reports*. 2023;9:6251-66. 10.1016/j.egypr.2023.05.236
 25. N.S. T. Hydrogeological features of the Yurubcheno-Tokhomskoye oil and gas condensate field. *Oil Industry Journal*. 2021; 2:50-5. 10.24887/0028-2448-2021-2-50-55
 26. Sutton RP, editor *Fundamental pVT calculations for associated and gas/condensate natural gas systems*. SPE Annual Technical Conference and Exhibition?; 2005: SPE. 10.2118/97099-PA
 27. Lee WJ, Wattenbarger RA. *Gas reservoir engineering*: Society of Petroleum Engineers; 1996.
 28. Standing MB, Katz DL. Density of natural gases. *Transactions of the AIME*. 1942;146(01):140-9. 10.2118/942140-G
 29. Hemmati-Sarapardeh A, Hajirezaie S, Soltanian MR, Mosavi A, Nabipour N, Shamshirband S, et al. Modeling natural gas compressibility factor using a hybrid group method of data handling. *Engineering Applications of Computational Fluid Mechanics*. 2020;14(1):27-37. 10.1080/19942060.2019.1679668
 30. Kareem LA, Iwalewa TM, Al-Marhoun M. New explicit correlation for the compressibility factor of natural gas: linearized z-factor isotherms. *Journal of Petroleum Exploration and Production Technology*. 2016;6:481-92. 10.1007/s13202-015-0209-3
 31. Mohammed N, Abbas AJ, Enyi GC, Suleiman SM, Edem DE, Abba MK. Alternating N₂ gas injection as a potential technique for enhanced gas recovery and CO₂ storage in consolidated rocks: an experimental study. *Journal of Petroleum Exploration and Production Technology*. 2020;10:3883-903. 10.1007/s13202-020-00935-z
 32. Zhang H, Diao R, Mostofi M, Evans B. Molecular simulation of enhanced CH₄ recovery and CO₂ storage by CO₂-N₂ mixture injection in deformable organic micropores. *Journal of Natural Gas Science and Engineering*. 2020;84:103658.
 33. Esfandian H, Esfandian N, Rezazadeh M. Modeling and comparison of optimized isotherm models for H₂, N₂, CO, CH₄ and CO₂ adsorption using cuckoo search optimization algorithm. *International Journal of Engineering, Transactions B: Applications*. 2020;33(5):712-9. 10.5829/IJE.2020.33.05B.01
 34. G.Y. K, A.A. V. Study of Conditions for Gas Hydrate and Asphaltene-Resin-Paraffin Deposits Formation in Mechanized Oil Production. *Bulletin of the Tomsk Polytechnic University Geo Assets Engineering*. 2023;334(10):61-75. 10.18799/24131830/2023/10/4181
 35. Azizi J, Azizi J, Davarnejad R. Prediction of gas hydrate formation using HYSYS software. *International Journal of Engineering*. 2014;27(9):1325-30. https://www.ije.ir/article_72369.html
 36. Korobov G.Y, PDV, Nguyen V.T. Mechanisms of formation of asphalt-resin and paraffin deposits and factors influencing their intensity. *Bulletin of the Tomsk Polytechnic University Geo Assets Engineering*. 2023;334(4):103-16. 10.18799/24131830/2023/4/3940
 37. Sinan Z, Junchang S, Guoqi W, Zheng D, Jieming W, Lei S, et al. Numerical simulation-based correction of relative permeability hysteresis in water-invaded underground gas storage during multi-cycle injection and production. *Petroleum Exploration and Development*. 2021;48(1):190-200. 10.1016/s1876-3804(21)60015-0
 38. Fatemi SM, Sohrabi M. Relative permeabilities hysteresis for oil/water, gas/water and gas/oil systems in mixed-wet rocks. *Journal of Petroleum Science and Engineering*. 2018;161:559-81. 10.1016/j.petrol.2017.11.014

COPYRIGHTS

©2024 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**

شعله ور شدن گاز همراه همچنان یک مشکل برای میداین نفت و گازی است که دسترسی به آنها دشوار است و از زیرساخت ها دور هستند. توسعه فعال تولید نفت و گاز در سیبری شرقی به این واقعیت منجر شده است که ظرفیت های حمل و نقل نمی تواند با توسعه میدان همگام شود. افزایش شعله ور شدن گازهای همراه منجر به افزایش قابل توجه گازهای گلخانه ای مانند دی اکسید کربن و متان می شود. راه حل احتمالی این مشکل، ذخیره گاز در سفره آب این میدان برای فروش آتی و کسب درآمد از طریق خط لوله اصلی گاز است. این مقاله فن آوری های اصلی استفاده از گاز مرتبط را تحلیل می کند و مشکل دور بودن از زیرساخت های حمل و نقل گاز میداین صعب العبور را نشان می دهد. یک فناوری موثر برای حل این مشکل، ایجاد ذخیره سازی موقت زیرزمینی گاز همراه در آبخوان میدان است. نتایج مدلسازی هیدرودینامیکی تحقق این فناوری با جایگزینی جزئی گاز بالشتک نشان داد که تزریق مشترک دی اکسید کربن و نیتروژن قبل از گاز هیدروکربن امکان افزایش نسبت بین گاز تولیدی و تزریقی را فراهم می کند که نشان دهنده کارایی بیشتری آن است. توصیه می شود برای پیاده سازی این فناوری، هنگام انتخاب محل تزریق زمین شناسی، روی سفره های زیرزمینی با دمای بالاتر از ۳۱.۲ درجه سانتی گراد تمرکز شود که باعث می شود دی اکسید کربن در شرایط مخزن در حالت فوق بحرانی باقی بماند.



Comprehensive Assessment Production Efficiency of Electric Rope Shovel through Operator Qualification Criteria

A. V. Mikhailov, D. A. Shibanov, A. E. Bessonov*, C. Bouguebrine

Department of Mechanical Engineering, Saint Petersburg Mining University, Russia

PAPER INFO

Paper history:

Received 10 December 2023

Received in revised form 10 February 2024

Accepted 02 March 2024

Keywords:

Ergatic System

Working Cycle Time

Electric Rope Shovel

Operator Qualification

Bucket Filling Ratio

ABSTRACT

The article examines methods for assessing the efficiency of electric mining excavators, emphasizing the inseparability of operational efficiency from the operator-machine ergatic system. It reviews methods for evaluating operator skills via experimental data and proposes a comprehensive approach to assess the excavator's operational efficiency and the operator's skill level. This method includes analyzing the machine's operating time and energy efficiency using a simulator, thereby offering a novel perspective on the dynamic interaction between human operators and automated systems. With the working cycle's duration measured by the ratio of average to nominal cycle times, and energy efficiency assessed through the comparison of specific energy consumption to theoretical values. The findings suggest prioritizing reductions in operating cycle time for suboptimal machine control and focusing on improving bucket fill rates to enhance energy efficiency. Moreover, the study underscores the potential for utilizing these methodologies in real-world applications, aiming to optimize the utilization of mining equipment and thereby significantly contribute to the advancement of operational methodologies in the mining sector.

doi: 10.5829/ije.2024.37.07a.03

Graphical Abstract

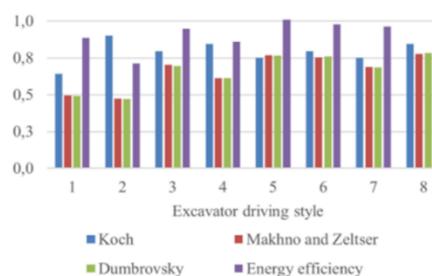
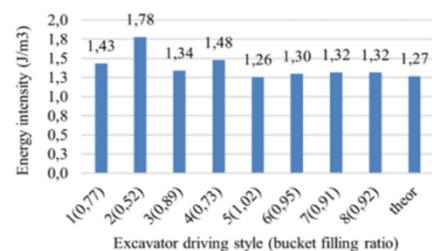
experiment at the EKG-18R training complex

analyse the experimental results

offering a new control coefficient for a quarry electric excavator

comparison of existing methods for evaluating operator qualification

offering a comprehensive approach to evaluating the production efficiency of equipment



*Corresponding Author Institutional Email: s215049@stud.spmi.ru (A. E. Bessonov)

Please cite this article as: Mikhailov AV, Shibanov DA, Bessonov AE, Bouguebrine C. Comprehensive Assessment Production Efficiency of Electric Rope Shovel through Operator Qualification Criteria. International Journal of Engineering, Transactions A: Basics. 2024;37(07):1231-8.

NOMENCLATURE

W_A	specific energy intensity for one cycle process (J/m ³)	m_{tt}	excavator turntable weight (kg)
N_Y	the total power of the excavator electric motors, (W)	m_s	soil weigh (kg)t
q_b	bucket capacity (m ³)	k_a	experimental positioning factor
$t_{a.c}$	average cycle time (s)	r	radius of the center of mass of rotation (m)
$k_{a,fill}$	average fill factor	ε	platform angular acceleration (1/s ²)
A_c	cycle work (J)	φ	rotation angle (rad)
A_d	work of digging resistance forces (J)	k_1	loosening factor
A_{lt}	lifting work (J)	K_k	Koch control factor
A_t	work on turning an excavator with soil (J)	t_{theor}	theoretic cycle time (s)
A_u	soil unloading work (J)	K_d	Dubrovsky control factor
$A_{c,t}$	empty bucket turning operation (J)	Q	excavator capacity (m ³ /h)
A_{lw}	work to lower the bucket to the beginning of the face (J)	Q_{theor}	theoretical excavator capacity (m ³ /h)
k_{sp}	specific resistance of soil to digging (N/m ²)	K_c	Makhno control factor
Y_s	soil volume weight (N/m ³)	$t_{sp,theor}$	specific theoretical cycle time (s/m ³)
g	gravitational acceleration (m/s ²)	$t_{sp,actual}$	specific actual cycle time (s/m ³)
H_{lift}	lifting height (m)	K_e	energy efficiency control factor
m_b	bucket weight (kg)	W_{theor}	theoretical specific energy intensity (J/m ³)

1. INTRODUCTION

In the current trend of mining industry development, there is a constant desire to improve the efficiency and sustainability of production processes. Electric rope shovels, specifically the EKG-18R model, with their significant power capacity, aging machinery, and the varying skill levels of operators, stand at the forefront of this developmental push. These shovels are no exception. This equipment is being developed by an intensive method, on the basis of which new technological solutions do not appear on it. The increasing volume of excavated mass in one cycle of work is due to an increasing dimensions of the machine and the energy intensity of the excavation process. In turn, various operator assistance systems are added: cameras, devices that signal a possible collision of the bucket with the caterpillar, etc. Which leads to the conclusion that the main focus for equipment modernization is to help the operator and level out the lack of experience and skills (1, 2).

There are several parameters that characterize the production efficiency of the equipment. These include: productivity, specific energy consumption, specific material consumption, unit cost, etc. Production efficiency is primarily affected by the operating conditions of the equipment (3). Having divided them into the main categories, it can be represented as: mining-geological and mining engineering, the quality of preparation of the face and rock mass, the technical condition of the machine, the organization of mining operations, climatic conditions, ergatic system (4-6). The

operating conditions of the equipment also affect the basic property of reliability, namely durability (7-9). As a consequence, it is necessary to select equipment and develop a technological scheme of development for its effective utilization under given operating conditions (10, 11). Additionally, advanced technologies play a crucial role in enhancing the production efficiency of equipment. Notably, intelligent control systems and methods based on artificial intelligence significantly impact not only the improvement of efficiency but also the enhancement of equipment reliability across various operational conditions (12, 13).

The ergatic system, or man-machine system, is the interaction of a person with equipment. The main criterion of this system is the qualification of the operator, which should reflect the experience and skills of managing this machine (14, 15). There are several methods for assessing the operator qualifications of scientists. These methods will be shown in the methodology section. However, these methods rely only on the average cycle time of the excavator or on its performance, which does not allow to fully evaluate the ergatic system (16, 17). There are also approaches to analyze the factors affecting the operator in an ergatic system, such as: temperature of the environment, fatigue, and well-being. This study deals directly with the evaluation of operator skill or machine control criteria (18-20).

Nowadays, in real production, the efficiency of equipment utilization is evaluated by the productivity of the machine. Such as the excavator's positioning, which directly impacts the machine's productivity. The

qualification of the operator is evaluated in the same way. However, this assessment is not objective, because it is considered together with the work of dump trucks. I.e. it is impossible to assess the production efficiency of only a dredging machine when at least one more object participates in the production cycle, as a rule, there are more dump trucks working on one excavator. Moreover, if take into account the performance of the machine itself, without the participation of third-party equipment, the ergatic system remains. Other criteria affecting the production efficiency of the equipment, except for the ergatic system, can be taken into account in calculations or by having reference values of operating time under given operating conditions.

Furthermore, the mining methods employed, such as the single backup or double backup methods, along with the rock quality and conditions of blasting operations, significantly influence the digging time (power) and other components of the excavator's working cycle duration. Consequently, conclude that the production efficiency of the equipment is inextricably linked to the ergatic system. As a result, by evaluating the production efficiency of the equipment it is possible to evaluate the level of operator's qualification. In this paper we proposed a comprehensive approach to assess the skill level of the operator, and, the production efficiency of the equipment, based on the evaluation of performance and energy efficiency of the equipment during operation. While extensive research exists in the field, practical application at the production level remains limited or non-indicative of actual efficiency. This gap underlines the necessity for innovative methodologies that accurately assess and enhance the real-world performance of mining equipment. The authors contribute by addressing the core challenges and obstacles in measuring production efficiency and operator proficiency, presenting initial achievements in overcoming these difficulties. Through a novel lens, this study introduces a methodological framework that not only evaluates the equipment's operational efficiency but also lays the groundwork for future research aimed at predicting the operational lifespan of electric rope shovels, thereby marking a significant advancement in the domain of mining machinery management (21-23).

2. METHODOLOGY

This study considers an integrated approach to the assessment of production efficiency taking into account the level of operator's skills directly at the site of extraction. The main criterion for evaluation is proposed to use the specific energy intensity of the excavation process for one cycle of work. This criterion combines the parameters directly related to the operator's skill level: working cycle time and bucket filling factor.

Further in the formulas will be used the average cycle time of the excavator and the average bucket filling factor, because in the calculations will be used directly these values based on the results of the experiment. Specific energy intensity of the excavation process for one working cycle as a ratio of the work performed to the bucket capacity given by Equation 1:

$$W_A = \frac{N_{y t_{a.c}}}{q_b} \quad (1)$$

To calculate the specific energy intensity of the process, will be used the average cycle time of the excavator to fill the dump truck. The data obtained for the calculations are presented in section 3 experiment. Also the bucket capacity is considered taking into account the average filling factor. As a consequence, the formula of specific energy intensity of the excavation process will take by Equation 2:

$$W_A = \frac{N_{y t_{a.c}}}{q_b k_{a.fill}} \quad (2)$$

The total power of the excavator drives multiplied by the average cycle time represents the total work done during the excavation cycle. During the excavation process, various actions are performed, which can be represented as work done by: digging (4), lifting (5), turning (6), turning with empty bucket (7), lowering (8). Also, in this cycle there is unloading work, but due to the specifics of the equipment it was taken as 0. Nevertheless, in Equation 3 given for the total work it is taken into account.

$$A_c = A_d + A_{it} + A_t + A_u + A_{e.t} + A_{tw} \quad (3)$$

$$A_d = k_{sp} \cdot q_b \cdot k_{a.fill} \cdot k_1 \quad (4)$$

$$A_{it} = (q_b \cdot Y_s + m_b \cdot g) \cdot H_{lift} \quad (5)$$

$$A_t = k_a (m_{it} + m_s + m_b) \cdot r^2 \cdot \varepsilon \cdot \varphi \quad (6)$$

$$A_{e.t} = k_a (m_{it} + m_b) \cdot r^2 \cdot \varepsilon \cdot \varphi \quad (7)$$

$$A_{tw} = m_b \cdot g \cdot H_{lift} \quad (8)$$

In the ground lifting formula Equation 5 the value of volumetric weight $Y_s = 2000 \text{ N/m}^3$ is assumed for the third category of soils, $g = 9,81 \text{ m/s}^2$. In the formula of rotation 6 angular acceleration of the rotary platform $\varepsilon = 1 \text{ 1/c}^2$, experimental positioning coefficient $k_a = 1.1$, rotation angle $\varphi = 90^\circ$. The bucket mass in all calculations is taken $m_b = 31000 \text{ kg}$, equal to the bucket mass of 18 m^3 for excavator EKG-18R. In the digging operation Equation 4, the volume of excavated rock mass is expressed through bucket capacity, average bucket fill

factor and ground loosening factor. The factors of soil resistance to digging k_{sp} and loosening k_l were assumed to be 0.25 and 1.25 for ground category III, respectively.

A mathematical model of the fourth coordinate, or working cycle time, for the selection of the necessary equipment was suggested by Balovnev (24, 25). Considering the parameters of production efficiency and operational conditions, the determination of the working cycle time emerges as a pivotal element in the selection of equipment. This approach facilitates the computation of production efficiency by integrating all referenced operational factors. Additionally, it was introduced a methodology for evaluating an operator's proficiency, which hinges on the proportion of theoretical cycle time to actual cycle time, with the theoretical aspect either deduced or furnished by the equipment manufacturer (9).

However, the method of calculation based on the operating cycle time in isolation from the bucket filling ratio is not representative over the entire range of variations of operating parameters. Excluding this problem, another method presented by Dombrovsky (26) suggests using the ratio of actual productivity to theoretical productivity. There is also another method described by Makhno (27). This methods are similar to Dombrovsky method, but it uses the ratio of specific theoretical cycle time (12) to specific actual cycle time (13). These methods also use the ladle filling ratio to estimate, in addition to the working cycle time. This allows the production efficiency to be evaluated over a larger range of operating hours .

$$K_c = \frac{T_{\text{theor}}}{T_{\text{actual}}} \quad (9)$$

$$K_d = \frac{Q_{\text{actual}}}{Q_{\text{theor}}} \quad (10)$$

$$K_c = \frac{T_{\text{sp.theor}}}{T_{\text{sp.actual}}} \quad (11)$$

$$T_{\text{sp.theor}} = \frac{t_{c.theor}}{q_b} \quad (12)$$

$$T_{\text{sp.actual}} = \frac{t_{a.c}}{q_b \cdot k_{a.fill}} \quad (13)$$

In this study it is proposed to use the ratio of the energy intensity of the excavation process to the energy intensity of the excavation process nominal (14). That is, the energy intensity with which the operator works on the real machine to the energy intensity calculated on the basis of nominal values. This indicator will make it possible to evaluate how energy efficient the equipment

is used at the production site. As a consequence, K_e will become one of the parts of a comprehensive approach to assessing the production efficiency of equipment and evaluating the skill level of the operator.

$$K_e = \frac{W_A}{W_{\text{theor}}} \quad (14)$$

By comparing and analyzing existing methods of operator qualification assessment, it is possible to understand their effectiveness and applicability in real production. It will also allow to define more precisely the vector of research, focusing on the problems not previously described.

3. EXPERIMENTAL

In the course of the study, an experiment was conducted to retrain the operator on a different type-size range of excavators. The experiment was conducted on the simulator complex designed to train operators working on the excavator EKG-18R. The range of simulator settings allows modeling various production conditions (type of minerals extracted, weighted average piece size, the share of oversize output and etc.). This complex fully simulates the work of the excavator, its main feature is the IDS system. This system displays in real time the load on the drives and displays instantaneous values of current and voltage of electric motors in time. IDS system allows to monitor the load on the lift, head, slewing and travel drives. In addition, it is possible to obtain data on the time of each equipment cycle, the number of cycles and the total volume of excavated material (28).

During the retraining process, the operator underwent a control exercise. Which consisted of filling a BELAZ-75304 with coal under normal operating conditions. Table 1 shows the results of these exercises. Since in this study was considered directly the evaluation of the operator's qualification and as a consequence the production efficiency of the equipment, the obtained data can not be considered in relation to the training time. Therefore, only an experimental database with different excavation rates under the same operating conditions is needed for further analysis (29). Thus, the results of the experiment are considered in relation to different operating styles of the equipment. The table also shows theoretical operating time parameters, i.e. calculated with respect to the data provided by the manufacturer. As the results of the experiment the main operating parameters necessary for further calculations are presented, they include productivity (Q), average bucket filling ratio ($k_{a.fill}$), number of cycles required to fill the BELAZ-75304, average cycle time ($t_{a.c}$) and radius of the center of mass of the rotating parts of the excavator (r).

TABLE 1. Experimentel results of equipment runtime

№	Q (J/m ³)	$k_{a,fill}$	Cycle number (e.g)	$t_{a,c}$ (s)	r (m)
1	1176	0,77	13	42	1,83
2	1131	0,52	18	30	1,58
3	1665	0,89	12	34	1,94
4	1469	0,73	14	32	1,80
5	1834	1,02	10	36	2,07
6	1825	0,95	11	34	2,00
7	1642	0,91	11	36	1,97
8	1879	0,92	11	32	1,97
theor	2400	1,00	10	27	2,05

4. RESULTS AND DISCUSSION

In the course of study, the work done by the excavator in one cycle was calculated. The results of calculations of digging work (9), lifting (5), turning (6), total work(3) and energy intensity of the excavation process for an average excavator cycle are shown in Table 2 with respect to different styles of equipment operation. The work on the return of the empty bucket (7) and its lowering (8) are equal for all styles of equipment operation. Accordingly, $A_{e,t} = 0,71$ MJ and $A_{lv} = 4,5$ MJ. Figure 1 shows the energy intensity of the excavation process in relation to different styles of excavator control and bucket filling ratio. Based on the graph we can conclude that the higher the bucket filling ratio, the less energy is required to excavate one cubic meter of rock.

Table 3 shows the results of equipment control coefficients calculated by the methods of Koch, Dumbrovsky, Makhno , and enery efficiency (14). Based on this data, a graph was created to compare these techniques as shown in Figure 2.

Having considered and compared the different criteria for evaluating operator qualification and as a consequence the production efficiency of the equipment.

TABLE 2. Calculated work and energy intensity of equipment

№	A_d (MJ)	A_t (MJ)	A_r (MJ)	A_c (MJ)	W_A (MJ/m ³)
1	3,47	8,72	2,43	19,90	1,43
2	2,36	7,40	1,79	16,83	1,78
3	3,99	9,35	2,76	21,38	1,34
4	3,27	8,49	2,31	19,35	1,48
5	4,61	10,10	3,19	23,17	1,26
6	4,28	9,69	2,96	22,20	1,30
7	4,13	9,51	2,86	21,77	1,32
8	4,15	9,54	2,87	21,84	1,32
theor	4,50	9,96	3,11	22,85	1,27

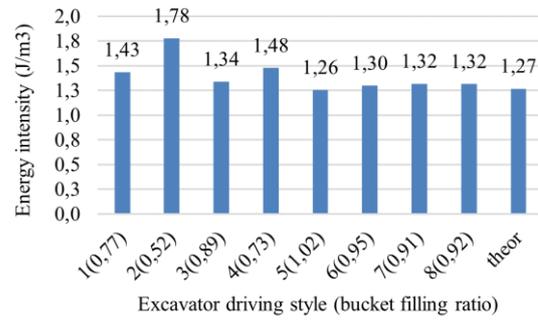


Figure 1. Dependence of energy intensity of excavation process on control style

TABLE 3. Estimated indicators of production efficiency of equipment

№	K_t	K_d	K_c	K_e
1	0,64	0,49	2,01	0,88
2	0,90	0,47	2,11	0,71
3	0,79	0,69	1,41	0,95
4	0,84	0,61	1,63	0,86
5	0,75	0,76	1,30	1,01
6	0,79	0,76	1,32	0,98
7	0,75	0,68	1,45	0,96
8	0,84	0,78	1,28	0,96

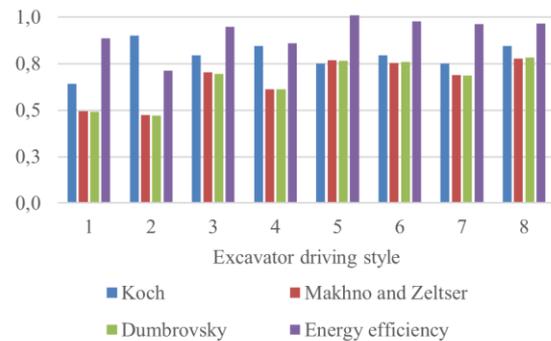


Figure 2. Dependence of production efficiency evaluation criteria on different excavator control styles

It can be seen that the approaches of Dumbrovsky, Makhno are similar in their results for different styles of equipment operation and differ slightly. Therefore, it can be concluded that these methods both are suitable for assessing operator qualification. However, in these approaches there is a problem of estimation of separate indicators of equipment operating time. It is fundamentally unclear which of the operating time indicators should be paid attention to in order to improve the production efficiency of the equipment.

Considering Koch's method, we can conclude that, as it was said earlier, the evaluation of cycle time does not cover all the parameters of equipment runtime. So in Figure 2, if we pay attention to the 2 styles of equipment operation, we can see that the performance and energy efficiency indicators are quite low relative to the others. However, the cycle time score is higher than the others. This is due to the fact that the bucket fill factor in this management style is low. Considering the other styles of equipment operation, we can say that this indicator correlates with Dumbrowski, Makhno indicators.

The energy efficiency of operation does not correlate with these parameters, as it is based mainly only on the bucket fill factor. At the same time, this parameter shows how efficiently the equipment is used in terms of energy consumed per cubic meter of production. Combining this approach together with Koch's method, it is possible to obtain a more indicative approach to assessing production efficiency.

Looking at Figure 2 in general, conclude that the most efficient control styles of the electric mining excavator are styles 3 and 5-8. The other control styles, in turn, have poor performance in terms of both productivity and energy efficiency.

Building on the foregoing discussion, particular emphasis should be placed on the energy efficiency criterion presented in this study. This criterion not only assesses the amount of energy expended per unit volume of extracted product but also, by integrating it with Koch's method, characterized by cycle time, will further allow for predicting the residual resource of the equipment. However, considering this research, combining these methods and delivering the results of these criteria directly during equipment operation, integrating them into the control system, this approach will enhance productivity by providing recommendations for equipment management. This includes advising on the necessity to reduce cycle times or increase the bucket fill factor, which will also enable the assessment of losses incurred by mining companies due to the use of unskilled labor.

5. CONCLUSION

This paper examined various techniques for assessing production efficiency and evaluating operator skill level based on different equipment control styles. The result of the study was comprehensive approach that includes the estimation of excavator cycle time based on Koch's method, as well as energy efficiency, considered as the ratio of production energy efficiency to theoretical evaluation of energy efficiency. Thanks to this approach, it is possible to assess the production efficiency of the equipment in operation from different aspects, and most importantly, to give recommendations for its

improvement. Thus, if the Koch control criterion is low, it is possible to conclude that it is necessary to reduce the operating cycle time. If the energy efficiency is low, the focus should be on increasing the bucket filling ratio. The integration of energy efficiency evaluation with Koch's cycle time analysis highlights the importance of assessing equipment performance from multiple dimensions. This approach facilitates a targeted improvement strategy by identifying specific areas of inefficiency. Moreover, it allows for a deeper understanding of how operator skill influences machine productivity and energy consumption, enabling more precise adjustments in training or operational procedures. The findings contribute to a nuanced perspective on optimizing electric rope shovel operations, suggesting a pathway for more efficient and effective use of resources.

6. REFERENCES

- Ivanov SL, Ivanova PV, Kuvshinkin SY. Promising model range career excavators operating time assessment in real operating conditions. *Journal of Mining Institute*. 2020;242(2):228-33. <https://doi.org/10.31897/PMI.2020.2.228>
- Ligotsky D. A review of mining and loading equipment currently used for open pit mining. *ARNP Journal of Engineering and Applied Sciences*. 2019;14(19):7154-8. <https://doi.org/10.36478/JEASCI.2019.7154.7158>
- Makarov VN, Anistratov KYu. Achieving the highest record monthly productivity of EKG-18 excavators at the open-pit mines of JSC Stroyservis. *Coal*, 2019;(1):20-6. <http://dx.doi.org/10.18796/0041-5790-2019-1-20-26>
- Lavrik A, Buslaev G, Dvoynikov M. Thermal Stabilization of Permafrost Using Thermal Coils Inside Foundation Piles. *Civil Engineering Journal*. 2023;9(4):927-38. <https://doi.org/10.28991/CEJ-2023-09-04-013>
- Basiri MH, Sharifi MR, Ostadi B. Reliability and risk assessment of electric cable shovel at Chadormalu iron ore mine in Iran. *International Journal of Engineering, Transactions A: Basics*. 2020;33(1):170-7. <https://doi.org/10.5829/IJE.2020.33.01A.20>
- Bolobov V, Akhmerov E, Rakiitin I. Influence of rock type on regularities of excavator bucket tooth crown wear. *MIAB Mining Inf Anal Bull*. 2022(6-2):189-204. https://doi.org/10.25018/0236_1493_2022_62_0_189
- Nasonov MY, Lykov YV, Trong, DD. The study of the resource and durability of metal structures of excavators after the expiration of the service life. *Coal*, 2020;(2 (1127)):13-7. <https://doi.org/10.18796/0041-5790-2020-2-13-17>
- El-Ghamry E, Abbas F, El-damcese M. Reliability analysis of three elements series and parallel systems under time-varying fuzzy failure rate. *International Journal of Engineering*. 2014;27(4):553-60. <https://doi.org/10.5829/idosi.ije.2014.27.04a.06>
- Gogolynskiy K, Gromyka D, Kremcheev E. A modelling of cyclic thermal and impact loads on excavator bucket. *International Review of Mechanical Engineering*. 2021;15(4):189-96. <https://doi.org/10.15866/ireme.v15i4.20699>
- Ivanov VV, Dzyurich DO. Justification of the technological scheme parameters for the development of flooded deposits of construction sand. *Записки Горного института*. 2022;253:33-40. <https://doi.org/10.31897/PMI.2022.3>

11. Kurganov VM, Gryaznov MV, Kolobanov SV. Assessment of operational reliability of quarry excavator-dump truck complexes. *Journal of Mining Institute*. 2020;241:10-21. <https://doi.org/10.31897/pmi.2020.1.10>
12. Morales LA, Fabara P, Pozo DF. An intelligent controller based on lambda for speed control of a three-phase inductor motor. *Emerging Science Journal*. 2023;7(3):676-90. <https://doi.org/10.28991/ESJ-2023-07-03-01>
13. Abdulwahid AH. Artificial intelligence-based control techniques for hvdc systems. *Emerging Science Journal*. 2023;7(2):643-53. <https://doi.org/10.28991/ESJ-2023-07-02-024>
14. Emelyanov AA, Ivanov SL, Shibanov DA. To the issue of assessing the impact of driver qualification on the technical condition of the excavator. *Mining information and Analytical Bulletin*. 2017(S38):442-53. <https://doi.org/10.25018/0236-1493-2017-12-38-442-453>
15. Velikanov VS. Mining excavator working equipment load forecasting according to a fuzzy-logistic model. *Записки Горного института*. 2020;241:29-36. <https://doi.org/10.31897/pmi.2020.1.29>
16. Baboli A, Ghodrattnama A, Tavakkoli-Moghaddam R. Comparing three proposed meta-heuristics to solve a new p-hub location-allocation problem. *International Journal of Engineering, Transactions C: Aspects*. 2013;26(9):1043-58. <https://doi.org/10.5829/idosi.ije.2013.26.09c.11>
17. Belikova DD, Morozov EV, Khisamutdinova EL. Optimizing control of mining machine power-units within the normal power setting range by means of engine oil quality monitoring. *Mining Informational and Analytical Bulletin*. 2021(6):95-103. https://doi.org/10.25018/0236_1493_2021_6_0_95
18. Velikanov V, Ilina E, Dyorina N. Structural and circuit design solution arguments of mine excavators ergonomics management. *Procedia Engineering*. 2016;150:1215-20. <https://doi.org/10.1016/j.proeng.2016.07.238>
19. Gorelik S, Grudinin V, Lecshinskiy V, Khaskelberg E. Method for assessing the influence of psychophysical state of drivers on control safety based on monitoring of vehicle movement parameters. *Transportation research procedia*. 2020;50:152-9. <https://doi.org/10.1016/j.trpro.2020.10.019>
20. Glemba K, Averianov Y. Substantiation of parameters and operation modes of device for thermal comfort of a mobile machine operator. *Procedia Engineering*. 2015;129:542-8. <https://doi.org/10.1016/j.proeng.2015.12.055>
21. Jain M, AGRAWAL C, Preeti C. Fuzzy reliability evaluation of a repairable system with imperfect coverage, reboot and common-cause shock failure. *International Journal of Engineering, Transactions C: Aspects*. 2012;25(3):231-8. <https://doi.org/10.5829/idosi.ije.2012.25.03c.07>
22. Korogodin AS, Ivanov SL. Maintenance and repair of drum mill trunnions of a floating mining equipment complex. *Sustainable development of mountain territories*. 2023;15(3):760-70. <https://doi.org/10.21177/1998-4502-2023-15-3-760-770>
23. Gromyka D, Gogolinskiy K. Method of state and residual resource assessment of excavator bucket tooth caps. *Russian Journal of Nondestructive Testing*. 2022;58(5):381-90. <https://doi.org/10.1134/S1061830922050035>
24. Balovnev V. Determination of optimal parameters and selection of road-building machines by the method of analysis of the fourth coordinate. *Moscow, MADI*. 2014;180.
25. Balovnev V. *Opredeleniye optimal'nykh parametrov i vybor zemleroynykh mashin v zavisimosti ot usloviy ekspluatatsii [Determination of optimal parameters and selection of earth-moving machines depending on operating conditions]*. M: MADI. 2010.
26. Dombrovsky N. *Multi-Bucket Excavators. Design, Theory and Calculation*. Mechanical Engineering: Moscow, Russia. 1972;432.
27. Makhno DE. To the methodology for assessing the level of qualification of an excavator driver. *Bulletin of the Irkutsk State Technical University*. 2011(12); 105-7.
28. Kostygova DM, Emelyanov AA. Simulation modeling of a quarry EKG-18R excavator manufactured by IZ-KARTEX LLC in the simulator for training machinists. *Mining Informational and Analytical Bulletin*, 2017(S23):177-84. <https://doi.org/10.25018/0236-1493-2017-10-23-177-184>
29. Bessonov AE, Shibanov DA, Mikhailov AV. The influence of the ergatic system on the working cycle of a quarry electric excavator. *Transport, mining and construction engineering: science and production*, 2022(15):136-41. <https://doi.org/10.26160/2658-3305-2022-15-136-141>

COPYRIGHTS

©2024 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 تجزیه و تحلیل روش های ارزیابی عملکرد موثر یک حفاری معدن الکتریکی را مورد بحث قرار می دهد. نتیجه گیری می شود که کارایی عملیاتی را نمی توان به صورت جداگانه از سیستم اپراتور-ماشین ارگاتیک در نظر گرفت. تعدادی از روش های ارزیابی سطح مهارت اپراتورها بر اساس داده های تجربی در نظر گرفته شده است. این مطالعه یک رویکرد یکپارچه برای ارزیابی کارایی عملیاتی یک حفاری معدن الکتریکی و در نتیجه سطح مهارت اپراتور را پیشنهاد کرد. این شامل ارزیابی زمان کار دستگاه و بهره وری انرژی فرآیند با استفاده از شبیه ساز بیل مکانیکی بود. مدت زمان چرخه کار بیل مکانیکی به عنوان نسبت متوسط زمان چرخه به اسمی برای یک نوع خاص از تجهیزات تخمین زده شده است. بهره وری انرژی فرآیند ارزیابی شده با نسبت شدت انرژی خاص در طول چرخه کاری حفاری به ارزش نظری شدت انرژی محاسبه شده بر اساس پارامترهای اسمی. بنابراین ، اگر معیار کنترل ماشین کم باشد ، می توانیم نتیجه بگیریم که کاهش زمان چرخه کار ضروری است ؛ اگر بهره وری انرژی فرآیند کم باشد ، باید توجه اصلی به افزایش درجه پر کردن سطل بیل مکانیکی شود.

of these models contains some parameters, that their values are playing an important role in grid analysis. Usually, some of these parameters are specified by their manufacturers and assumed to be constant. However, some changes in power or terminal voltage may cause special operational point that needs to update these parameters.

The bases of the ways for determination of SG parameters in practical mode are discussed in literature (1-3) that used in many analysis. In addition, Mouni et al. (2) have conducted parameter estimation by applying a short circuit and least the square method, which may cause damage to the SG's structure. The other methods that are related to the direction of this research discussed by Chowdhury and Senroy (4) used five measuring quantities of SG and importing them in the unscented Kalman filter (UKF); although an example of estimating the state of a power system using the Kalman filter are given in literature (3) in a basic form. However, it is not possible to measure all these quantities at a low cost in practice. Appropriate estimation was made by Ghahremani and Kamwa (5) Ren et al. (6), for just dynamic states of SG with low order models and extended Kalman filter (EKF). Huang et al. (7) introduced fourth-order model parameters estimation with square root unscented Kalman filter (SRUKF) that they do not contain more details in comparison with high-order models of SGs. To fix this issue, Geraldi et al. (8) Zhao and Mili (9), Valverde et al. (10) estimated parameters or dynamic states with UKF. However, in UKF algorithms some coefficients need to be set accurately and may be necessary to fix them in other equations or different SG's types. For both parameter and state estimation, good results with a modified type of UKF to be achieved. Rouhani and Abur (11) dond parameter estimation for fully regulated SG, but uses a noncomplex model like two axes model with UKF. González-Cagigal et al.(12) , Li et al. (13) used cubature kalman filter (CKF) to estimate the dynamic state of SG in detailed. CKF algorithm for implementation needs a powerful computing platform, but can clearly say that the results of this filter are one of the most exact results, also the phasor measurement units (PMUs) are widely used in data acquisition of synchronous generator operation parameters, which can capture the dynamic response of generators. For many reasons it is hard to transmit gigantic volumes of data to the information center due to limited communication bandwidth, so to reduce communication pressure, an improved regularized particle filter (IRPF) is designed to guarantee the estimation performance by Bai et al. (14).

According to reviews, this research is the first analysis to challenge the CDKF performance in estimating both parameters and dynamic states for two different dynamic models. It should be noted that in this research, the estimation process will be implemented on SGs behaviour consisting of increasing excitation

voltage and input torque for generator A and instant short circuit for another generator (i.e. B). The measurements also used in this work is the easily accessible outputs of SG. Furthermore, the algorithm used has good stability, and even it does not need to specify parameters like UKF, and it is compatible with a wide range of dynamic equations. Also Using the CDKF filter in the hybrid form, where the discrete measurements are placed next to the continuous state equations that are known as HCDKF, and because discrete measurements are used, the answers are closer to reality.

The remaining parts of this work are organised as follows. Section 2 is allocated to describe and formulate the CDKF algorithm and after that in section 3 presents the models of SG used in the estimation process. Match the models with the CDKF algorithm for dual estimation described in section 4. Two study cases, including generators A and B with their details for testing the accuracy of the process, is presented in section 5. The conclusion and recommendation are obtained from the results presented in section 6.

2. CENTRAL DIFFERENCE KALMAN FILTER

The basis of all Kalman filter derivatives goes back to its linear type and is used to filter noisy data, produce nonobservable states and forecast future state (15). But for nonlinear states, these particle filters have the ability to estimate states with different tension that can be considered (16, 17). In nonlinear forms of this filter, the central difference Kalman filter (CDKF) is one of the exact forms with good results. For implementing this filter to SG's dynamic, a complete description of the states and measurements is required. Equations 1 and 2 represent the usual form of the states and measurement, Also, special attention should be paid to the discontinuity and continuity of these equations with respect to time, which will be explained more (18, 19).

$$\dot{x}(t) = f(x(t), u(t)) + v(t) \quad (1)$$

$$y(t_k) = h(x(t_k), u(t_k)) + n(t_k) \quad (2)$$

where $x(t)$ represents a state vector, $u(t)$ is the system input, and $v(t)$ is the process noise which is assumed to be Gaussian with covariance R_v . These values together form a continuous nonlinear function as f . In the second equation, $y(t_k)$ represents available measurement which is described by the function h . The measurement noise $n(t_k)$ is assumed to be Gaussian with the covariance R_n at instant t_k . Since this filter is supposed to be implemented on generator equations, a hybrid form of the filter should be mentioned. Generators like many nonlinear systems have continuous state equations and discrete measurement equation. Because sensors which record the measurements based on a sample time and finally

generate a data series with a specified step, so the measurement equation is modified as follows (20):

$$y_k = h(x_k, u_k) + n_k \quad (3)$$

If the EKF algorithm was used to estimate, Equation 1 should linearize and then enter into the process (21). However, in CDKF, the state equation evaluates in a continuous nonlinear form, that is based on Sterling's polynomial interpolation functions (16). This method is known for approximating nonlinear function and uses a finite number of the main function evaluations instead of analytical derivatives. For scalar form and if supposed that $g(x)$ is a nonlinear function with a random variable x and its mean \bar{x} , the 2nd order case is given by:

$$g(x) = g(\bar{x}) + \tilde{D}_{\Delta x} g + \frac{1}{2!} \tilde{D}^2_{\Delta x} g \quad (4)$$

where $\tilde{D}_{\Delta x} g$ is the first and $\tilde{D}^2_{\Delta x} g$ is the second-order central divided difference operators and given by:

$$\tilde{D}_{\Delta x} g = (x - \bar{x}) \frac{g(\bar{x}+h) - g(\bar{x}-h)}{2h} \quad (5)$$

$$\tilde{D}^2_{\Delta x} g = (x - \bar{x})^2 \frac{g(\bar{x}+h) + g(\bar{x}-h) - 2g(\bar{x})}{h^2} \quad (6)$$

h is known as interval length in these equations, except that, the derivatives are not analytically solved by extending the sterling formula to make a multi-dimensional mode. It's necessary to provide a new variable that is stochastically decoupled. So z can be expressed as Equation 7 and then put in a nonlinear function to create $\tilde{g}(z)$ in Equation 8.

$$z = S_x^{-1} x \quad (7)$$

$$\tilde{g}(z) = g(S_x z) = g(x) \quad (8)$$

where S_x is obtained from the Cholesky factorization of covariance matrix x called P_x . Furthermore, by putting this transformation in the sterling formula instead $\tilde{D}_{\Delta x} g$ & $\tilde{D}^2_{\Delta x} g$, it's easy to calculate mean, covariance and cross-covariance, which is summarized in two stages in the following. At first, it is necessary to specify initial values as below:

$$\hat{x}_0 = E[x_0] \quad (9)$$

$$P_{x0} = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T] \quad (10)$$

where \hat{x}_0 is the initial estimation or mean value for the first step and P_{x0} is the initial estimation covariance that can have a significant influenced on the estimation result.

2. 1. Time Update Time update in CDKF is similar to UKF and starts by specifying the number of sigma points which have $2L+1$ points for each instant k . L is the state space dimension or in other words, the number of states. \hat{x}_{k-1} is the previous value of estimation to calculate sigma points (x_{k-1}) as follows:

$$x_{k-1} = [\hat{x}_{k-1} \quad \hat{x}_{k-1} + h\sqrt{P_{x_{k-1}}} \quad \hat{x}_{k-1} - h\sqrt{P_{x_{k-1}}}] \quad (11)$$

Now it's time to propagate sigma points through a state equation that can obtain,

$$x_{k/k-1} = f(x_{k-1}, u_{k-1}) \quad (12)$$

Then, using these points to calculate a priori estimation of the mean (\hat{x}_k) and covariance (\hat{P}_{x_k}) as follows:

$$\hat{x}_k = \sum_{i=0}^{2L} w_i^{(m)} x_{i,k/k-1} \quad (13)$$

$$P_{x_k} = \sum_{i=1}^L [w_i^{(c1)} (x_{i,k/k-1} - \hat{x}_{k/k-1})^2 + w_i^{(c2)} (x_{i,k/k-1} + x_{i+L,k/k-1} - 2\hat{x}_{k/k-1})^2] + R_v \quad (14)$$

where $w_i^{(m)}$ is the weighting vector of mean and $w_i^{(c1)}$ & $w_i^{(c2)}$ are weighting vectors of covariance and calculated by Equation 15. In addition, the power of two in Equation 14 represents the outer product.

$$\begin{cases} w_0^{(m)} = \frac{h^2 - L}{h^2} \\ w_i^{(m)} = \frac{1}{2h^2} \quad i = 1, \dots, 2L \\ w_i^{(c1)} = \frac{1}{4h^2} \quad i = 1, \dots, 2L \\ w_i^{(c2)} = \frac{h^2 - 1}{4h^2} \quad i = 1, \dots, 2L \end{cases} \quad (15)$$

2. 2. Measurement Update In this stage, sigma points must be recalculated, but this time with prior estimation (\hat{x}_k) and its covariance (\hat{P}_{x_k}) that comes from the previous stage as follows:

$$x_{k/k-1}^* = [\hat{x}_k \quad \hat{x}_k + h\sqrt{P_{x_k}^-} \quad \hat{x}_k - h\sqrt{P_{x_k}^-}] \quad (16)$$

Then new sigma points ($x_{k/k-1}^*$) are propagated through measurement function as Equation 16.

$$y_{k/k-1} = h(x_{k/k-1}^*) \quad (17)$$

With the weights were calculated in Equation 15, mean (\hat{y}_k), covariance (p_{y_k}) and cross-covariance ($p_{x_k y_k}$) Can be easily found in Equation 18.

$$\begin{cases} \hat{y}_k = \sum_{i=0}^{2L} w_i^{(m)} y_{i,k/k-1} \\ \hat{P}_{y_k} = \sum_{i=1}^L [w_i^{(c1)} (y_{i,k/k-1} - \hat{y}_{k/k-1})^2 + w_i^{(c2)} (y_{i,k/k-1} + y_{i+L,k/k-1} - 2\hat{y}_{k/k-1})^2] + R_n \\ p_{x_k y_k} = \sqrt{w_1^{(c1)} P_{x_k}^-} [y_{1:k,k/k-1} - y_{1+1:2L,k/k-1}]^T \end{cases} \quad (18)$$

And in the last step, one of the filter's output that named Kalman gain (k) is calculated in Equation 19.

$$K_k = p_{x_k y_k} p_{y_k}^{-1} \quad (19)$$

Then the final results of the estimation process (\hat{x}_k) as a posterior estimation and its covariance (p_{x_k}) Can be obtained,

$$\hat{x}_k = \hat{x}_k + k_k (y_k - \hat{y}_k) \quad (20)$$

$$p_{x_k} = p_{x_k}^- - k_k p_{y_k}^- k_k^T (21)$$

where these values are the final estimation for the k th instant, and the first step for $k+1$ th.

3. SYNCHRONOUS GENERATOR MODELS

One of the most important parts in dual estimation is recognizing the dynamic model of the system under discussion. In this work, two models of SG are described that if not done correctly, can cause bad results. The third and seventh order model of SG presented in the next two sections.

3.1. Third-Order Dynamic Model The third-order model is the flexible and accessible model in the dynamic analysis (19, 22). It's not very detailed and to achieve the equations, some assumptions must be taken. For example, neglecting the dynamic changes of stator and damper windings, stator resistance is assumed to be zero and also rotor speed is considered to be equal with synchronous speed (1p.u.). However, using this model is favored in the filter and makes it, easy to estimate. The equations in this model consist of rotor angle (δ), rotor speed (ω) and transient internal voltage of the armature (e'_q) that are calculated as follows:

$$\dot{\delta} = \omega (22)$$

$$\dot{\omega} = \frac{1}{J} (T_m - T_e - D\omega) (23)$$

$$\dot{e}'_q = \frac{1}{T'_{do}} (E_{fd} - e'_q - (x_d - x'_d)i_d) (24)$$

To complete the above equations, mention the direct and quadrature current (i_d & i_q) and electrical torque (T_e) is required, which represent in Equations 25 to 27, respectively. Unknown parameters and states that should be estimated are discussed separately in the following sections.

$$i_d = \frac{e'_q - v \cos \delta}{x'_d} (25)$$

$$i_q = \frac{v \sin \delta}{x_q} (26)$$

$$T_e = P_e \cong \frac{v}{x'_d} e'_q \sin \delta + \frac{v^2}{2} \left(\frac{1}{x_q} \right) (27)$$

After determining the state equations, it's time to specify the measurement equations. There are many quantities in a generator that have a certain equation, but that's not enough alone. Quantities are selected, which can be easily measured. So in this work terminal current (i_t), terminal voltage (v_t), electrical power (p_e) and rotor speed (ω_r) are used. The following procedure can be followed to extract their equations. Equation 28, stator current is defined.

$$i_t = \sqrt{i_d^2 + i_q^2} (28)$$

For v_t , it's needed to calculate direct and quadrature voltage as follows:

$$v_d = \left(1 - \frac{x_e}{x_q}\right) v \sin \delta (29)$$

$$v_q = \left(1 - \frac{x_e}{x'_q}\right) v \cos \delta \frac{x_e e'_q}{x'_d} (30)$$

and then put them in Equation 31 to form the following equation.

$$v_t = \sqrt{v_d^2 + v_q^2} (31)$$

The Equations 23 and 27 can be used for ω_r and p_e respectively to establish a good connection between the states and measurement equations. In addition, this combination gives accurate results in the estimation process. All equations in this section are implemented on generator A and the details are given in the appendix.

3.2. Seventh-Order Dynamic Model The seventh-order model has more details than the third-order model. That is why implementing the estimation algorithm for this model is a challenging issue. There are five electrical equations and two mechanical equations in this model which form the dynamic states. Unlike the third-order model, the damper winding behavior is not neglected here. Krause et al. (23) is used to review these equations. The magnetic flux of the stator, rotor and damper windings are considered as electrical equations and written as follows:

$$\dot{\phi}_{qs} = \omega_b [v_q - \frac{\omega_r}{\omega_b} \phi_{ds} + \frac{r_s}{x_{ls}} (\phi_{mq} - \phi_{qs})] (32)$$

$$\dot{\phi}_{ds} = \omega_b [v_d - \frac{\omega_r}{\omega_b} \phi_{qs} + \frac{r_s}{x_{ls}} (\phi_{md} - \phi_{ds})] (33)$$

$$\dot{\phi}_{kq} = \omega_b [v_{kq} + \frac{r_{kq}}{x_{lkq}} (\phi_{md} - \phi_{kq})] (34)$$

$$\dot{\phi}_{kd} = \omega_b [v_{kd} + \frac{r_{kd}}{x_{lkd}} (\phi_{md} - \phi_{kd})] (35)$$

$$\dot{\phi}_{fd} = \omega_b [v_{fd} + \frac{r_{fd}}{x_{lfd}} (\phi_{md} - \phi_{fd})] (36)$$

In coordinates transformation from abc to qdo the magnetic flux of the o axis (ϕ_{os}) becomes zero. ϕ_{md} and ϕ_{mq} are defined as middle variables to make writing equations easier and calculated as follows:

$$\begin{cases} \phi_{mq} = x_{aq} \left(\frac{\phi_{qs}}{x_{ls}} + \frac{\phi_{kq}}{x_{lkq}} \right) \\ \phi_{md} = x_{ad} \left(\frac{\phi_{ds}}{x_{ls}} + \frac{\phi_{kd}}{x_{lkq}} + \frac{\phi_{fd}}{x_{lfd}} \right) \\ x_{aq} = \left(\frac{1}{x_{mq}} + \frac{1}{x_{ls}} + \frac{1}{x_{lkq}} \right)^{-1} \\ x_{ad} = \left(\frac{1}{x_{md}} + \frac{1}{x_{ls}} + \frac{1}{x_{lkq}} + \frac{1}{x_{lfd}} \right)^{-1} \end{cases} (37)$$

and mechanical equations also calculated as:

$$\left(\frac{\dot{\omega}_r}{\omega_b}\right) = \frac{1}{2H} (T_L - T_e) \quad (38)$$

$$\dot{\delta} = \omega_b \left(\frac{\omega_r}{\omega_b} - 1\right) \quad (39)$$

where ω_b is the synchronous speed and equal to $2\pi f$. Now, and after completing the state equations, it is time to choose the measurable quantities.

In this model, rotor speed, field winding current (i_{fd}), stator voltage and current were chosen, and they are calculated according to Equations 39, 40, 28 and 31, respectively. Except that the direct and quadrature voltages are calculated with Equations 41 and 42.

$$i_{fd} = \frac{1}{x_{jfd}} (\phi_{fd} - \phi_{md}) \quad (40)$$

$$v_d = v \sin \delta - x_e i_q \quad (41)$$

$$v_q = v \cos \delta - x_e i_d \quad (42)$$

As it was reviewed, this model has seven states and nine parameters which should be estimated separately. These equations together, form the structure of generator B, which its real values for simulation are listed in the appendix.

4. IMPLEMENTATION OF CDKF

The CDKF is the most accurate filter among all sigma point Kalman filters and uses the Sterling interpolation formula. This filter has the computational complexity in the order of EKF but achieves a second or higher-order accuracy in the posterior mean and covariance which can be observed in some recent works (24). Any filter from sigma points filters or even a linear or nonlinear form of them can be used for state, parameter and dual estimation. However, for each of them, it is necessary to modify the base filter's structure. There are two approaches for dual estimation. One of them uses two filters that work together to estimate states and parameters. The other method uses just one filter, and the equations are written in such a way that states and parameters are estimated simultaneously. The second type is used in this work. Dual estimation has its difficulties; for example, González-Cagigal et al. (12) used two steps to estimate. In the first step, implement the algorithm on a simple model and then add other components and estimate whole parameters. However, in this work, the estimation process is implemented on complete forms of equations. Furthermore, to understand the structure of dual estimation in the form of a joint filter, Qi et al. (25) have used. According to the state equations from section (3), unknown states in third and seventh-order models of SG can be written as:

$$x_{(3th-order)}^T = [\delta, \omega_r, e'_q]$$

and

$$x_{(7th-order)}^T = [\varphi_{qs}, \varphi_{ds}, \varphi_{kq}, \varphi_{kd}, \varphi_{fd}, \omega_r, \delta]$$

respectively. In addition, the unknown parameters vector (w) for the mentioned models are

$$w_{(3th-order)}^T = [j, D, T'_{do}, x_d, x_q, x'_d]$$

and

$$w_{(7th-order)}^T = [r_s, r_{kq}, r_{kd}, r_{fd}, x_{ls}, x_{lkq}, x_{lkd}, x_{lfd}, H]$$

then the augmented vector of states and parameters ($x_{aug} = [x^T, w^T]^T$) must be formed, which is an important part of the estimation process. The size of the augment vector for the third-order model is $L_{(3th-order)} = 9$ and $L_{(7th-order)} = 16$ for the other. Now the augment vector should be replaced with the state vector in Equations 3 and 1 to perform the new structure as follows:

$$\begin{bmatrix} x_k \\ w_k \end{bmatrix} = \begin{bmatrix} f(x_k, u_k, w_k) \\ I w_k \end{bmatrix} \begin{bmatrix} v_k \\ r_k \end{bmatrix} \quad (43)$$

$$y_k = [I \quad 0 \quad \dots \quad 0] \begin{bmatrix} x_k \\ w_k \end{bmatrix} + n_k \quad (44)$$

where r_k is the parameters noise and depending on their equation, their values are set. For the measurement vector $y_{(3th-order)}^T = [\omega_r, I_s, v_d, p_e]$ is used for generator A and $y_{(7th-order)}^T = [\omega_r, I_s, v_d, i_f]$ is used for generator B, which their equations have already been specified. Finally, it should be noted that the input vector (u) consists of excitation voltage (v_f) and input torque (T_L) in case (1) and terminal voltage in case (2).

5. CASE STUDIES

CDKF algorithm is tested on generators A & B, which reviewed their equations in the previous sections. For measuring the output signals of generators, their structure was simulated in Matlab's code environment. To better determination of each case's details, two sections are considered.

5.1. Case1: Generator A As shown in Figure 1, generator A connected to an infinite bus in a steady state condition through a line with impedance $j0.2$ p.u. Generator outputs, which should enter to estimation process, are recorded by sensors at sample time $\Delta t = 1$ ms. They will be ready to use after the noise is added to them, which is shown in Figure 2. The additional noise has a Gaussian distribution with zero mean and standard deviation $R_n = 10^{-5}$.

The algorithm performance will be challenged in two changes. The first one is increasing the field voltage by 20% its real value at second 2, and after 10 seconds, the next change will happen with an increase of 2% in input

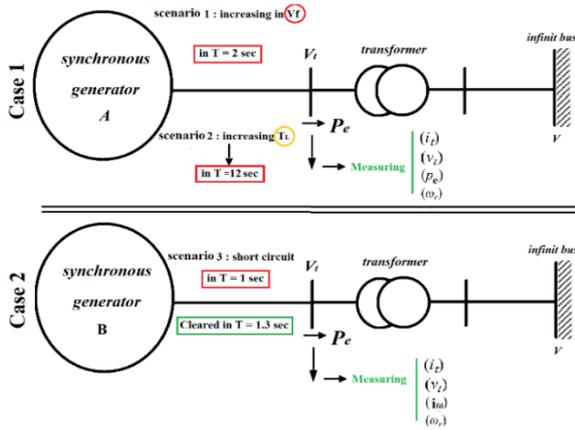


Figure 1. Structure of the system under study

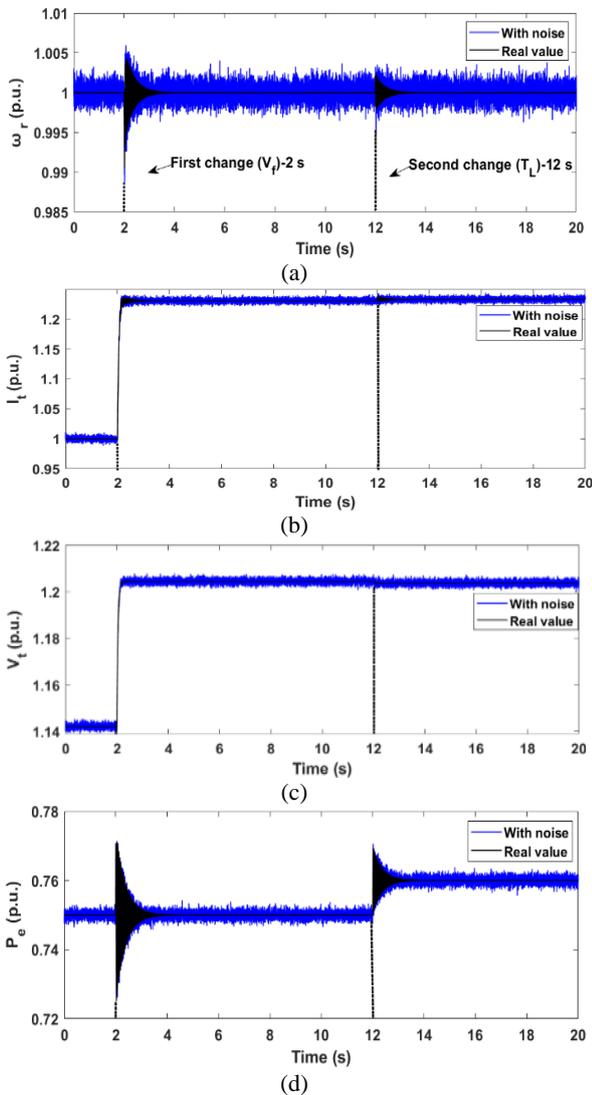


Figure 2. Diagrams of measurement changes for 20 seconds in generator A (a) Real and noisy values of rotor speed (b) Real and noisy values of terminal current (c) Real and noisy values of terminal voltage (d) Real and noisy values of electrical power

torque. Steady state calculation before changes take effect, listed in the appendix.

In this case, process noise (r_k) that being in state equations, is specified by its covariance matrix (Q_x) and also the parameter noise with covariance matrix (Q_w). These covariance matrixes combined each other as a diagonal form that, being ready for the dual estimation purpose as $Q_{CDKF} = \text{diag}(Q_x^T, Q_w^T)$. The usual value of Q_x in the UKF algorithm is 10^{-6} like Qi et al. (25), but in this work for better results, Q_{CDKF} consider as:

$$Q_{CDKF}^T = \text{diag}([10^{-6}, 10^{-6}, 10^{-7}, 10^{-5}, 10^{-5}, 10^{-6}, 10^{-5}, 10^{-5}])$$

The initial value of parameters is an important issue in the implementation algorithm. For this case, initial values are set at 80% and 130% of their real values, and for both of them, estimation results for states and parameters are shown in Figures 3, 4 and 5. Their results are named HCDKF, because the filter reformed for the SG's equations in hybrid structure and defined in section 2. In addition, the initial estimation error covariance, which indicates the confidence to the initial parameters is a diagonal matrix as:

$$P_{x_0}^T = \text{diag}([10^{-5}, 10^{-5}, 10^{-5}, 10^{-3}, 10^{-3}, 10^{-4}, 10^{-3}, 10^{-3}])$$

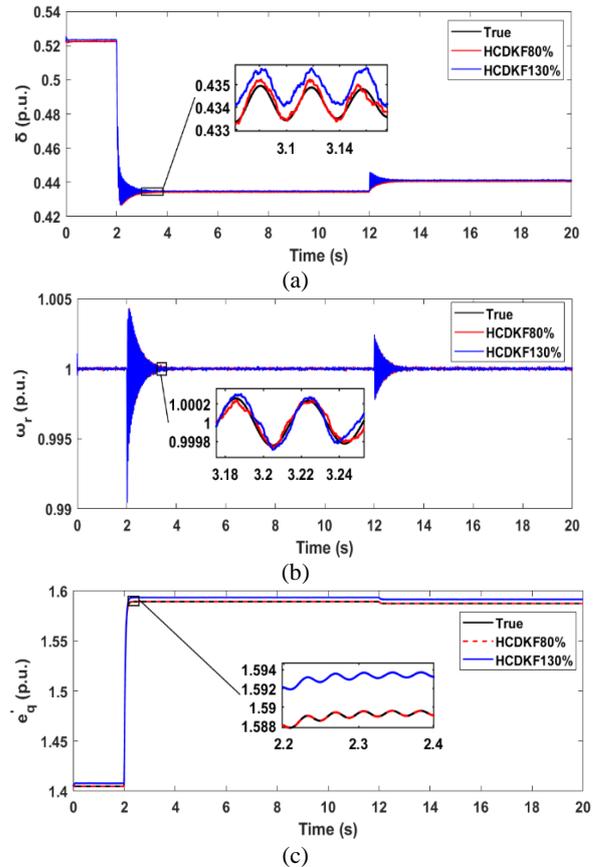


Figure 3. Diagrams of state estimations in generator A (a) Diagrams of true and estimated values of δ (b) Diagrams of true and estimated values of ω_r (c) Diagrams of true and estimated values of e'_q

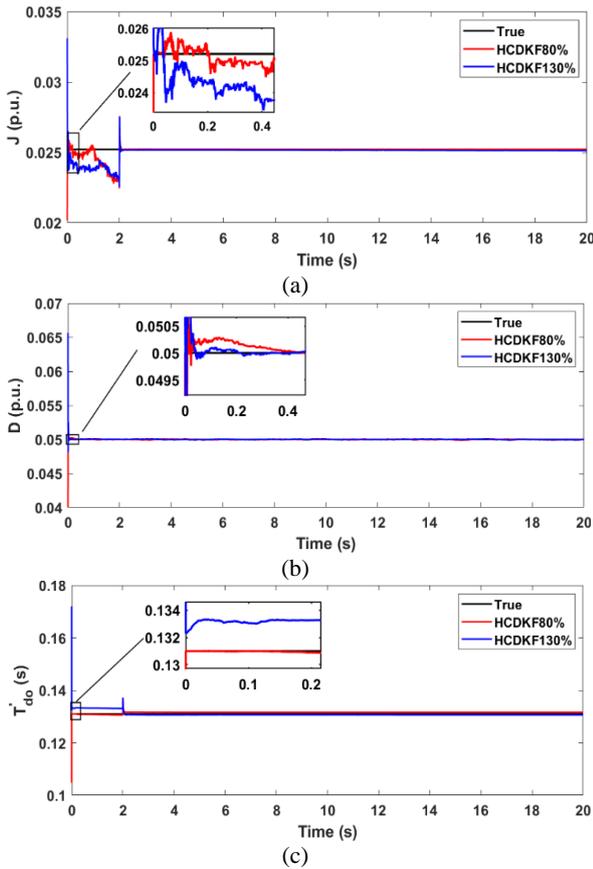


Figure 4. Diagrams of parameters estimation in generator A (a) Diagrams of true and estimated values of J (b) Diagrams of true and estimated values of D (c) Diagrams of true and estimated values of T'_{do}

For better comparison, the final estimated values in the second 20 are listed in Table 1. Two rows of this table are allocated to UKF estimation with similar noise to

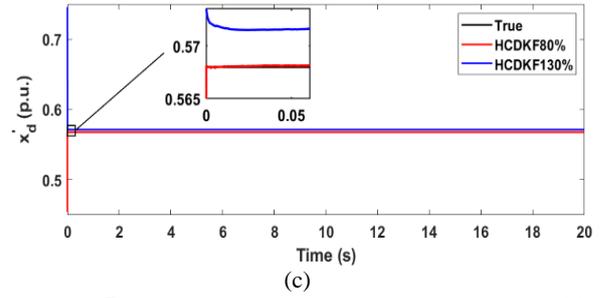
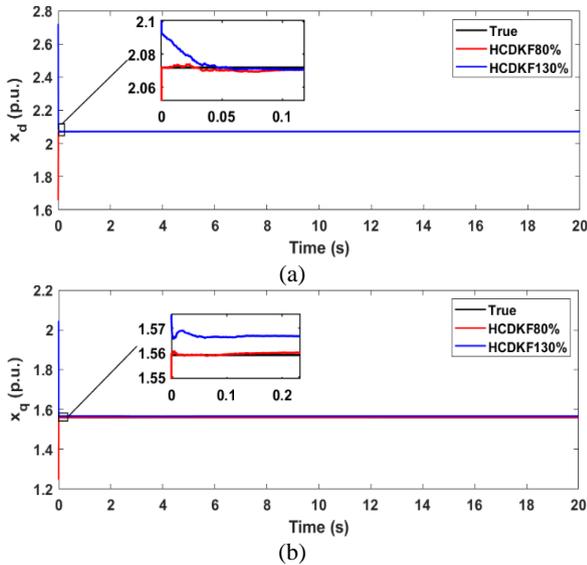


Figure 5. Diagrams parameter estimations in generator A (a) Diagrams of true and estimated values of x'_d (b) Diagrams of true and estimated values of x_q (c) Diagrams of true and estimated values of x'_d

CDKF and tunable parameters $\alpha = (0.5)$, $\beta = (2)$ and $\kappa = (-6)$. The only tunable parameter in CDKF is h, which is equal to $\sqrt{3}$. The time duration (20 S) seems to be sufficient because the filter responses are being fixed. As it seen, estimation values are very close to the real values, also tracking states by the filter is admirable. On the other hand, to evaluate the estimation accuracy, relative errors for the last data in the filter's memory are calculated and listed in Table 2.

TABLE 1. Estimations of CDKF and UKF for generator A parameters in two initial conditions

Parameters	Real values	CDKF estimations		UKF estimations	
		80% initial valve	130% initial value	80% initial valve	130% initial value
J (p.u.)	0.0252	0.025193	0.0251217	0.025258	0.026209
D (p.u.)	0.05	0.049956	0.050007	0.050016	0.050336
T'_{do} (s)	0.1310	0.131622	0.130642	0.12787	0.126155
x_d (p.u.)	2.072	2.07196	2.071871	2.07199	2.092736
x_q (p.u.)	1.559	1.55959	1.56673	1.55335	1.56293
x'_d (p.u.)	0.568	0.568036	0.571634	0.56776	0.56858

TABLE 2. Relative errors for CDKF and UKF estimations in generator A with two initial conditions

Parameters	CDKF relative errors (%)		UKF relative errors (%)	
	80% initial valve	130% initial value	80% initial valve	130% initial value
J (p.u.)	0.0253	0.3104	0.2337	4.004
D (p.u.)	0.08161	0.0159	0.0321	0.6735
T'_{do} (s)	0.4753	0.2727	2.3819	3.6981
x_d (p.u.)	0.0015	0.062	3.53×10^{-5}	1.0008
x_q (p.u.)	0.0383	0.4961	0.3472	0.2525
x'_d (p.u.)	0.6440	0.6399	0.0409	0.1073

5. 2. Case2: Generator B For generator B, using the same connection to the grid as generator A, but the difference is that, in this case, a three-phase short circuit occurred at generator terminal in second 1 and cleared in 0.3 second later. Measurements change with the added noise to them showed in Figure 6. The noises have a Gaussian distribution with zero mean and standard deviation $R_n = 10^{-5}$.

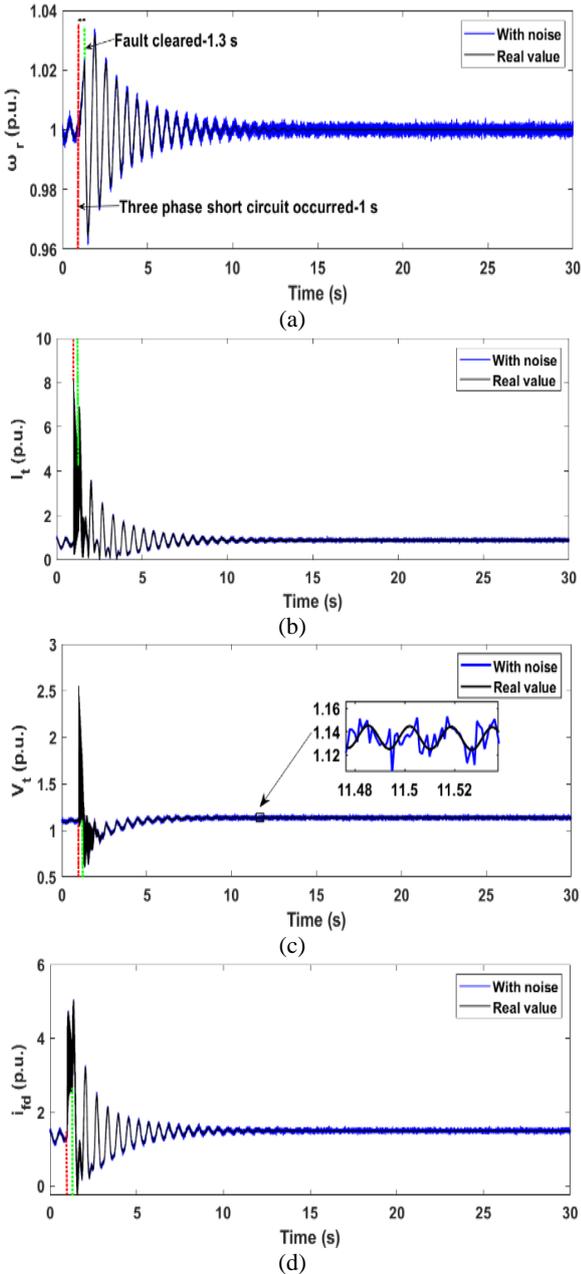


Figure 6. Diagrams of measurement changes for 30 seconds in generator B a) Real and noisy values of rotor speed b) Real and noisy values of terminal current c) Real and noisy values of terminal voltage d) Real and noisy values of filed current

Since this model has more details, the generator's operation is simulated for 30 seconds and filter should estimate states and parameters in this duration. Process and measurement noises are defined by 16×16 diagonal matrix (Q_{CDKF}) and in this case,

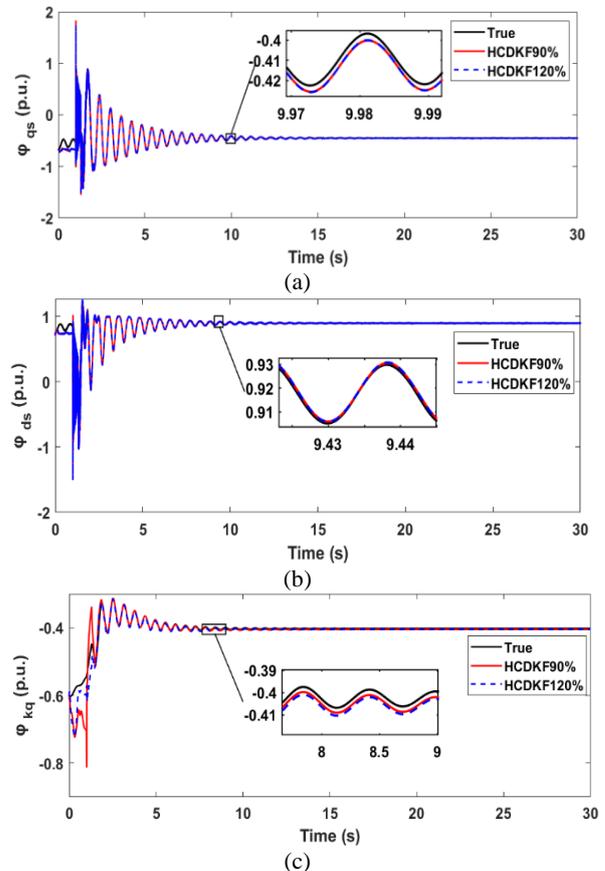
$$Q_{CDKF}^T = \text{diag} ([10^{-6}, 10^{-5}, 10^{-5}, 10^{-5}, 10^{-6}, 10^{-6}, 10^{-6}, 10^{-7}, 10^{-8}, 10^{-7}, 10^{-8}, 10^{-6}, 10^{-6}, 10^{-6}, 10^{-6}, 10^{-7}])$$

The initial values of parameters set 90% and 120% of their real values and the covariance error of the initial estimation (P_{x0}) defines as:

$$P_{x0}^T = \text{diag} ([10^{-6}, 10^{-6}, 10^{-6}, 10^{-6}, 10^{-6}, 10^{-6}, 10^{-6}, 10^{-4}, 10^4, 10^4, 10^4, 10^4, 10^4, 10^4, 10^4, 10^4])$$

with these values for noise and initial objects, as mentioned above, Figures 7 and 8 show the states and parameter estimation with both initial values of parameters. And Figures 9 and 10 show the performance of the filter to estimate the parameters. The result figures for this model, which are estimated by only four measurements signal are very acceptable.

The last estimation in 30 seconds for CDKF with $h = \sqrt{3}$ and for UKF with $\alpha = (1)$, $\beta = (2)$ and $\kappa = (-13)$ with similar noises to the CDKF process are listed in Table 3. For better comparison, relative and definite errors are



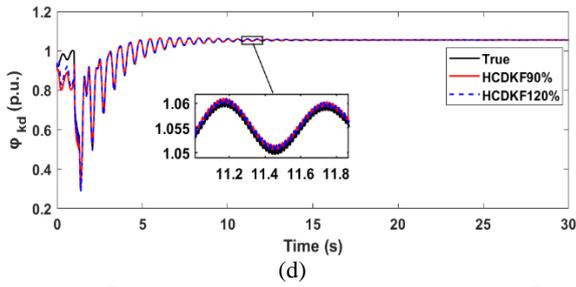


Figure 7. Diagrams of state estimations in generator B (a) Diagrams of true and estimated values of φ_{qs} (b) Diagrams of true and estimated values of φ_{ds} (c) Diagrams of true and estimated values of φ_{kq} (d) Diagrams of true and estimated values of φ_{kd}

presented in Table 4; which shows CDKF has done its job properly and in some parameters better than UKF

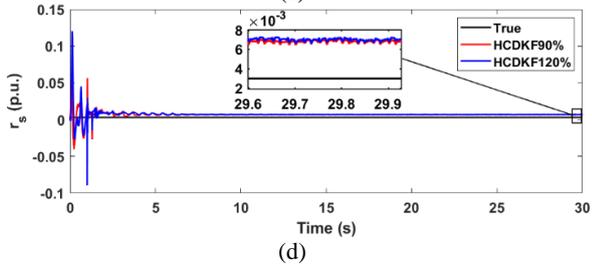
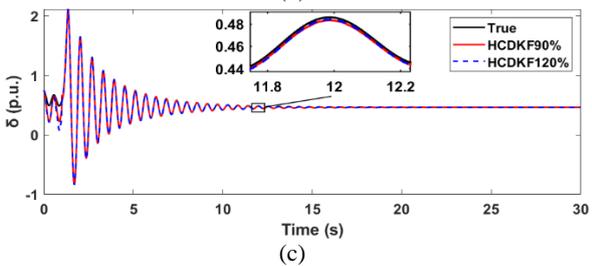
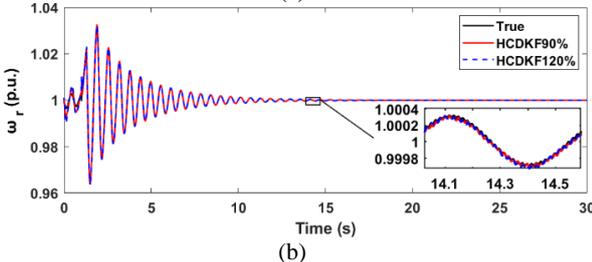
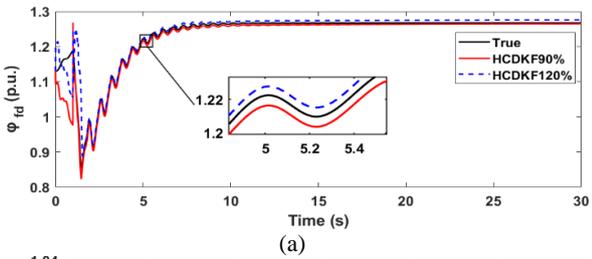


Figure 8. Diagrams of state and parameter estimations in generator B (a) Diagrams of true and estimated values of φ_{fd} (b) Diagrams of true and estimated values of ω_r (c) Diagrams of true and estimated values of δ (d) Diagrams of true and estimated values of r_s

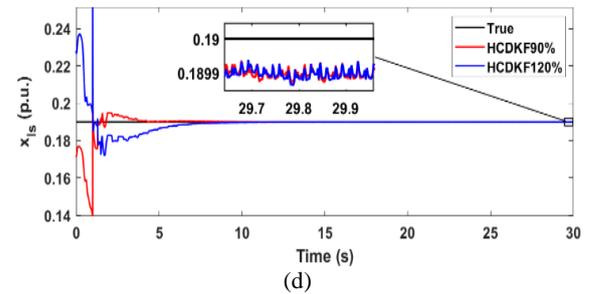
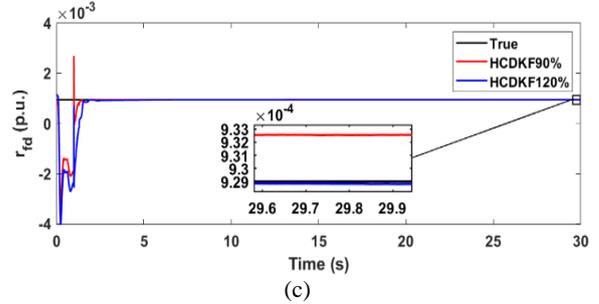
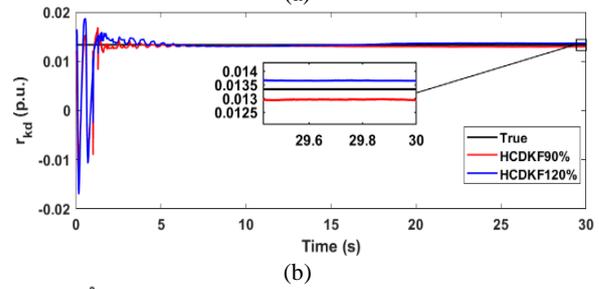
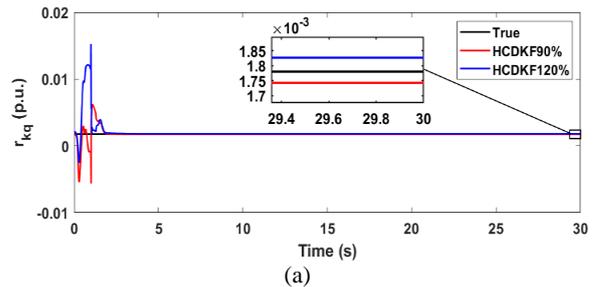
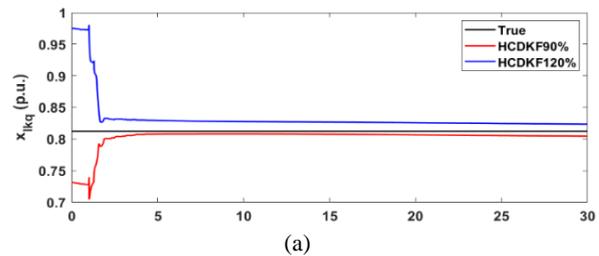


Figure 9. Diagrams of parameter estimations in generator B (a) Diagrams of true and estimated values of r_{kq} (b) Diagrams of true and estimated values of r_{kd} (c) Diagrams of true and estimated values of r_{fd} (d) Diagrams of true and estimated values of x_{1s}

except r_s . However, the important thing about the filter is the number of tunable parameters and its compatibility with many structures.



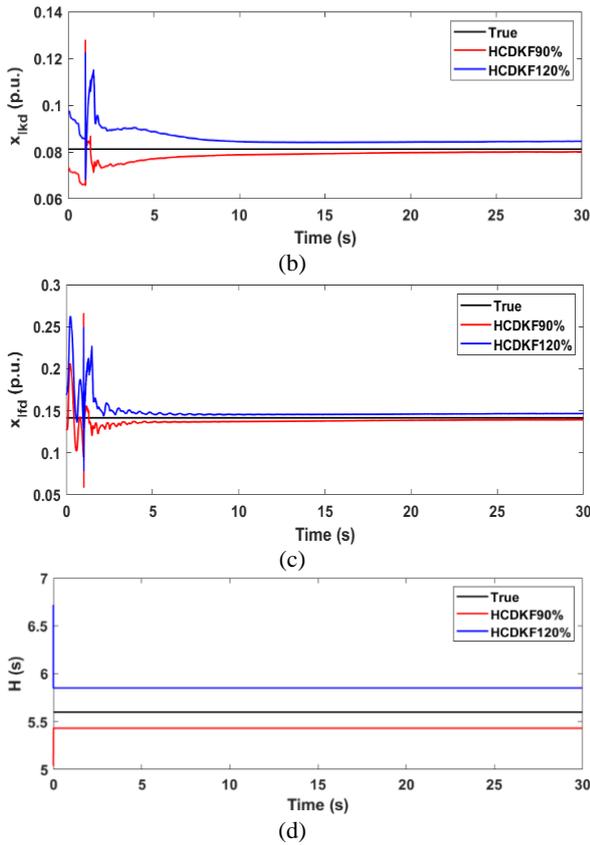


Figure 10. Diagrams of parameter estimations in generator B (a) Diagrams of true and estimated values of x_{ikd} (b) Diagrams of true and estimated values of x_{lkd} (c) Diagrams of true and estimated values of x_{lfd} (d) Diagrams of true and estimated values of H

Steady states equations and real values of generators A and B defined in Tables 5 and 6.

TABLE 3. Estimations of CDKF and UKF for generator B parameters in two initial conditions

Parameters	Real values	CDKF estimations		UKF estimations	
		90% initial valve	120% initial value	90% initial valve	120% initial value
r_s (p.u.)	0.003	0.00676	0.006953	0.0306	0.00287
r_{kq} (p.u.)	0.0178	0.001742	0.001826	0.001753	0.00167
r_{kd} (p.u.)	0.0133	0.012947	0.013647	0.013618	0.01267
r_{fd} (p.u.)	0.000929	0.000932	0.000928	0.000932	0.000893
x_{ls} (p.u.)	0.19	0.18988	0.189887	0.18989	0.18060
x_{lkq} (p.u.)	0.8125	0.80454	0.82359	0.8176	0.77187
x_{lkd} (p.u.)	0.08125	0.080166	0.08463	0.08320	0.07719
x_{lfd} (p.u.)	0.1414	0.139176	0.14650	0.14417	0.14037
H (s)	5.6	5.4320	5.8520	6.1600	5.3200

TABLE 4. Relative errors for CDKF and UKF estimations in generator B with two initial conditions

Parameters	Real values	CDKF relative errors (%)		UKF relative errors (%)	
		90% initial valve	120% initial value	90% initial valve	120% initial value
r_s (p.u.)	0.003	125.25	131.79	2.003	4.2141
r_{kq} (p.u.)	0.0178	2.1175	2.6159	1.447	5.6456
r_{kd} (p.u.)	0.0133	2.9508	2.3097	2.095	4.9994
r_{fd} (p.u.)	0.000929	0.3828	0.0204	0.3532	3.8986
x_{ls} (p.u.)	0.19	0.0610	0.0593	1.959	4.945
x_{lkq} (p.u.)	0.8125	0.9796	1.3652	1.997	4.9996
x_{lkd} (p.u.)	0.08125	1.3337	4.1670	2.067	4.9939
x_{lfd} (p.u.)	0.1414	1.5528	3.6079	0.0919	0.7246
H (s)	5.6	2.9997	4.5000	1.998	5.000

6.CONCLUSION

In this paper, a CDKF algorithm is implemented on two dynamic models of different generators, with the purpose of dual estimation of states and parameters.

Although many kinds of Kalman estimation algorithms EKF, UKF and CKF with different levels of accuracy tested on SG's equations. In each of them, some parameters have to be set. While in CDKF just by tuning one parameter can achieve good results the same as other methods. In addition, the 7-order model used in this work has 16 unknown states and parameters which are estimated only by four external measurements from accessible quantities of the generator.

The case studies have shown that the algorithm has acceptable accuracy in different working conditions and also in different initial values of parameters for the first step of estimation. Furthermore, in some parameters, it works better than UKF.

Also, due to the volume of equations and computational complexity of the CDKF algorithm, special attention can be paid to the implementation of this process on a low-cost and powerful computing platform such as PIC as future activities, and its results will be evaluated in the form of another research.

7. REFERENCES

1. Yaghibi H, Ansari K, Mashhadi HR. Stator turn-to-turn fault detection of synchronous generator using total harmonic distortion (THD) analyzing of magnetic flux linkage. Iranian Journal of Science and Technology Transactions of Electrical Engineering. 2013;37(E2):161. <https://doi.org/10.22099/IJSTE.2013.1886>

2. Mouni E, Tnani S, Champenois G. Synchronous generator modelling and parameters estimation using least squares method. *Simulation Modelling Practice and Theory*. 2008;16(6):678-89. <https://doi.org/10.1016/j.simpat.2008.04.005>
3. Mahdavi Tabatabayee N, Kenarangui R. On line electric power systems state estimation using kalman filtering (research note). *International Journal of Engineering*. 1995;8(4):233-5.
4. Chowdhury SD, Senroy N, editors. PMU data based online parameter estimation of synchronous generator. 2016 IEEE 6th International Conference on Power Systems (ICPS); 2016: IEEE.
5. Ghahremani E, Kamwa I. Dynamic state estimation in power system by applying the extended Kalman filter with unknown inputs to phasor measurements. *IEEE Transactions on Power Systems*. 2011;26(4):2556-66. <https://doi.org/10.1109/tpwrs.2011.2145396>
6. Ren P, Lev-Ari H, Abur A, editors. Robust continuous-discrete extended Kalman filter for estimating machine states with model uncertainties. 2016 Power Systems Computation Conference (PSCC); 2016: IEEE.
7. Huang M, Li W, Yan W. Estimating parameters of synchronous generators using square-root unscented Kalman filter. *Electric Power Systems Research*. 2010;80(9):1137-44. <https://doi.org/10.1016/j.epsr.2010.03.007>
8. Geraldi Jr EL, Fernandes TC, Ramos RA. A UKF-based approach to estimate parameters of a three-phase synchronous generator model. *Energy Systems*. 2018;9(3):573-603. <https://doi.org/10.1007/s12667-018-0280-1>
9. Zhao J, Mili L. A decentralized H-infinity unscented Kalman filter for dynamic state estimation against uncertainties. *IEEE Transactions on Smart Grid*. 2018;10(5):4870-80. <https://doi.org/10.1109/tsg.2018.2870327>
10. Valverde G, Kyriakides E, Heydt GT, Terzija V. Nonlinear estimation of synchronous machine parameters using operating data. *IEEE Transactions on Energy Conversion*. 2011;26(3):831-9. <https://doi.org/10.1109/tec.2011.2141136>
11. Rouhani A, Abur A. Constrained iterated unscented Kalman filter for dynamic state and parameter estimation. *IEEE Transactions on Power Systems*. 2017;33(3):2404-14. <https://doi.org/10.1109/tpwrs.2017.2764005>
12. González-Cagigal M, Rosendo-Macías J, Gómez-Expósito A. Parameter estimation of fully regulated synchronous generators using unscented Kalman filters. *Electric power systems research*. 2019;168:210-7. <https://doi.org/10.1016/j.epsr.2018.11.018>
13. Li Y, Li J, Qi J, Chen L. Robust cubature Kalman filter for dynamic state estimation of synchronous machines under unknown measurement noise statistics. *IEEE Access*. 2019;7:29139-48. <https://doi.org/10.1109/access.2019.2900228>
14. Bai X, Qin F, Ge L, Zeng L, Zheng X. Dynamic state estimation for synchronous generator with communication constraints: An improved regularized particle filter approach. *IEEE Transactions on Sustainable Computing*. 2022. <https://doi.org/10.1109/TSUSC.2022.3221090>
15. Siddharth D, Saini D, Singh P. An efficient approach for edge detection technique using kalman filter with artificial neural network. *International Journal of Engineering, Transactions C: Aspects*. 2021;34(12):2604-10. <https://doi.org/10.5829/ije.2021.34.12c.04>
16. Van Der Merwe R. Sigma-point Kalman filters for probabilistic inference in dynamic state-space models: Oregon Health & Science University; 2004.
17. Hooshmand M, Yaghobi H, Jazaeri M. Irradiation and Temperature Estimation with a New Extended Kalman Particle Filter for Maximum Power Point Tracking in Photovoltaic Systems. *International Journal of Engineering, Transactions C: Aspects*. 2023;36(6):1099-113. <https://doi.org/10.5829/IJE.2023.36.06C.08>
18. Nejadi Z, Faraji A, Abedi M. Robust three stage central difference kalman filter for helicopter unmanned aerial vehicle actuators fault estimation. *International Journal of Engineering, Transactions B: Applications*. 2021;34(5):1290-6. <https://doi.org/10.5829/ije.2021.34.05b.23>
19. Dehghani M, Nikravesh S. Nonlinear state space model identification of synchronous generators. *Electric Power Systems Research*. 2008;78(5):926-40. <https://doi.org/10.1016/j.epsr.2007.07.001>
20. Sahel Hanane K, Abderrazak L, Adlene R, Mohamed A, Mohamed K. Fuzzy Logic Control of Maximum Power Point Tracking Controller in an Autonomous Hybrid Power Generation System by Extended Kalman Filter for Battery State of Charge Estimation. *International Journal of Engineering, Transactions B: Applications*. 2023;36(2):199-214. <https://doi.org/10.5829/IJE.2023.36.02B.02>
21. Niazi S. Estimation of LOS rates for target tracking problems using EKF and UKF algorithms-a comparative study. *International Journal of Engineering, Transactions B: Applications*. 2015;28(2):172-9. <https://doi.org/10.5829/idosi.ije.2015.28.02b.02>
22. Nasab MR, Yaghobi H. A real-time out-of-step protection strategy based on instantaneous active power deviation. *IEEE Transactions on Power Delivery*. 2020;36(6):3590-600. <https://doi.org/10.1109/TPWRD.2020.3045276>
23. Krause PC, Wasynczuk O, Sudhoff SD, Pekarek S. Analysis of electric machinery and drive systems: Wiley Online Library; 2002.
24. Ebrahimi F, Abedi M. Design of a robust central difference Kalman filter in the presence of uncertainties and unknown measurement errors. *Signal processing*. 2020;172:107533. <https://doi.org/10.1016/j.sigpro.2020.107533>
25. Qi J, Sun K, Wang J, Liu H. Dynamic state estimation for multi-machine power system by unscented Kalman filter with enhanced numerical stability. *IEEE Transactions on Smart Grid*. 2016;9(2):1184-96. <https://doi.org/10.1109/tsg.2016.2580584>

8. APPENDIX

<p>The main variables of 3-order model are:</p> <p>e'_q Transient internal voltage of armature δ Rotor angle ω_r Rotor speed E_{fd} The equivalent EMF in the excitation coil T_e Output electrical torque v_t Generator terminal voltage V Infinite bus voltage i_d, i_q Direct and quadrature axis stator currents T'_{do} Direct axis transient time constant x'_d Direct axis transient reactance J Rotor inertia D Damping factor x_d Direct axis reactance x_q Quadrature axis reactance</p>	<p>The main variables of 7-order model are:</p> <p>ϕ_{qs} Quadrature axis magnetic flux ϕ_{ds} Direct axis magnetic flux ϕ_{kq} Quadrature axis damper magnetic flux ϕ_{kd} Direct axis damper magnetic flux ϕ_{fd} Field winding magnetic flux i_{fd} Field winding current r_s Stator resistance r_{kq} Quadrature axis damper resistance r_{kd} Direct axis damper resistance r_{fd} Field winding resistance x_{ls} Stator leakage reactance x_{lkq} Quadrature axis damper leakage reactance x_{lkd} Direct axis damper leakage reactance x_{lfd} Field winding leakage reactance H Inertia constant v_q Quadrature axis voltage v_d Direct axis voltage v_{kq} Quadrature axis damper voltage v_{kd} Direct axis damper voltage v_{fd} Field winding voltage θ Generator voltage angle</p>
--	--

TABLE 5. Generator A [20]

Real values	Steady state calculations
$\omega_b = 120 \times \pi$ (rad/s)	$v = 1$ (p.u.)
$S = 1$ (p.u.)	$x_e = 0.2$ (p.u.)
$T_m = 0.8$ (p.u.)	$i_s = \frac{s}{v} = 1$ (p.u.)
$J = 0.0252$ (p.u.)	$\theta = \cos^{-1} \left(\frac{(T_m \cdot D \cdot I)}{s} \right) = 0.7227$ (p.u.)
$D = 0.05$ (p.u.)	$\delta = \tan^{-1} \left(\frac{x_q \times i_s \times \cos \theta}{v + x_q \times \sin \theta} \right) = 0.5223$ (p.u.)
$x'_d = 0.568$ (p.u.)	$\omega = 1$ (pu)
$x_d = 2.072$ (p.u.)	$E_{fd} = v \times \cos \delta + x_q \times i_s \times \sin(\delta + \theta) = 2.8297$ (p.u.)
$x_q = 1.559$ (p.u.)	$e'_q = v \times \cos \delta + x'_d \times i_s \times \sin(\delta + \theta) = 1.4048$ (p.u.)
$T'_{do} = 0.1310$ (s)	

TABLE 6. Generator B [21]

Real values	Steady state calculations
$S = 835 \times 1000000$ (VA)	$\theta = \cos^{-1}(PF) = 0.4510$ (p.u.)
$v_{LL} = 26000$ (v)	$v_{as} = 1$ (p.u.)
$v_s = \frac{v_{LL}}{\sqrt{3}}$ (v)	$i_{as} = \cos \theta - (\sin \theta \times I) i = 0.90 - 0.435 i$ (p.u.)
$PF = 0.9$	$E_a = v_{as} + (r_s \times x_q \times I) i \times i_{as} = 1.787 + 1.618 i$ (p.u.)
$\omega_b = 377$ (rad/s)	$\delta = \angle(E_a) = 0.7359$ (p.u.)
$P = 2$	$i_{ds} = -\sin(-\theta - \delta) = 0.9272$ (p.u.)
$H = 5.6$ (s)	$i_{qs} = \cos(-\theta - \delta) = 0.3745$ (p.u.)
$x_e = 0.2$ (p.u.)	$x_{md} = x_d + x_{ls} = 1.61$ (p.u.)
$r_s = 0.003$ (p.u.)	$x_{mq} = x_q + x_{ls} = 1.61$ (p.u.)
$x_{ls} = 0.19$ (p.u.)	$E_{fd} = E_a + (x_d - x_q) \times i_{ds} = 2.4114$ (p.u.)
$x_d = 1.8$ (p.u.)	$i_{fd} = \frac{E_{fd}}{x_{md}} = 1.4977$ (p.u.)
$r_{fd} = 0.000929$ (p.u.)	$i_{kq} = i_{kd} = i_{os} = 0$ (p.u.)
$x_{lfd} = 0.1414$ (p.u.)	$\phi_{qs} = -x_q \times i_{qs} + x_{mq} \times i_{kq} = -0.6741$ (p.u.)
$r_{kd} = 0.01334$ (p.u.)	$\phi_{ds} = -x_d \times i_{ds} + x_{md} \times i_{kd} + x_{md} \times i_{fd} = 0.7423$ (p.u.)
$x_{lkd} = 0.08125$ (p.u.)	$\phi_{kq} = x_{lkq} \times i_{kq} + x_{mq} \times (-i_{qs} + i_{kq}) = -0.6029$ (p.u.)
$x_q = 1.8$ (p.u.)	$\phi_{kd} = x_{lkd} \times i_{kd} + x_{md} \times (-i_{ds} + i_{fd} + i_{kd}) = 0.9185$ (p.u.)
$r_{kq} = 0.00178$ (p.u.)	$\phi_{fd} = x_{lfd} \times i_{fd} + x_{md} \times (-i_{ds} + i_{fd} + i_{kd}) = 0.1303$ (p.u.)
$x_{lkq} = 0.8125$ (p.u.)	$\omega_r = 1$ (p.u.)
	$T_m = \phi_{ds} \times i_{qs} - \phi_{qs} \times i_{ds} = 0.9030$ (p.u.)

COPYRIGHTS

©2024 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 این مقاله، از فیلتر کالمن تفاضل مرکزی (CDKF) برای تخمین پارامترهای دو مدل مختلف ژنراتور سنکرون در شرایط وجود نویز استفاده شده است. قابل ذکر است که مدل های مختلفی از ژنراتور سنکرون با سطوح مختلف دقت برای استفاده در الگوریتم های تخمین وجود دارد. الگوریتم تخمین در این مقاله از تعداد کمتری ورودی جهت اندازه گیری برای تخمین حالت ها و پارامترهای ناشناس برای دو مدل دقیق از ژنراتور سنکرون استفاده می کند. فیلتر کالمن تفاضل مرکزی (CDKF)، از اعضای خانواده فیلترهای کالمن بوده که همانند فیلتر کالمن خنثی (UKF)، از نقاط سیگما برای مدل سازی معادلات غیر خطی استفاده می کند. فیلتر کالمن تفاضلی (CDKF) نتایج بهتری نسبت به فیلتر کالمن خنثی ارائه می دهد. در این تحقیق، با استفاده از دو مدل ژنراتور سنکرون با پارامترهای مختلف در سه سناریوی متفاوت، توانایی فیلتر کالمن تفاضل مرکزی به چالش کشیده می شود، که نشان می دهد این روش، بسیار کارا و قابل اطمینان است.



Leaching Potential and Effectiveness of Pervious Mortar Filters in Bacteria and Turbidity Removal from Surface Water

E. Yogafanny^{a,b}, R. Triatmadja^{*a,c}, F. Nurrochmad^a, I. Supraba^a

^a Department of Civil and Environmental Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia

^b Department of Environmental Engineering, Universitas Pembangunan Nasional Veteran Yogyakarta, Indonesia

^c Center for Environmental Studies, Universitas Gadjah Mada, Yogyakarta, Indonesia

PAPER INFO

Paper history:

Received 24 August 2023

Received in revised form 22 December 2023

Accepted 27 December 2023

Keywords:

Pervious Mortar

Water Filter

Turbidity Removal

Bacteria Removal

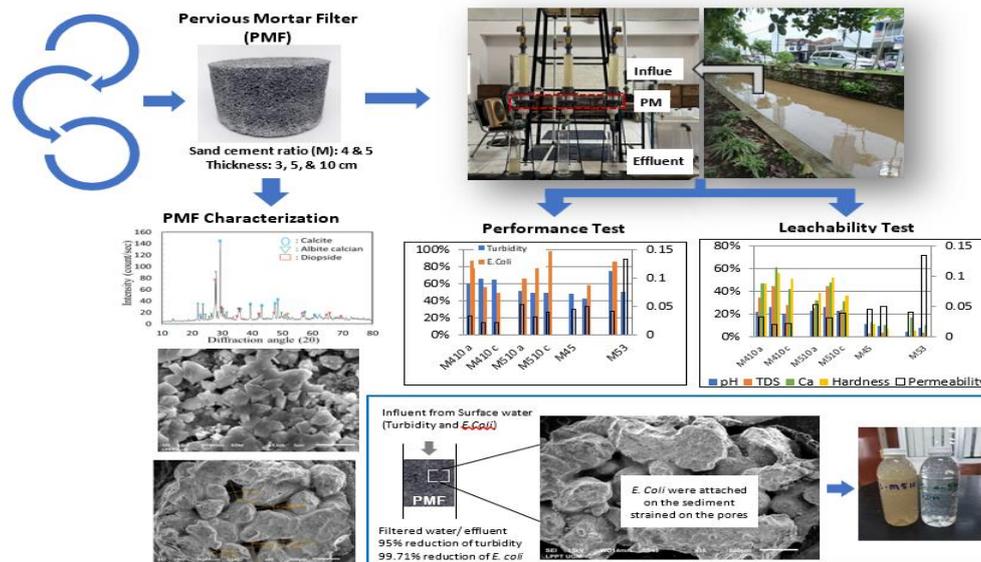
Leaching

ABSTRACT

A pervious mortar filter (PMF) is a modification of pervious mortar and pervious concrete designed as a water filter that, based on its physical characteristics, can reduce turbidity and bacteria. However, chemically, it contains minerals that can dissolve upon contact with water and be found in effluent. This study aimed to determine the performance of PMF in treating surface water by reducing turbidity and *Escherichia coli* and to assess its leaching potential. PMF specimens were created by mixing sand (0.6–0.85 mm), cement, and water with sand-to-cement ratios (M) of 4 and 5 and a water-to-cement ratio (w/c) of 0.4. Each mixture was then molded into pipes with a diameter of 8.2 cm and different thicknesses: 3, 5, and 10 cm. Raw surface water was used for the performance and leaching tests. Results showed that PMF effectively removed 95% turbidity and 99.71% *E. coli*, which increased with the filtration duration. PMF reduced *E. coli* more effectively when designed with a thickness of 10 cm than 5 or 3 cm because it would provide more surface areas for suspended solids and bacteria to attach and be retained. More substantial increases (mean %) of pH, hardness, calcium ions, and TDS were observed in PMF M4 with a thickness of 10 cm than in thinner ones because it contained more cement that would dissolve when in contact with water.

doi: 10.5829/ije.2024.37.07a.05

Graphical Abstract



*Corresponding Author Email: radianta@ugm.ac.id (R. Triatmadja)

Please cite this article as: Yogafanny E, Triatmadja R, Nurrochmad F, Supraba I. Leaching Potential and Effectiveness of Pervious Mortar Filters in Bacteria and Turbidity Removal from Surface Water. International Journal of Engineering, Transactions A: Basics. 2024;37(07):1252-62.

1. INTRODUCTION

Pervious concrete is a composite of cement, coarse aggregates, few to no fine aggregates, water (1), and other admixtures, additives, or organic or recycled materials as substitutes for aggregates (2, 3). Its first use in water purification can be traced back to last decades, when Park and Tia (4) stated about the application of eco-concrete as a water purifier. Then, they observed substantial reductions in total nitrogen and phosphorus concentrations in water when testing a pervious concrete filter that was made of small sand particles and thus had many pores or voids (4). The application of pervious concrete as a water filter was further solidified after being patented by Majersky in 2008 (5). Numerous studies have since confirmed its ability to reduce water pollutants, including BOD, COD, and heavy metals (5). Each pore in pervious concrete is a medium that physically captures dissolved nitrogen and phosphorus in water by filtration and chemically removes dissolved phosphorus by adsorption (6). Pervious concrete has shown promising results as a water purifier and had a potentially significant role in addressing water pollution challenges.

Pervious mortar filter (PMF) is an innovative technology that is currently under development. Unlike previous concrete, PMF does not use gravels, additives, and admixtures to ensure good-quality water treatment. Furthermore, PMF differs from conventional mortar in that it does not solely add cement, fine aggregates, and water to the mixture (7) but applies specific sand-to-cement and water-to-cement ratios. PMF mixes fine sand, cement, and water with a sand-to-cement ratio (M) of 4 or 5, resulting in a smaller pore size and permeability lower than pervious concrete but higher than conventional mortar. Due to the small pores and their hydraulic characteristics, this porous composite can transmit water while effectively entrapping the pollutants carried, such as suspended solids, bacteria, and heavy metals. One method commonly used in surface water treatment especially for turbidity removal is coagulation using alum and ferric chloride along with calcium hydroxide as a coagulant to reduce turbidity (8). PMF can be an alternative method to reduce turbidity without using these chemicals.

Many scholars have explored the application of fine aggregates in mortar/ concrete as water filter. Taghizadeh et al. (9) designed a water filter using only fine sand, later termed porous concrete. In Indonesia, the development of PMF technology is initiated as a water filter by using sand with a grain size of 1–2 mm as the aggregate. Subsequently, comprehensive research by Kamulyan et al. (10) investigated the use of fine aggregates with a particle size of < 2 mm.

The ability of PMF to reduce various impurities has been well-established in the literature. Pervious mortar

proved highly effective in reducing turbidity, with a reported 90–95% reduction (10). PMF with 10 cm diameter and 20 cm thickness created using sand aggregate sized 0.15–0.30 mm and a sand-to-cement ratio of 4 significantly reduced turbidity to <5 NTU and impressively removed *Escherichia coli* by (*E. coli*) 98.71% or 2-log removal value (LRV). Therefore, to contribute to PMF development, the current research assessed the filtration performance of PMF made of small (0.6–0.85 mm) and uniformly sized aggregates with a sand-to-cement ratio of 4 (Figure 1).

The pollutant removal performance of PMF stems from its unique physical and hydraulic characteristics, such as pore structure and geometry. PMF composed of aggregates smaller than 1 mm with the appropriate mix design (sand-to-cement ratio of 4 or 5) has low permeability, prolonging contact with water and the substances it carries. This creates opportunities for effective entrapment and retention of impurities within the filter's pores. However, the extended contact duration also allows for dissolution and subsequent transportation of the dissolved elements, such as Ca^{2+} , into the water. This means pervious mortar or concrete can dissolve or leach if exposed to specific water qualities for an extended period (11). Leaching is a significant degradation observed in cement-based composites, representing one of the primary factors responsible for their mechanical alteration (12–16). This process starts when cement paste comes into contact with pure or acidic water (low pH), causing hydrolysis (17). Conversely, dissolution involves transporting the concrete's compounds or ions due to the reaction between water and portlandite ($\text{Ca}(\text{OH})_2$). Portlandite is a mineral formed during the curing of cement by hydration (6, 17). Leaching evolves through the dissolution and precipitation of minerals in cement-based composites, such as PMF. The chemical reactions are expressed in Equations 1 to 4.

Formation of carbonic acid:

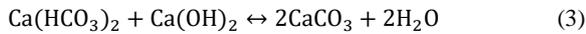


Figure 1. Pervious mortar filter with the dimension of 8.2 cm (d, diameter) x 5 cm (h, or thickness), made of sand with a grain size of 0.6–0.85 mm and a sand-to-cement ratio (M) of 4

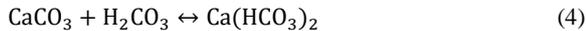
Dissolution of calcium hydroxide:



Formation of calcium carbonate:



Dissolution of calcium carbonate:



Numerous studies have investigated the interplay between the leaching of calcium ions (Ca^{2+}) and its impact on concrete's porosity and tensile strength, and vice versa. Exploring the effect of porosity in Portland cement on the leaching process, Haga et al. (18) found that increased pore volume led to faster dissolution of portlandite present in the cement-based composite. They concluded that diffusion regulates the transport of dissolved substances or constituents and that the primary components dissolved in hardened ordinary Portland cement (OPC) are portlandite and C-S-H gel. Solpuker et al. (19) estimated the leaching potential and the ability of pervious concrete to retain trace metals. Using the column method, they passed water with a pH of 4.3 through the pervious concrete, but the generated wastewater was basic (approximately pH of ~10). The conductivity decreased rapidly within the first 50 hours and then slowly afterward. In the early stages, there was substantial leaching of trace metals, which diminished significantly after 50 hours and then gradually declined over time.

Various researchers have extensively studied leaching in concrete or pervious concrete, primarily with immersion tests (14, 18, 19). The column method has received less attention by comparison. In this research, the column method was first adjusted to align with the function and utilization of pervious mortar as a water filter and then employed to assess the leaching potential of composites with smaller aggregate sizes, specifically pervious mortar. In this method, water flowed dynamically into the pervious mortar until saturation. Based on the above description, this study aimed to investigate the effectiveness of pervious mortar filters in treating surface water by removing impurities, i.e., suspended solids and bacteria, and their leaching potential.

2. MATERIALS AND METHODS

2.1. Preparation of Specimens

A pervious mortar filter (PMF) comprises sand, cement, and water without admixtures. The sand used was collected from the Progo River (Indonesia), with a particle size ranging from 0.6 to 0.85 mm, while the cement used was Portland composite cement (PCC) "Semen Gresik" (Indonesia). No admixtures or additives were incorporated into the PMF mixture to prevent the risk of heavy metal

dissolution (19). The physical and chemical characteristics of the sand and cement are presented in Table 1. Specimens were made with two sand-to-cement ratios (M) of 4 and 5 and a water-to-cement ratio of 0.4, as shown in Table 2.

The mixing and stirring stages were modified from the previous concrete production process in Park and Tia (4), which included the use of a Hobart mortar mixer and the 2-minute mixing time of the dry matter (instead of 1 minute). These modifications were necessary because the current research used sand with a much smaller grain size, thus creating a larger surface area. First, half of the sand and half of the cement were added into the mixer and stirred for 2 minutes. Subsequently, the remaining halves of both materials were added to the mix and stirred again for 2 minutes until homogenous. Then, water was gradually added while stirring for 2 minutes to create the final PMF mixture.

The next stage was molding on a vibrating table. This method was modified and informed by preliminary experiments conducted before commencing this study. Approximately 200 g of the final mixture was carefully poured into sand-layered PVC molds, made by cutting 3-inch PVC pipes into varying heights or thicknesses: 3, 5, and 10 cm. A stainless-steel ballast weighing 225 g was positioned on top of the mixture on the vibrating table and then vibrated at 50 Hz for 1 minute. This step was repeated until the mixture slightly exceeded the mold's rim (approximately three fillings to create a specimen with 5 cm thickness), which was later leveled off using a ruler. Afterward, the filled mold was inverted, weighed, and vibrated again for 1 minute. Finally, it was returned to its initial position and vibrated for 30 seconds. The specimens were left to dry for 24 hours at room temperature (approximately 26°C), followed by curing with a moist cloth cover for 90 days. Afterward, the PMF was ready for use as a filter.

All the 60 specimens were tested for porosity. Results show that porosity remained consistent across the different thicknesses and M variations. For the performance and leachability tests, several specimens were chosen according to its porosity. By this consideration, three specimens of 10 cm thickness were selected from M410 (namely M410a, M410b, and M410c) or M510 (namely M510a, M510b, and M510c). One specimen that represent the mean porosity on PMF with 5 and 3 cm thickness for both M ratio were chosen to be tested for performance and leachability test. According to the preliminary study, the thicker PMFs (10 cm) perform better than the thinner ones. Hence, for this performance and leachability test, the PMF with 10 cm thickness is evaluated in more detail by using three samples on each M. While for the 5 and 3 cm PMF, only one sample is used to be observed. In this experiment, ten specimens of PMF were chosen i.e. M410a, M410b, M410c, M510a, M510b, M510c, M45, M55, M43, and M53.

TABLE 1. Physical and chemical characteristics of the sand and cement inputted into the pervious mortar filter (PMF) mix design

Sand (Progo River)		Cement (PCC)	
Element	Value	Element	Value
SiO ₂ (mass%)	39.10	SiO ₂ (mass%)	18.52
P ₂ O ₅ (mass%)	3.768	P ₂ O ₅ (mass%)	6.65
SO ₃ (mass%)	0.509	SO ₃ (mass%)	2.051
CaO (mass%)	9.958	CaO (mass%)	67.53
TiO ₂ (mass%)	4.351	TiO ₂ (mass%)	0.633
MnO (mass%)	0.532	MnO (mass%)	0.121
Fe ₂ O ₃ (mass%)	31.70	Fe ₂ O ₃ (mass%)	4.23
CuO (mg/kg)	654.8	CuO (mg/kg)	513
ZnO (mg/kg)	660	ZnO (mg/kg)	359
Rb ₂ O (mg/kg)	187.6	Rb ₂ O (mg/kg)	73.7
SrO (mass%)	0.199	SrO (mg/kg)	519
BaO (mass%)	0.102	BaO (mg/kg)	382
Al ₂ O ₃ (mass%)	7.65	As ₂ O ₃ (mg/kg)	75.9
K ₂ O (mass%)	1.981	NiO (mg/kg)	684
Density (kg/m ³)	2,840	Density (kg/m ³)	2,960
Fineness modulus	2.7		
Uniformity coef.	1.6		

TABLE 2. The mixture design of the PMF specimens

Criteria	Value and unit
Size of sand	0.6–0.85 mm
Type of cement	Portland Composite Cement (PCC), Semen Gresik (Indonesia)
Sand-to-cement ratio (M)	4 and 5
Water-to-cement ratio	0.4
Specimen's diameter	8.2 cm (3-inch PVC pipe)
Specimen's thickness	3, 5, and 10 cm
Curing type	Moist cloth cover
Curing duration	90 days

2. 2. Performance and Leaching Tests Figure 2 presents the laboratory equipment used in the performance and leaching tests. The performance or effectiveness test of the PMF specimens was assessed based on changes in the turbidity and *E. coli* concentration after filtration. The influent was water samples collected every morning during the experiment (January 8 to March 8, 2023) from the drainage channel “Selokan Mataram” (north of the Faculty of Forestry, Gadjah Mada University (UGM), Indonesia). Samples were tested for turbidity at the Laboratory of Sanitary and

**Figure 2.** Equipment for the performance and leaching tests of PMF specimens

Environmental Engineering (UGM) and *E. coli* presence at the Faculty of Geography (UGM). Meanwhile, the leaching test was specifically modified to the function of PMF as a water filter. The modifications made to the several steps of the leaching test performed by Vadas et al. (12) were as follows:

a. Column filtration

Mineral dissolution occurs when water flows into and directly interacts with PMF. The leaching potential of pervious mortar, acting as a filter, was analyzed by comparing the water quality before (influent) and after (effluent) passing through 100-day-old PMF specimens designed in this study. The tested water quality parameters were pH, TDS, Ca²⁺, and total hardness (CaCO₃), which are indicators of mineral leaching from the filters during water treatment.

b. X-ray diffraction (XRD) and scanning electron microscope-energy dispersive X-ray (SEM-EDX)

XRD was employed to determine the minerals formed during the curing of pervious mortar. The minerals present in PMF were used as a basis for analyzing the chemical elements or compounds found in the effluent. XRD was conducted on one of the test specimens (M4 and M5) and cement that had been allowed to set for 90 days. This test was performed at the UGM-Integrated Research and Testing Institute (LPPT). In addition, SEM-EDX was employed to determine the morphology of the PMF composites to help visually identify their constituent minerals and pore sizes.

2. 3. Water Quality Test Equipment

a. Turbidity

The equipment used to measure turbidity is HACH 2100Q Turbidimeter. Turbidity is read by the unit of Nephelometric Turbidity Unit (NTU). A turbidimeter is a tool for turbidity testing with optical properties due to light dispersion and can be expressed as a ratio of reflected light to incident light. The intensity of light reflected by a solid suspension is a function of concentration if other conditions are constant.

b. Total Dissolved Solid (TDS) and pH

The equipment used to measure these two parameters are a portable multiparameter - HACH sensION 65 with the pH electrode – HACH, and conductivity/TDS electrode – HACH

c. Ca^{2+} and total hardness (CaCO_3)

Water hardness is analyzed using the complexometric method, the principle of which is based on the formation of soluble complex compounds between metal ions and complex-forming substances, namely the formation of Ca with EDTA.

The Sodium salt ethylene diamine tetra acetate (EDTA) will react with certain metal cations to form soluble chelate complex compounds. At pH 10.0 + 0.1, the calcium and magnesium ions in the test sample will react with the Eriochrome Black T (EBT) indicator and form a purplish red solution. If Na_2EDTA is added as a titrant, the calcium and magnesium ions will form a complex compound, the indicator molecules are released again, and at the end point of the titration the solution will change color from purplish red to blue. From this method, total hardness (Ca and Mg) can be obtained.

Calcium can be determined directly with EDTA, if the pH of the test sample is made high enough (12-13), so that magnesium will precipitate as magnesium hydroxide and at the end point of the titration the Eriochrome Black T (EBT) indicator will only react with calcium to form a blue solution. From this method the calcium (Ca) concentration in the water can be obtained.

d. *E.coli*

This research uses the Most Probable Number (MPN) method which consists of an presumptive test, confirmation test, and refinement test, using solution concentrations of 0.1 ml, 1 ml and 10 ml.

The initial step of the experiment, namely a presumptive test, was to sterilize the equipment using an autoclave at a temperature of 121°C for 30 minutes. After that, the media was prepared, the made media were Lactose Broth (LB), Brilliant Green Lactose Bilebroth (BGLB), and Eosin Methylene Blue (EMB).

Next is the confirmation test. At this stage, 9 culture tubes containing sterile LB media equipped with Durham tubes were then poured into the water samples with a dropper in different volumes, namely 10 ml, 1 ml and 0.1 ml in each of the 3 test tubes, then incubated for 24 hours at 37 °C. In the confirmation test, each culture tube containing 10 ml of Brilliant Green Lactose Bilebroth (BGLB) equipped with a Durham tube is prepared, positive samples are added. Pour water into the 1 ml lactose culture that has been incubated and is considered positive, the tube is incubated for 24 hours at 45°C.

The final stage is the refinement test, samples that are positive in the confirmation test are inoculated using a loop needle onto the surface of Eosin Methylene Blue (EMB) media in a zig-zag manner and then incubated at a temperature of 37°C for 24 hours. Colony growth was observed on Eosin 63 Methylene Blue (EMB) media. Colonies that show a metallic sheen are colonies of

Escherichia coli bacteria. After all tests are completed, the *Escherichia coli* MPN value is determined by matching the analysis results with the MPN table (20).

3. RESULTS AND DISCUSSION

3. 1. Characteristics of PMF Specimens

The same M value represents an identical composition of cement and sand. PMFs with M4 (PMF-M4) and PMF-M5 were characterized and analyzed morphologically and chemically using XRD and SEM-EDX tests. Figure 3 shows that PMF-M4 consisted of three major minerals: calcite, albite calcian, and diopside. Further details on their morphology and visual appearance were obtained from the SEM-EDX test, as presented in Figures 5 and 6.

Figure 6 also shows the presence of capillary and gel pores. All data obtained from the XRD and SEM-EDX tests provided the basis for analyzing the leaching potential of PMF when used as a filter.

The X-ray diffractograms in Figure 7 identified calcite, albite calcian, and diopside as the mineral constituent of PMF-M5. Figures 8 and 9 provide further information on their mineral morphology and the distribution of capillary and gel pores.

3. 2. Performance of PMF Specimens

Pervious concretes act as an effective water filter by retaining suspended solids and adsorbing metals and other chemicals present in water (5). Similarly, porous mortars possess comparable physical and hydraulic



Figure 3. Examples of PMF specimens during the production process

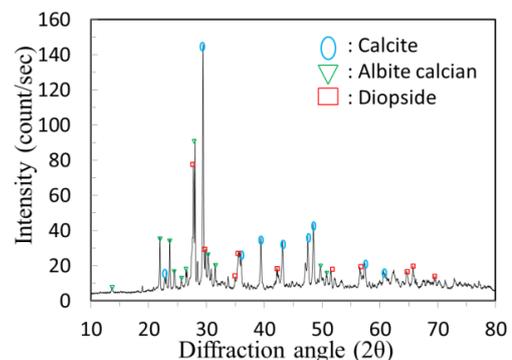


Figure 4. X-ray diffractograms showing the mineral composition of PMFs with M4 (PMF-M4)

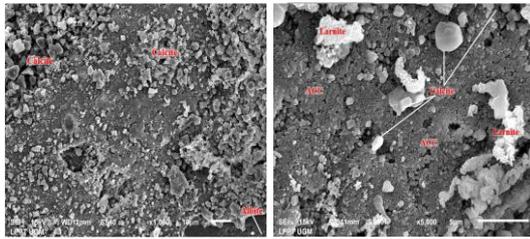


Figure 5. SEM of PMF-M4 at 1000x (left) and 5000x magnifications (right)

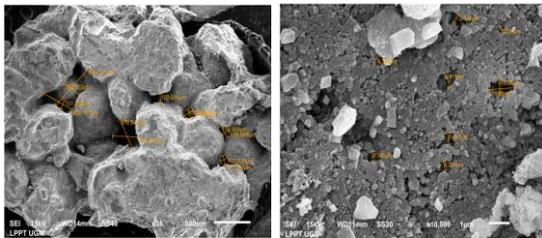


Figure 6. Identification of capillary pores (left) and gel pores (right) in PMF-M4

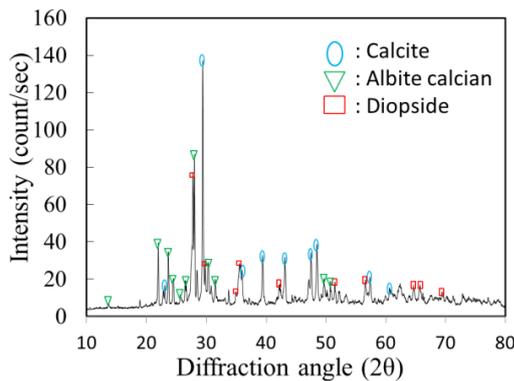


Figure 7. X-ray diffractograms showing the mineral composition of PMF-M5

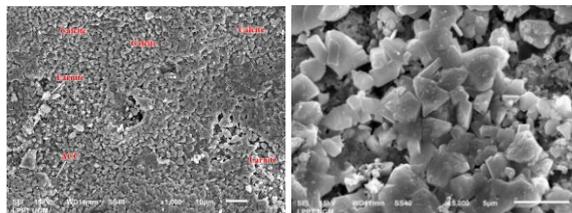


Figure 8. SEM of PMF-M5 at 1000x (left) and 5000x magnifications (right)

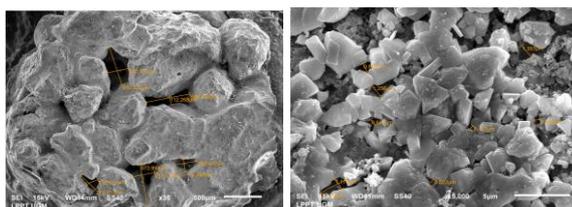


Figure 9. Identification of capillary pores (left) and gel pores (right) in PMF-M5

characteristics, enabling them to filter water effectively. Previous studies have demonstrated the ability of pervious mortars to entrap and remove suspended particles (5, 9), heavy metals, such as Cu, Pb, Cd, and Zn (12), and bacteria from raw water, resulting in better-quality effluents. In the mortar’s composition, sand and cement form pores that facilitate the storage and retention of water impurities. This study analyzed the performance of PMFs using two water quality parameters: turbidity and *E. coli* bacteria. Further, to complement previous research findings, it observed the time taken to perform filtration until saturation. The performance test equipment, process, and output of filtration of water samples collected from the drainage channel “Selokan Mataram” with the PMF specimens are shown in Figure 10.

Based on the performance test shown in Table 3, the highest to the lowest maximum effectiveness of PMF in reducing turbidity was 95% (10 cm/M5), 93% (10 cm/M4), 78% (3 cm/M4), 77% (5 cm/M4), 74% (3 cm/M5), and 66% (5 cm/M5). PMF also proved effective in removing *E. coli* by up to 99% (10 cm/M4 and M5), 58% (5 cm/M5), and 86% (3 cm/M4). Table 4 shows the influent and effluent turbidity and *E. coli* concentrations for all PMF specimens.

Figure 11 demonstrates the relationship between the permeability and the performance of PMF specimens in lowering turbidity and *E. coli*. Permeability was measured using the constant head method with the pressure level set at 30 cm. The graph generally indicates that for PMFs with the same thickness, lower permeability coincided with greater turbidity reduction and, thus, higher effectiveness. However, this trend did not apply to performance in reducing *E. coli*.

Among the specimens, M510a produced the highest effectiveness with the most significant turbidity reduction of up to 95%, although a fairly wide range of

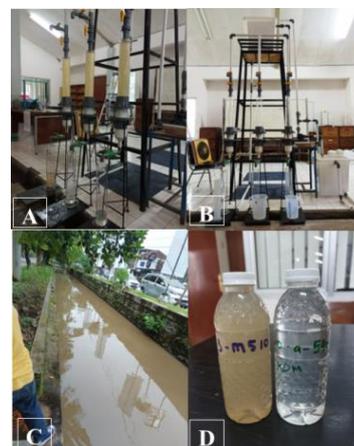


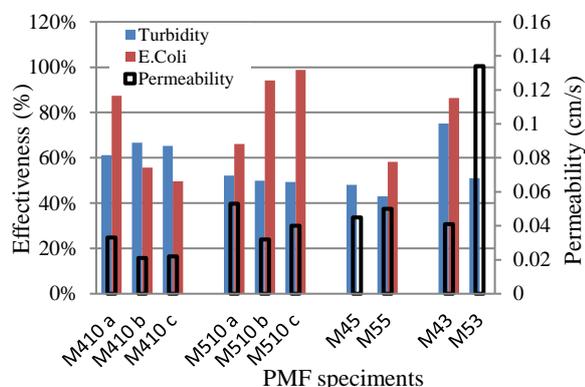
Figure 10. Process (A) and equipment setup of the PMF performance test (B), water in the drainage channel “Selokan Mataram” as influent (C), and water samples before and after filtration with PMF (D)

TABLE 3. Permeability and effectiveness of PMF specimens in reducing turbidity and *E. coli*

PMF	Turbidity (%)		<i>E. coli</i> (%)		Permeability (cm/s)	Vol. of treated water (L)
	min	max	min	max		
M410a	33	93	0	93.75	0.033	83.9
M410b	42	79	0	86.36	0.021	44.83
M410c	53	84	0	99.36	0.022	98.27
M510a	8	95	0	99.71	0.053	86.6
M510b	20	80	91.25	97.08	0.032	46.56
M510c	6	92	98.83	98.75	0.040	70.13
M45	34	77	0	0	0.045	133.36
M55	28	66	0	58.18	0.050	229.17
M43	73	78	0	86.36	0.041	31.05
M53	28	74	0	0	0.134	75.89

TABLE 4. Turbidity and *E. coli* concentration in the influent (before) and effluent (after filtration)

PMF	Filtration time (hour)	Turbidity (NTU)		<i>E. coli</i> (MPN/100ml)	
		Inf	Eff	Inf	Eff
		M410a	7.5	184	8.26
M410b	5.5	184	23.9	2400	150
M410c	8.5	173	29.7	28	7
M510a	9	156	13.8	2400	7
M510b	4.4	156	49.2	2400	7
M510c	6.9	156	19.1	2400	3
M45	17.5	121	20.8	1100	1100
M55	21	121	18.4	1100	1100
M43	5.25	182	72.1	460	150
M53	5.75	323	73	1100	1100

**Figure 11.** Permeability and effectiveness of PMF specimens in removing turbidity and *E. coli*

effectiveness was observed (minimum effectiveness of 8%). For comparison, M43 exhibited moderate effectiveness by lowering turbidity levels by 73–78%, similar to Hosseini & Togholi (21), who reported that pervious concrete successfully reduced total suspended solids by 75 to 85%. Performance variations in M510a indicate that PMFs with 10 cm thickness reduced turbidity to a varying degree. This occurred because the filter required time to accumulate suspended particles on its pore walls, contributing to increasing turbidity reduction over time. The longer the PMF operates, the higher its effectiveness in lowering turbidity before reaching saturation.

PMF-M4 showed a higher average of turbidity reduction than PMF-M5 of the same thickness (5 or 3 cm). This could be attributed to the larger proportion of cement in the M4 mix that created a less porous filter with a better ability to retain suspended particles. The maximum turbidity reduction of PMF-M4 and PMF-M5 with 10 cm thickness varied from 79% to 95%, with no noticeable difference between both sand-to-cement ratios.

PMFs with thicknesses of 5 and 3 cm showed a maximum turbidity reduction of 66% to 78%, although the latter performed better than the former. This might stem from its smaller pore area, which enabled suspended solids to fill the pores almost immediately and hold onto other suspended particles, increasing effectiveness. However, due to the limited pore area, this PMF became clogged or saturated more quickly, thus transmitting and treating less water than the one with a thickness of 5 cm (see Table 3).

The biological parameter observed to estimate the amount of bacteria in the water was *E. coli*, a species found in the environment, food, and the intestines of humans and animals. While most *E. coli* are not harmful, a small percentage of certain kinds of this microorganism can result in pneumonia, urinary tract infections, and diarrhea. Therefore, detecting their presence in a clean water supply is essential. Based on the analysis results presented in Table 3, the PMF specimens could remove *E. coli* from the influent up to 99.71%. The zero percentage indicates no removal of *E. coli*, possibly because the filters were relatively new. In filters that have been operated for a long time, many bacteria are attached and growing on the walls of their pores and can later bind *E. coli*. In other words, with more frequent use, the PMF improves its effectiveness in removing *E. coli* from water.

No removal of *E. coli* concentration was primarily observed in PMFs with a thickness of 5 cm. As seen in Table 3, they processed or filtered the highest amount of treated water compared to PMFs with thicknesses of 10 and 3 cm. These results suggest that pore condition or effective porosity significantly affects how much water PMFs can filter. Nonetheless, more information than the total porosity presented in the table is required to analyze

the effect of porosity on the filter's ability to treat water. Also, further studies are needed to identify the size and number of open and closed pores to support water filtration analysis with PMF.

Based on its minimum and maximum values, M510 consistently demonstrated good effectiveness in removing about 97.08 to 99.71% of *E. coli*. These results did not substantially differ from the maximum effectiveness of M410, which varied from 86.36 to 99.36%. In conclusion, the sand-to-cement ratio (M) does not directly affect the filter's ability to retain bacteria. Instead, thickness is believed to have a more significant influence. Generally, 10-cm thick PMFs exhibited better *E. coli* reduction than those designed with thicknesses of 5 and 3 cm. Thicker PMFs provide more space to entrap and retain suspended particles and bacteria. These results found that thicker PMFs had higher effectiveness in removing *E. coli*, specifically those with a thickness of 15 cm and a sand-to-cement ratio of 4

3. 3. Leaching Potential of PMF Specimens

Water can be filtered effectively by pervious concrete and pervious mortar. However, because of the chemical composition of their basic materials, there is a chance of disintegration when exposed to water. Disintegration occurs when the calcium ions in the cement detach and flow together with water into the filter's surroundings, hence called calcium leaching (22, 23). Calcium leaching potentially affects the strength and durability of all cement-based composites (24). Calcium contents, total hardness (CaCO_3), total dissolved solids (TDS), and pH levels of the effluent may increase due to water contact with the chemical composition of pervious concrete or pervious mortar filters during filtration (5).

This research investigated how sand-to-cement ratio (M) and thickness affected the effluent's pH, hardness, TDS, and calcium levels. In doing so, it adds more evidence of the leaching potential to earlier studies on PMF, as demonstrated by changes in these parameter values. The leaching test aimed to determine the leaching patterns of PMF when exposed to the inflow of natural water sources. Therefore, it used the same influent as the performance test.

Table 5 shows the mean percentage of increase in total hardness, Ca^{2+} , TDS, and pH. It indicates a correlation between smaller sand-to-cement ratios and higher increases in Ca^{2+} and hardness levels. The percentage of increase for each parameter was generally higher in thicker (10 cm) than in thinner PMF specimens (3 cm), indicating a higher water alkalinity level. Lee et al. (25) found that pH increased more substantially in thicker (20 cm) than thinner pervious concrete (10 cm) because pervious concrete contains hydrated cement products that easily dissolve in water, releasing substances like OH^- that can increase the water pH. Thicker pervious concrete has more cement, resulting in

the dissolution of more products of cement hydration reactions that leads to effluent with increased pH. The current study shares some similarities with Lee et al. (25), as it found pH, hardness, Ca^{2+} , and TDS generally increased more significantly in PMFs with a thickness of 10 cm than 3 or 5 cm.

Each PMF has its unique capacity to drain water, thus creating different durations of filtration, times to saturation, and filtered water amounts. For all the specimens, the pH levels of the influent ranged from 6.7 to 7.5. Meanwhile, the pH levels of the effluent were as follows: 8.5–10.5 for M410, 8.5–10.6 for M510, 7.5–8.5 for M45, 7.1–9 for M55, 7.3–7.4 for M43, and 7.2–8.5 for M53. The TDS levels of the influent ranged from 87 to 119.5 mg/l and the TDS levels of the effluent were as follows: 100.9–565 mg/l for M410, 96.3–582 mg/l for M510, 109.3–119.5 mg/l for M45, 109.4–133.6 mg/l for M55, 103.2–106.1 mg/l for M43, and 94.4–114.1 mg/l for M53. The Ca^{2+} levels of the influent ranged from 15.84 to 25.6 mg/l and the Ca^{2+} levels of the effluent were as follows: 22.97–89.09 mg/l for M410, 28.67–106.5 mg/l for M510, 25.6–32.77 mg/l for M45, 24.58–40.96 mg/l for M55, 22.53–27.14 mg/l for M43, and 17.41–29.7 mg/l for M53. The total hardness (CaCO_3) levels of the influent ranged from 27.72 to 69.3 mg/l and the total hardness levels of the effluent were as follows: 43.56–387.52 mg/l for M410, 51.2–230.4 mg/l for M510, 46.08–61.44 mg/l for M45, 43.52–66.56 mg/l for M55, 61.44–76.8 mg/l for M43, and 55.04–70.4 mg/l for M53.

Increased pH levels in the effluent are consistent with the outcome of water treatment using pervious concrete (12, 25, 26). Vadas et al. (12) stated that pervious concrete could entrap suspended particulates, which later enhanced the retention of heavy metals, such as Cu, Zn, Cd, and Pb, effectively reducing their concentrations by up to 95%. This process was followed by an increase in

TABLE 5. Mean percentages of increase in pH, TDS, Ca^{2+} , and total hardness (CaCO_3), after filtration using all the PMF specimens with their varying porosity characteristics

PMF	Average of increment percentage (%)			
	pH	TDS	Ca^{2+}	Hardness
M410a	22%	35%	47%	47%
M410b	26%	45%	61%	56%
M410c	20%	28%	42%	51%
M510a	23%	25%	32%	39%
M510b	26%	44%	48%	52%
M510c	23%	24%	32%	36%
M45	12%	3%	13%	11%
M55	10%	4%	10%	7%
M43	5%	2%	17%	5%
M53	8%	4%	10%	0%

pH levels from 7 (influent) to a maximum of 7.8 (effluent) (12). Lee et al. (25) also observed an increase in pH from 2 in the influent to 7 in the effluent after filtration with pervious concrete.

3. 4. Analysis of Variance (ANOVA) ANOVA was used in this research to compare variances across the means (or average) of different groups (ratio M and thickness). Tables 6 and 7 show the results of analysis of variance (ANOVA) with SPSS for performance of PMF and leachability of PMF.

TABLE 6. ANOVA on performance of PMF

Turbidity removal				
Variables	Mean	*Significant value		
		Sig.	No Sig.	
Ratio S/C (M)	4	61.46%		0.163
	5	47.97%		
	3	63.08%		
Thickness of PMF (cm)	5	45.24%	0.390	
	10	58.37%		
<i>E.coli</i> removal				
Variables	Mean	*Significant value		
		Sig.	No Sig.	
Ratio S/C (M)	4	37.68%		0.648
	5	45.90%		
	3	21.59%		
Thickness of PMF (cm)	5	6.46%	0.002	
	10	68.38%		

*Significant if the significant value is < 0.05

TABLE 7. ANOVA on leachability of PMF

pH increment				
Variables	Mean	*Significant value		
		Sig.	No Sig.	
Ratio S/C (M)	4	18.04%		0.722
	5	16.66%		
	3	6.53%		
Thickness of PMF (cm)	5	10.55%	0.000	
	10	23.46%		
TDS increment				
Variables	Mean	*Significant value		
		Sig.	No Sig.	
Ratio S/C (M)	4	22.67%		0.609
	5	17.22%		

Thickness of PMF (cm)	3	2.81%	0.011
	5	3.95%	
	10	33.39%	
Ca ²⁺ increment			
Variables	Mean	*Significant value	
		Sig.	No Sig.
Ratio S/C (M)	4	35.76%	0.179
	5	23.48%	
	3	13.83%	
Thickness of PMF (cm)	5	11.61%	0.001
	10	44.09%	
Hardness increment			
Variables	Mean	*Significant value	
		Sig.	No Sig.
Ratio S/C (M)	4	34.69%	0.302
	5	23.48%	
	3	2.27%	
Thickness of PMF (cm)	5	9.19%	0.000
	10	47.33%	

*Significant if the significant value is < 0.05

Based on the results of the ANOVA significance value, there was a significant difference in the average reduction in *E. coli* as well as an increase in pH, TDS, Ca, and hardness based on PMF thickness. However, there is no significant difference in the average values of all variables based on the sand-cement ratio (M).

The significant difference in variable mean values based on PMF thickness is in accordance with findings in other studies that filter thickness greatly influences *E. coli* removal and leachability of PMF. The thicker the pervious concrete, the higher the pH of the effluent in pervious concrete filter (25, 26).

4. CONCLUSION

The effectiveness of pervious mortar filters (PMFs) in reducing turbidity and *E. coli* is caused by the physical filtration process. The filter needs to first entrap some suspended solids from water and allow them to accumulate on its pore walls. This buildup further captures suspended particles from subsequent influents, increasing effectiveness in reducing turbidity. Among PMFs with the same thickness, those with M4 have higher effectiveness than M5. This can be attributed to the higher presence of cement as an adhesive material in M4, which makes the filters less porous and enhances the retention of suspended particles as well as *E. coli*. In addition, thicker PMFs (10 cm) generally have higher

effectiveness in removing *E. coli* than the thinner ones (5 and 3 cm) because they provide more surface area and, thus, more chances for suspended solids and bacteria to attach and be retained.

The leaching potential of PMFs can be seen from changes in calcium ions, pH, total dissolved solids (TDS), and hardness between the influent and the effluent. The effluent shows substantially higher parameter values after filtration with the 10-cm thick PMFs than with the thinner ones, suggesting more leaching in the former than the latter. In addition, among PMFs with the same thickness, a smaller sand-to-cement ratio (M4) produces a higher percentage of increase in pH, TDS, Ca^{2+} , and hardness than M5, especially Ca^{2+} and hardness. Calcium ions are released from reactions with water in the form of calcium carbonate. The trend of water hardness in the effluent mirrors that of Ca^{2+} .

5. ACKNOWLEDGMENT

This work received funding from the Indonesia Endowment Fund for Education awarded to the first author (Grant no: KET 248/LPDP.4/2021). The authors would gratefully acknowledge the assistance and support provided by the Hydraulics Laboratory of the Department of Civil and Environmental Engineering, Universitas Gadjah Mada.

6. REFERENCES

- Pachideh G, Gholhaki M, Moshtagh A. Performance of porous pavement containing different types of pozzolans. *International Journal of Engineering, Transactions C: Aspects.* 2019;32(9):1277-83. <https://doi.org/10.5829/ije.2019.32.09c.07>
- Shafabakhsh G, Ahmadi S. Evaluation of coal waste ash and rice husk ash on properties of pervious concrete pavement. *International Journal of Engineering, Transactions B: Applications.* 2016;29(2):192-201. <https://doi.org/10.5829/idosi.ije.2016.29.02b.08>
- Shantveerayya K, Kumar CL M, Shwetha K, Jima F, Fufa K. Performance Evaluation of Hollow Concrete Blocks Made with Sawdust Replacement of Sand: Case Study of Adama, Ethiopia. *International Journal of Engineering, Transactions C: Aspects.* 2022;35(6):1119-26. <https://doi.org/10.5829/ije.2022.35.06c.03>
- Park S-B, Tia M. An experimental study on the water-purification properties of porous concrete. *Cement and concrete research.* 2004;34(2):177-84. [https://doi.org/10.1016/S0008-8846\(03\)00223-0](https://doi.org/10.1016/S0008-8846(03)00223-0)
- Majersky G. Concrete filtering systems and methods. Google Patents; 2008.
- Kim G, Jang J, Khalid HR, Lee H-K. Water purification characteristics of pervious concrete fabricated with CSA cement and bottom ash aggregates. *Construction and Building Materials.* 2017;136:1-8. <https://doi.org/10.1016/j.conbuildmat.2017.01.020>
- Jebadurai S, Tensing D, Hemalatha G, Siva R. Experimental investigation of toughness enhancement in cement mortar. *International Journal of Engineering, Transactions B: Applications.* 2018;31(11):1824-9. <https://doi.org/10.5829/ije.2018.31.11b.04>
- Davarpanah L, Abdollahzadeh Sharghi E. Performance of Continuous Electrocoagulation Process for Turbidity Removal from Sand Filter Backwash Water. *International Journal of Engineering, Transactions C: Aspects.* 2019;32(12):1696-703. <https://doi.org/10.5829/ije.2019.32.12c.01>
- Taghizadeh M, Torabian A, Borghei M, Hassani A. Feasibility study of water purification using vertical porous concrete filter. *International Journal of Environmental Science & Technology.* 2007;4:505-12. <https://doi.org/10.1007/BF03325987>
- Kamulyan B, Nurrochmad F, Triatmadja R, Sunjoto S, editors. Capacity of concrete sand filter to treat high turbid water. *Int Conf Sustain Dev Water Wastewater Treat*; 2009.
- Hartwich P, Vollpracht A. Influence of leachate composition on the leaching behaviour of concrete. *Cement and Concrete Research.* 2017;100:423-34. <https://doi.org/10.1016/j.cemconres.2017.07.002>
- Vadas TM, Smith M, Luan H. Leaching and retention of dissolved metals in particulate loaded pervious concrete columns. *Journal of Environmental Management.* 2017;190:1-8. <https://doi.org/10.1016/j.jenvman.2016.12.047>
- Ueno T, Yanaka A, Okazaki S, Matsumoto N, Yoshida H. Physical Property and Heavy Metal Leaching Behavior of Concrete Mixed With Woody Ash. *GEOMATE Journal.* 2023;24(102):77-84. <https://doi.org/10.21660/2023.102.s8530>
- Liang T, Zhou J, Wu Q. Experimental investigation on leaching behavior of ultra-high performance concrete submitted to a flow environment. *Construction and Building Materials.* 2023;372:130843. <https://doi.org/10.1016/j.conbuildmat.2023.130843>
- Overmann S, Lin X, Vollpracht A. Investigations on the leaching behavior of fresh concrete—A review. *Construction and Building Materials.* 2021;272:121390. <https://doi.org/10.1016/j.conbuildmat.2020.121390>
- Wang Y, Tao M, Feng D, Jiao Y, Niu Y, Wang Z. Effect of leaching behavior on the geometric and hydraulic characteristics of concrete fracture. *Materials.* 2022;15(13):4584. <https://doi.org/10.3390/ma15134584>
- Youssari FZ, Taleb O, Benosman AS. Towards understanding the behavior of fiber-reinforced concrete in aggressive environments: Acid attacks and leaching. *Construction and Building Materials.* 2023;368:130444. <https://doi.org/10.1016/j.conbuildmat.2023.130444>
- Haga K, Sutou S, Hironaga M, Tanaka S, Nagasaki S. Effects of porosity on leaching of Ca from hardened ordinary Portland cement paste. *Cement and concrete research.* 2005;35(9):1764-75. <https://doi.org/10.1016/j.cemconres.2004.06.034>
- Solpuker U, Sheets J, Kim Y, Schwartz F. Leaching potential of pervious concrete and immobilization of Cu, Pb and Zn using pervious concrete. *Journal of contaminant hydrology.* 2014;161:35-48. <https://doi.org/10.1016/j.jconhyd.2014.03.002>
- Nurjannah L, Novita DA. Uji bakteri Coliform dan Escherichia coli pada air minum isi ulang dan air sumur di kabupaten Cirebon. *Jurnal Ilmu Alam Indonesia.* 2018;1(1).
- Hosseini SA, Toghroli A. Effect of mixing Nano-silica and Perlite with pervious concrete for nitrate removal from the contaminated water. *Advances in concrete construction.* 2021;11(6):531-44. <https://doi.org/10.12989/acc.2021.11.6.531>
- Long W-J, Ye T-H, Li L-X, Feng G-L. Electrochemical characterization and inhibiting mechanism on calcium leaching of graphene oxide reinforced cement composites. *Nanomaterials.* 2019;9(2):288. <https://doi.org/10.3390/nano9020288>
- Müllauer W, Beddoe RE, Heinz D. Leaching behaviour of major and trace elements from concrete: effect of fly ash and GGBS.

- Cement and Concrete Composites. 2015;58:129-39. <https://doi.org/10.1016/j.cemconcomp.2015.02.002>
24. Lin W, Cheng A, Huang R, Chen C, Zhou X. Effect of calcium leaching on the properties of cement-based composites. Journal of Wuhan University of Technology-Mater Sci Ed. 2011;26:990-7. <https://doi.org/10.1007/s11595-011-0350-x>
25. Lee M-G, Tia M, Chuang S-H, Huang Y, Chiang C-L. Pollution and purification study of the pervious concrete pavement material. Journal of Materials in Civil Engineering. 2014;26(8):04014035. [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0000916](https://doi.org/10.1061/(ASCE)MT.1943-5533.0000916)
26. Alighardashi A, Mehrani MJ, Ramezani-pour AM. Pervious concrete reactive barrier containing nano-silica for nitrate removal from contaminated water. Environmental Science and Pollution Research. 2018;25:29481-92. <https://doi.org/10.1007/s11356-018-3008-9>

COPYRIGHTS

©2024 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

فیلتر ملات گذر (PMF) اصلاحی از ملات گذرا و بتن گذری است که به عنوان فیلتر آب طراحی شده است که بر اساس ویژگی های فیزیکی آن می تواند کدورت و باکتری ها را کاهش دهد. با این حال، از نظر شیمیایی، حاوی مواد معدنی است که می تواند در تماس با آب حل شود و در پساب یافت شود. این مطالعه با هدف تعیین عملکرد PMF در تصفیه آب های سطحی با کاهش کدورت و اثرشیا کلی و ارزیابی پتانسیل آبشویی آن انجام شد. نمونه های PMF با اختلاط ماسه (۰.۶-۰.۸۵ میلی متر)، سیمان و آب با نسبت ماسه به سیمان 4 (M) و ۵ و نسبت آب به سیمان 0.4 (w/c) ایجاد شدند. سپس هر مخلوط به لوله هایی با قطر ۸.۲ سانتی متر و ضخامت های مختلف: ۳، ۵ و ۱۰ سانتی متر قالب گیری شد. از آب سطحی خام برای آزمایش عملکرد و شستشو استفاده شد. نتایج نشان داد که PMF به طور موثر ۹۵٪ کدورت و ۹۹.۷۱٪ *E. coli* را حذف کرد که با مدت زمان فیلتراسیون افزایش یافت. PMF زمانی که با ضخامت ۱۰ سانتی متر از ۵ یا ۳ سانتی متر طراحی می شد، *E. coli* را به طور مؤثرتری کاهش داد، زیرا سطح بیشتری را برای چسبیدن و نگهداری مواد معلق و باکتری ها فراهم می کرد. افزایش قابل توجهی (میانگین درصد) pH، سختی، یون های کلسیم و TDS در PMF M4 با ضخامت ۱۰ سانتی متر نسبت به نمونه های نازک تر مشاهده شد، زیرا حاوی سیمان بیشتری بود که در تماس با آب حل می شد.



An Innovative and Reliable Hybrid Cooling Method for Electric Vehicle Motors

J. Selvan^a, S. Manavalla^{*b}

^a School of Mechanical Engineering, Vellore Institute of Technology Chennai, India

^b Electric Vehicles Incubation, Testing and Research Centre, Vellore Institute of Technology Chennai, India

PAPER INFO

Paper history:

Received 05 November 2023

Received in revised form 14 December 2023

Accepted 30 December 2023

Keywords:

Electric Vehicle Motor Cooling

Phase Change Material

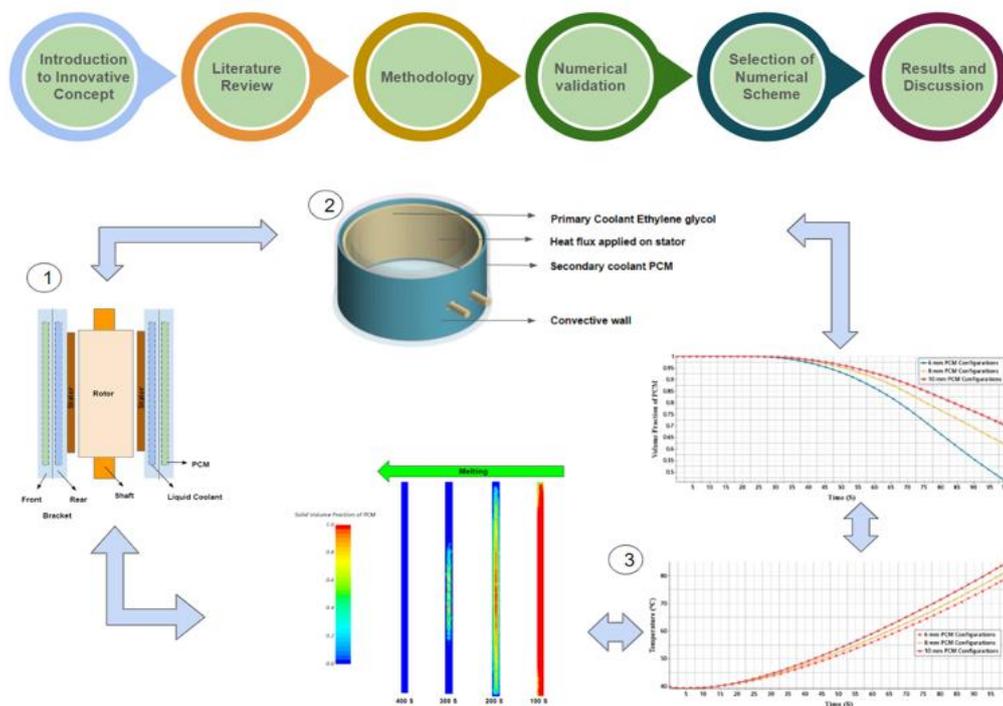
Computational Fluid Dynamics

ABSTRACT

Air, liquid, and oil are commonly used for the cooling of electric vehicle motors. Phase change materials (PCM) are not extensively used apart from electronic components. In general, coolants like air, liquid, and oil were separately used as independent coolants for the motor. In this research, two cooling channels were added to the motor cooling. The liquid is used as a primary coolant, while PCM is used as a secondary coolant. This novel method of cooling helps to keep the bracket temperature under the allowable limit even though the liquid cooling is operating at its lowest operating point. In the study, the primary focus was on the PCM coolant channel by keeping the liquid coolant under one particular operating condition to study the PCM in detail. The thickness of the PCM had an influence on the motor cooling. Three different PCM channels were studied with thicknesses of 6 mm, 8 mm, and 10 mm. The best cases were identified with a 6 mm PCM thickness, which is better in terms of heat transfer improvement of 6% observed.

doi: 10.5829/ije.2024.37.07a.06

Graphical Abstract



*Corresponding Author Email: manavalla.sreekanth@vit.ac.in (S. Manavalla)

Please cite this article as: Selvan J, Manavalla S. An Innovative and Reliable Hybrid Cooling Method for Electric Vehicle Motors. International Journal of Engineering, Transactions A: Basics. 2024;37(07):1263-73.

NOMENCLATURE			
A_{mush}	mushy zone (m^2)	\vec{v}	fluid velocity (m/s)
c_p	Specific heat at constant pressure J/(kg °C)	β	liquid volume fraction
h_{ref}	Reference enthalpy (J)	ε	constant
h	Sensible enthalpy (kJ)	ρ	Density (kg/m ³)
ΔH	latent enthalpy (J)	\vec{v}_p	solid velocity (m/s)
L	Liquid	φ	turbulence parameter
S	Source term		

1. INTRODUCTION

In recent years, electric vehicles (EVs) have gained significant popularity and have become a hot topic in the automotive industry. This surge in interest can be attributed to several factors, including the growing concern for the environment, the need to reduce dependence on fossil fuels, and advancements in technology. EVs offer a promising solution to these challenges by providing a cleaner and more sustainable mode of transportation.

Motor thermal management plays a crucial role in ensuring the optimal performance and longevity of electric vehicle motors. As electric motors convert electrical energy into mechanical energy, they generate heat as a byproduct. Excessive heat can have detrimental effects on the motor's efficiency, reliability, and overall performance. Therefore, effective thermal management is essential to maintaining the motor within the appropriate temperature range.

The basic functional requirements of any automobile vehicle that would be expected from the customer's side are

- Acceleration
- High energy and fuel economy while traveling
- Maximum speed, speed control and ride comfort
- Minimum filling/recharge time of fuel or energy
- Minimum pollution while running
- Longer Vehicle/Battery Life

The drive cycle of a fuel-powered automobile or electric vehicle (EV) completely depends on traffic conditions, road conditions, road terrain, and the load carried by the vehicle. The drive cycle also depends on the various inputs from the driver based on the road conditions. The running of an EV also affects the working condition of the electric motor, which needs to undergo frequent sudden stoppages, high acceleration, and deceleration, along with different environmental conditions. Electric motors in industrial applications operate at rated speeds and under rated conditions. Thus, the industrial electrical motor design approach cannot be applied to electric vehicles. Electric motors used in e-vehicles must meet the basic vehicle requirements for efficient operation (1).

Computational fluid dynamics (CFD) can provide detailed studies of fluid flow that cannot be obtained analytically or experimentally. Since CFD models, which have demonstrated compelling performance, can save money on prototype manufacturing. CFD has become very popular since the introduction of high-speed

computing, it is necessary to use either experimental or CFD simulation for flow visualization. Flow visualization using transparent materials was attempted earlier. Although it is simple and inexpensive, it is not suitable for colorless gases such as air. In such cases, CFD simulations can be used to visualize the flow.

Using analysis, Bellettre et al. studied (2) a model in transient state for an auto-synchronous e-machine stator and identified problematic locations on the copper winding, leading to the selection of a phase change material as the cooling system. The conventional method of testing induction motors, as analyzed by Huai et al. (3); therefore, it is becoming more important to use mathematical modeling tools and computational experiments to estimate temperature rise. This is especially relevant in the thermal study of electrical machines. Boglietti et al. (4) performed extensive surveys on the history and current approaches. The developments and the latest methods introduced over the last decade are examined in detail and contrasted to demonstrate their strengths and shortcomings. Hosain and Fdhila (5) optimized design parameters for inlet and groove geometry to increase cooling performance. Yabiku et al. (6) proposed a technique for measuring the rise in temperature of critical sections of induction e-machines under stop and start conditions by accounting for the impact of material. Cavazzuti et al. (7) introduced a powerful brushless synchronous e-machine with magnets to conduct a thermal analysis using a lumped parameter approach. Research by Kuria and Hwang (8) explains three-dimensional CFD simulations conducted over a fan-cooled brushless DC motor in a fully enclosed environment. The machine was analyzed to determine the temperatures of the critical machine components and the effects of various thermal parameters. Kim et al. (9), Kim and Kim (10), and Kim (11) suggested a few design points to enhance the E-machine thermal design and then verified the improvement of the cooling efficiency. Nategh et al. (12) conducted simulations to analyze the flow of fluids in the cooling channels by using partial modeling. The output of the numerical simulation is used as feedback to the lumped parameter approach, which will explain the transfer of heat to the fluid. He found the approaches useful because of their simplicity, since complete machine analysis is avoided. Zhou et al. (13) used a technique to construct efficient and robust electrical machine thermal models that were used for real-time thermal observation and modeling.

According to Lim and Kim (14), the solid shaft of a 35 kW high-capacity wheel e-channel machine's design was examined for oil splash cooling and numerical analysis and practical testing were used to assess the machine's cooling efficiency. Huang et al. (15) provided heat transfer modeling of a high output brushless e-machine with rotor magnets and heat pipes with liquid cooling passage. Pechánek and Bouzak (16) used CFD to analyze the two water jackets for the e-machine cooling system and explained how heat transfer can be enhanced. Li (17) proposed a numerical model for solving the thermal problem, and the developed model is in good agreement with the experiments.

Jayakumar and Sreekanth (18) discussed the novel approach of cooling an e-machine using a baffle and its arrangements. Selvan and Manavalla (19) studied the phase-change material as the primary coolant for cooling the e-machine and discussed its practical feasibility. The PCM as a secondary coolant for the motor cooling was not studied in detail before, and either liquid coolant or oil coolant was used in the past for the cooling. In this study, we are going to introduce a new concept in the cooling of the motor by adding two different cooling channels to protect the motor from overheating. The primary coolant is the liquid coolant, and the secondary coolant is the PCM. The secondary coolant is introduced to protect the motor when the liquid coolant is running at lower flow rates or in the event of coolant pump failure. The secondary coolant is the primary focus of this study,

and it is thoroughly investigated to understand its impact on the overall system.

Table 1 gives an overview of cooling methods applied by different researchers and engineering techniques to design the cooling system. Most of research focus on the air cooled machine followed by liquid, oil, heat pipe and design optimization, testing were carried out using analytical, experimental & CFD/FEA.

Phase change materials (PCM) have not been thoroughly investigated as a cooling medium for e-machines. This study investigates the OM35 PCM [20] as a coolant for the stator and brackets to assess the benefits of PCM. Phase Change Materials (PCMs) are substances that can store and release thermal energy during the process of melting and solidification. While PCMs offer several advantages such as high-energy storage density and long-term stability, they also have a few drawbacks that make them less desirable in certain applications. By highlighting these drawbacks, we can better understand why OM35 is a preferred choice.

OM35, a specific PCM, addresses some of these drawbacks. It has a wide operating temperature range (-35°C to 35°C), high thermal conductivity, and minimal super cooling and sub cooling effects. Additionally, OM35 exhibits minimal volume change during phase transition and demonstrates good cycling stability. These features make OM35 a favorable choice for various applications, including economical operation, building insulation, energy storage systems, and thermal

TABLE 1. Cooling method and engineering analysis technique used by various researchers

Reference	Cooling method				Engineering analysis technique				Motor rating, kW
	Air	Liquid	Oil	Heat Pipe	Lumped/ Mathematical	Experimental	CFD/FEA		
Lee et al. (1)	x	x	✓	x	x	✓	✓	16.7	
Bellette et al. (2)	x	✓	x	x	x	✓	✓	-	
Huai et al. (3)	✓	x	x	x	✓	x	x	1.5	
Boglietti et al. (4)	✓	x	x	x	✓	x	✓	-	
Hosain and Fdhila (5)	✓	x	x	x	x	x	✓	-	
Yabiku et al. (6)	✓	x	x	x	✓	x	x	16	
Cavazzuti et al. (7)	✓	x	x	x	✓	x	x	-	
Kuria and Hwang (8)	✓	x	x	x	x	✓	✓	-	
Kim et al. (9)	✓	x	x	x	x	✓	✓	10,25	
Kim and Kim (10)	✓	x	x	x	x	✓	✓	-	
Kim (11)	✓	x	x	x	x	✓	✓	5	
Nategh et al. (12)	✓	✓	x	x	✓	✓	✓	4	
Zhou et al. (13)	✓	x	x	x	x	✓	✓	145	
Lim et al. (14)	x	x	✓	x	x	✓	✓	35	
Huang et al. (15)	✓	✓	x	✓	✓	x	✓	-	
Pechanek and Bouzak (16)	x	✓	x	x	x	✓	✓	-	
Li et al. (17)	✓	x	x	x	x	✓	✓	-	
Jayakumar and Sreekanth (18), Selvan and Manavalla (19)	x	✓	x	x	x	✓	✓	18	

management. Other applications that use hybrid cooling applications such as "air-cooled condensers, cooling towers, surface condensers, and air-cooled heat exchangers", which were considered hybrid cooling.

The aim of this paper is to determine if the PCM can be utilized for cooling the e-machine. From the literature review so far, studies on e-machines have primarily been based on a liquid cooling system. To the best of our knowledge, PCM, along with a liquid coolant, was never studied earlier. Thus, in this article, PCM, is along with a coolant, is included in the bracket to enhance heat transfer, and its performance is studied.

2. GEOMETRY

The geometry consisting of the front and rear brackets, coolant passage, and stator stack. Figure 1 shows the schematic sketch of the proposed cooling method. To simplify the CFD analysis, only the components needed for bracket cooling are considered. The front and rear brackets hold the primary and secondary coolants, respectively.

The secondary coolant is where the innovation begins by adding the extra coolant to prevent the motor from overheating when the primary coolant is unable to adequately cool the motor or the pump is unable to deliver the coolant.

A heat source of 18 kW (19) was applied to the stator stack at an ambient temperature of 40 °C. The brackets were applied with a convective heat transfer coefficient of 10 W/m²K, as shown in Figure 2. The primary coolant of 0.01 kg/s was supplied, and the secondary coolant of PCM thickness was varied from 6 mm to 10 mm, as shown in Figure 3.

The bracket consists of primary and secondary coolants as shown in Figure 1. The primary coolant jacket is filled with the liquid coolant, and the secondary coolant

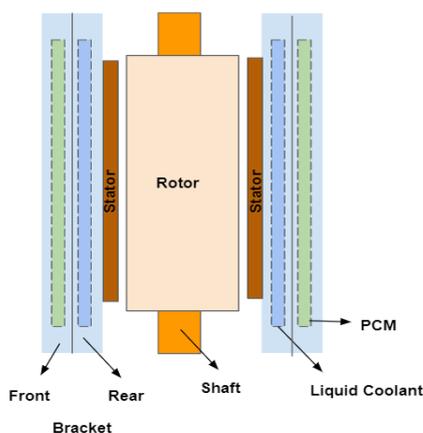


Figure 1. Schematic sketch of the cooling concept

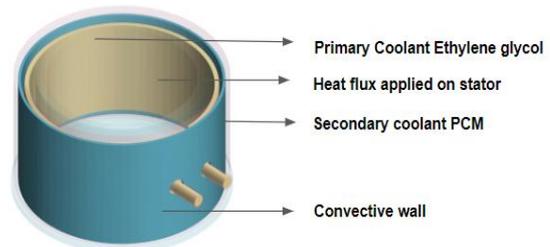


Figure 2. Simplified model of E-machine

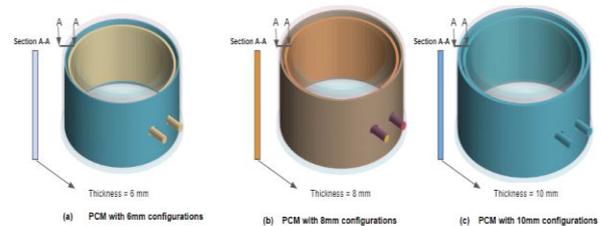


Figure 3. PCM Coolant configurations

is filled with PCM. The rotating system consists of a rotor and a shaft. The stator stack is assembled inside the brackets.

3. CONCEPT OF TWO DIFFERENT COOLING CHANNELS

The intention behind this proposed cooling system is to improve the heat transfer of the front bracket. The advantage of improving the heat transfer of the front bracket is that it will help to improve inverter cooling where the inverter is mounted on the top of the front bracket. In general, the temperatures in the brackets are roughly between 180 - 200 °C while using primary coolant alone in the system (18, 19). With the proposed method, we can reduce the operating temperatures of the brackets, which will help the inverter's operating temperature be lower. In order to simplify the analysis, the inverter is not considered in this simulation, and it is assumed that lowering the bracket temperature will help the inverter electronics cooling.

The main goal of this research is to analyze how the PCM behaves in terms of absorbing and releasing heat when placed inside the bracket, as depicted in Figure 4. The speed at which the phase change occurs directly affects the heat transfer capability of the surroundings. To avoid manufacturing constraints and excessive costs, a minimum PCM thickness of 6 mm has been chosen. Going below this threshold would result in higher manufacturing costs. Therefore, this study focuses on numerical analysis and does not include thicknesses below 6 mm.

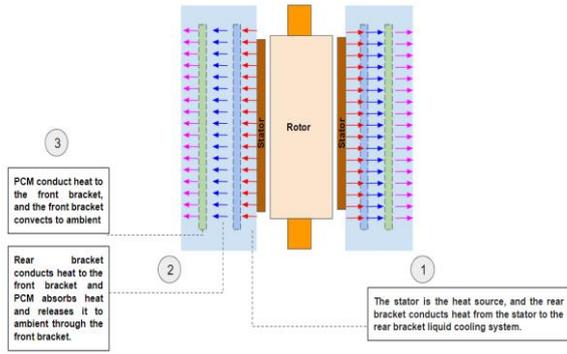


Figure 4. Concept of two different cooling system

4. METHODOLOGY

STAR-CCM+ models the solidification or melting process using an enthalpy method. The melt interface is not specifically tracked in this method. Instead, each cell in the domain has a value that separates the solid and liquid fractions present in the cell volume. At every iteration, the ratio of the medium is calculated using an enthalpy balance.

The "pseudo" porous medium that represents the mushy zone has a porosity that ranges from 1 to 0 when the substance solidifies. The porosity is zero when the material has entirely solidified in a cell, which causes the velocities to likewise be zero.

Enthalpy of a material is given by,

$$H = h + \Delta H \quad (1)$$

where,

$$h = h_{ref} + \int_{T_{ref}}^T c_p dT \quad (2)$$

The Latent Heat content can now be written in terms of the latent heat of the material, L:

$$\Delta H = \beta L \quad (3)$$

The latent heat content can vary between zero (for a solid) and L (for a liquid).

For solidification/melting problems, the energy equation is written as

$$\frac{\partial}{\partial t}(\rho H) + \nabla \cdot (\rho \vec{v} H) = \nabla \cdot (k \nabla T) + S \quad (4)$$

The momentum sink due to the reduced porosity in the mushy zone takes the following form: The momentum sink in the mushy zone is a result of the reduced porosity in that region. This reduced porosity affects the flow of momentum through the mushy zone, leading to a sink in momentum. The specific behavior and characteristics of the momentum sink in the mushy zone can vary depending on the specific situation and conditions. The momentum sink depends on the liquid volume fraction β of the mushy zone area with respect to its relative velocity (20).

$$S = \frac{(1-\beta)^2}{(\beta^2 + \epsilon)} A_{mush} (\vec{v} - \vec{v}_p) \quad (5)$$

$$S = \frac{(1-\beta)^2}{(\beta^2 + \epsilon)} A_{mush} \varphi \quad (6)$$

5. NUMERICAL VALIDATION OF CFD RESULTS WITH EXPERIMENT

The phase change of PCM was validated with experimental results to ensure the CFD results are logical and the results can be used for further research. The numerical model was tested by experiments, and a publication related to this research topic was found to validate the results. Pradeep and Venugopal (21) include experimental findings that closely resemble the scenario we are attempting to replicate. This scenario involves applying a heat source to a solid material and utilizing phase change material (PCM) to enhance the cooling process. The boundary condition is similar to the experimental setup (21) used to validate the numerical scheme. The temperature probes T1 to T8 were monitored and compared to validate the results, as shown below in Table 2 and Figure 5.

6. GRID AND TIME STEP INDEPENDENCY STUDIES

The solution obtained from the studies should be independent of the grid and time step. To confirm this, separate analyses were carried out for different grid sizes. For time dependency studies, different time steps were studied by monitoring the temperature as the field variable. The grid's independence was initially established for a grid size of 7 million cells. Beyond this, the grid was independent of bracket temperature for this study, as shown in Figure 6. After that, the time step was varied in steps of 0.001 s, and the bracket temperature was monitored. For a time step of 0.002 s and lower, the temperature remained the same as shown in Figure 7. Polyhedral cells were used, as shown in Figure 8.

TABLE 2. Comparison of temperatures as measured by Pradeep and Venugopal (21) and computed in the present work

Temperature Probe	Experiment	CFD	Deviation from Experiment, %
T1	55 °C	57 °C	3.64%
T2	48 °C	50 °C	4.17%
T3	27 °C	25 °C	7.41%
T4	27 °C	25 °C	7.41%
T5	27 °C	25 °C	7.41%
T6	27 °C	26 °C	3.70%
T7	27 °C	26 °C	3.70%
T8	27 °C	26 °C	3.70%

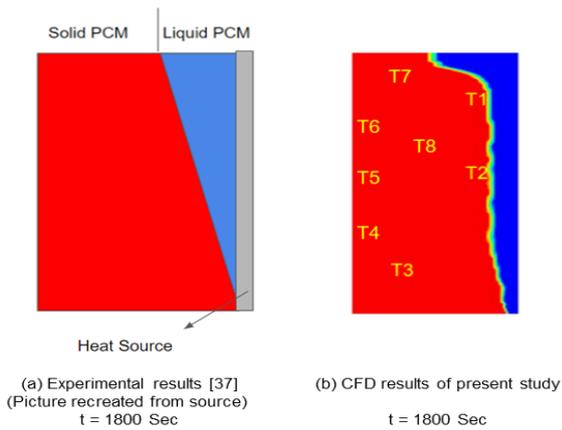


Figure 5. Comparison of CFD and Experimental results

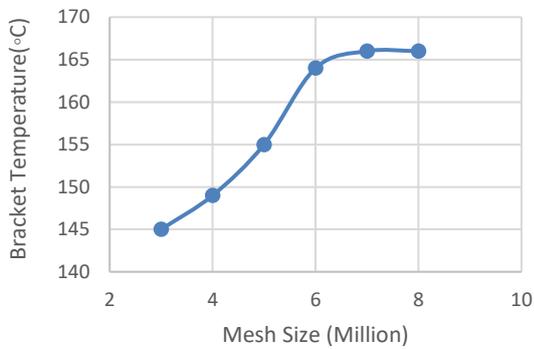


Figure 6. Results of grid independence study

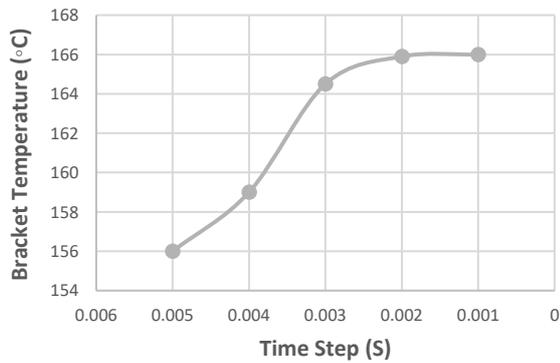


Figure 7. Results of time step independence study

7. NUMERICAL SCHEME SELECTION

In the present study, in order to simulate the change of phase, the multiphase unsteady solver was selected. Polyhedral mesh elements are used to capture the change of phases. Further, the multiphase solver with melting and solidification schemes was used to capture the melting process. The solidification process was not captured in this study. OM35 PCM was used as a

secondary coolant, and for the primary coolant, ethylene glycol was used. The material properties are shown below in Table 3 (19, 21). The volume of fractions and the temperature of brackets were monitored for convergence. Along with this continuity, momentum and energy equation residuals are monitored for convergence of $1e-2$.

8. RESULTS AND DISCUSSION

To understand the effect of primary and secondary coolant systems on the e-machine, the following parameters were monitored: The primary coolant is the liquid coolant system, where the inlet and outlet temperatures were monitored, and for the secondary cooling system, the PCM was monitored with respect to volume fraction and bracket temperature. Along with it, several PCM configurations' heat capacities were compared. Based on the above parameters, the best performing configurations were run under cyclic load conditions. In this analysis, we examined the fluctuating operating conditions and heat load of the machine, which is known as a cyclic load. To understand the complete thermal distribution along the PCM and liquid coolant passage, the unwrap section was used to show the temperature distribution between the inlet and outlet. The primary liquid coolant was kept at the lowest flow rate of 0.01 kg/s by assuming that the secondary coolant should protect the machine in the event the primary coolant fails because of pump failure.

8. 1. Different PCM Configurations are Compared for Volume Fraction

The different PCM configurations were compared with respect to the volume fraction of PCM over a period of time, as shown in Figure 9. The y axis shows the volume fraction of PCM and the x axis shows the time period. The unsteady simulations were run for a period of 100 s. This is to understand the best performing PCM configurations. After that, real world conditions were simulated by applying the cyclic load conditions.

TABLE 3. Material properties of the coolant (21, 22)

Properties	OM35	Ethylene glycol (Ethane-1,2-diol)
Melting temperature (°C)	46.5	-
Density–Solid (kg/m ³)	747	-
Density–liquid (kg/m ³)	709.5	1060
Specific heat–Solid (J/kg.K)	1650	-
Specific heat–liquid (J/kg.K)	2219	4182
Thermal conductivity–S, (W/m.K)	0.17	
Thermal conductivity–L, (W/m.K)	0.16	0.26
Viscosity–(kg/m.s)	0.01602	0.01983

The PCM is used as a secondary coolant on the rear bracket. As discussed before, the primary coolant will take on the heat load, and the excess heat will be transferred into the PCM, which will act as a secondary coolant to protect the outer casing. So it is very clear that the PCM remains solid for 30 s, as shown in Figure 9. The primary coolant is taking the heat load up to 30 s. After that, the excess heat is transferred to the PCM and the phase change starts at 30 s. In this case, three different configurations were studied with different PCM thicknesses of 6 mm, 8 mm, and 10 mm. The phase change is fastest in 6 mm configurations, followed by 8 mm, and slowest in 10 mm configurations. The initial temperature of PCM is 40 °C, and once it reaches 44 °C, the phase change of PCM begins with a 2% volume change for 6 mm and less than a 1% change for the remaining configurations. After 50 s, only 8% of PCM converts into liquid, and the remaining 92% of PCM remains solid for 6 mm configurations, and for 8 mm and 10 mm configurations, this remains less than 5%. From 50s until 60s, there is a faster change of phase in the PCM with an additional 10% change. This continues for 100 s, and it is very linear progress for PCM in 6 mm configurations, whereas this is not observed in 8 mm or 10 mm. At 100 s, the 6 mm configurations transform into 50% PCM liquid, whereas the other configurations at 40% and 30% phase change do not. Based on the results, 6 mm of PCM performs better in a given time period when compared with other configurations.

The PCM remains almost solid for 30 s, and the section view in Figure 10 shows the phase change starting from the corner of the PCM and later progressing to the center. At 50 s, the phase change happens at the top and bottom corners faster compared to the center. Until 50 s, the change in volume of PCM is very minimal. After 50 s, the phase change is faster, and at the end of 100 s, over 50% of the PCM has been transformed into liquid. The 6 mm configurations of PCM are faster for the phase change, which is better for the heat transfer.

The volume fractions of 8 mm and 10 mm PCM thickness are shown in Figures 11 and 12 respectively. The phase change was not happening until 30 s, which is almost negligible. After 30 s, the phase change starts,

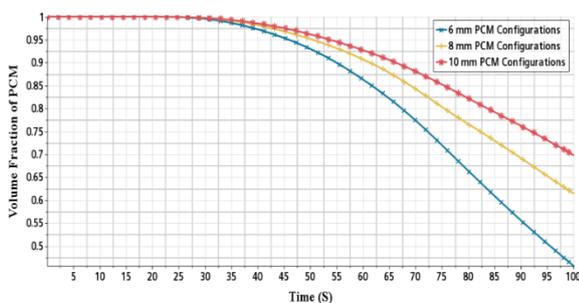


Figure 9. Different PCM configurations are compared for volume of fraction

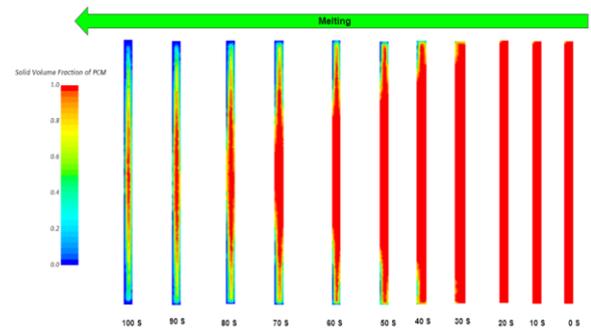


Figure 10. The volume fraction of 6 mm PCM configurations as a function of time

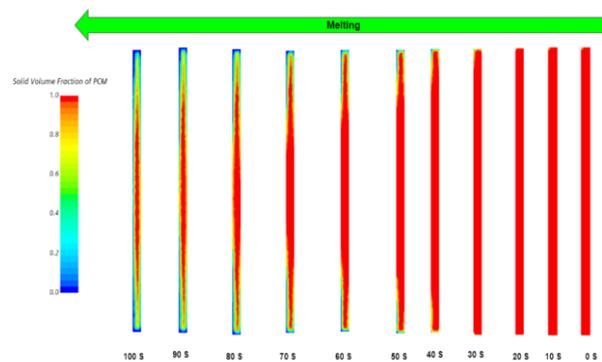


Figure 11. The volume fraction of 8 mm PCM configurations as a function of time

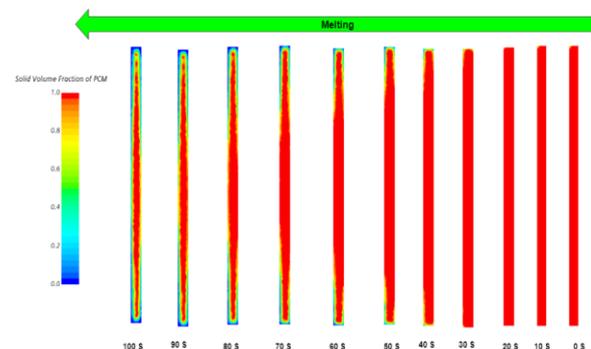


Figure 12. The volume fraction of 10 mm PCM configurations as a function of time

which is slower in both the 8 mm and 10 mm configurations. The phase change pattern looks similar to 6 mm configurations, but it is slower by 5 to 10 % throughout the time period.

8. 2. Influence on Bracket Temperature for Different PCM Configurations

The comparison of bracket temperatures for different PCM configurations was made to understand the thermal behavior. Three different PCM configurations were compared for temperature on the bracket, as shown in Figure 13. For

the first 10 seconds, the temperature of the bracket remains the same as the initial temperature. After 10 s, the temperature starts to increase. Every 5 seconds, the temperature of the bracket rises by 2 °C. For 20 seconds, the temperature of the brackets remains nearly constant across all configurations. After 20 seconds, the temperature of the bracket increases linearly, with 6 mm configurations having the lowest rate of temperature rise, followed by 8 mm configurations, and finally 10 mm configurations. From 0 s to 50 s, the temperature difference between the configurations is roughly 1 °C. After that, the temperature difference starts to increase from 2 to 5 °C. Based on the results, the best performing configuration is 6 mm, which is better in terms of volume fraction and temperature behavior. The initial temperature was 40 °C, and it ended with 78 °C, 82 °C, and 84 °C for three different configurations, i.e., 6 mm, 8 mm, and 10 mm configurations.

Different configurations' heat capacities were observed, as shown in Figure 14. The lowest heat transfer rate of 80 kW was observed for 6 mm configurations, and more than 100 kW was observed for 8 mm and 10 mm configurations, which means the amount of heat energy required to change the phase is reduced and heat is conducted through the brackets faster compared to other configurations. The volume fraction and temperature behavior results show that the 6 mm configuration performs better and helps to keep the temperature of the brackets as low as possible. Henceforth, the 6 mm configurations are simulated for cyclic load conditions.

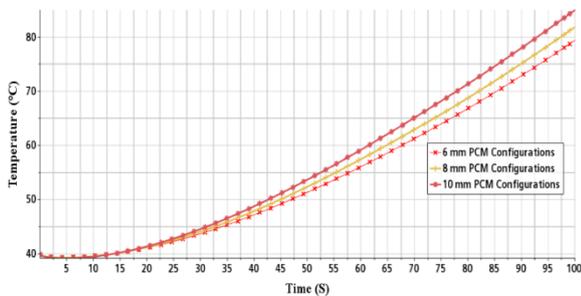


Figure 13. Bracket temperature of different PCM configurations as a function of time

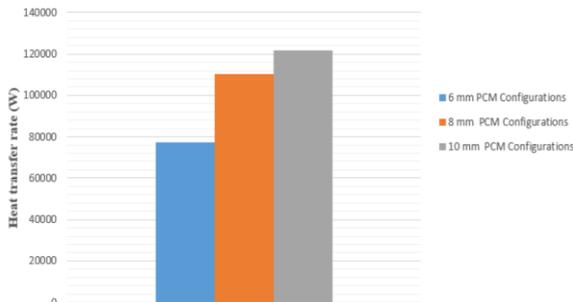


Figure 14. The heat capacity comparison of different PCM configurations

The primary liquid coolant was kept at an inlet temperature of 40 °C, and the outlet temperature was monitored for three different configurations as shown in Figure 15. According to the results, the temperature of 6 mm configurations is 79 °C, while that of 8 mm and 10 mm configurations is around 81 °C and 85 °C, respectively. The 10 mm configuration increases the overall temperature of the system when compared with the 8 mm and 6 mm configurations.

As shown in Figure 16, the PCM is unwrapped to demonstrate the thermal behaviours of various configurations. The lowest temperature is observed in 6 mm configurations, where we have the lowest temperature in the middle of the PCM and the highest temperature on the top and bottom of the wall. But when we study the other configurations, the bottom wall has a higher temperature, which shows the thermal gradient is higher and the uniform melting of PCM will not happen in the cases of 8 mm and 10 mm because of the difference in temperature between the top and bottom walls.

By comparing the bracket temperature results with previously published work [18,19] the present study shows better heat transfer, which results in lower bracket temperature as shown in Table 4.

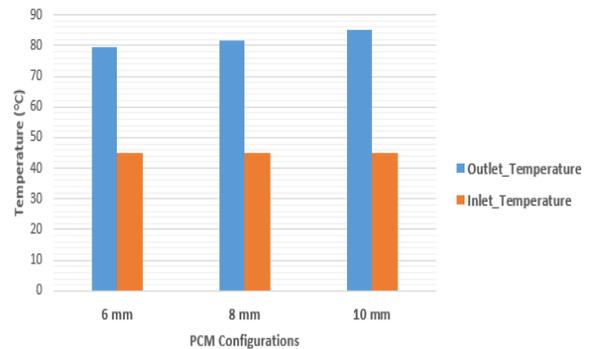


Figure 15. The Inlet and outlet temperature comparison for liquid coolant

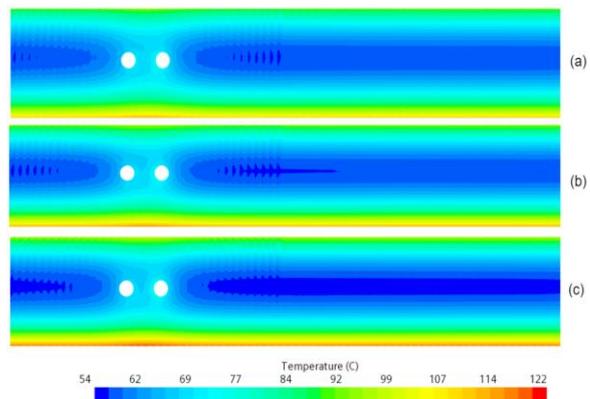


Figure 16. Different PCM configurations Unwrap (a) 6 mm (b) 8 mm (c) 10 mm are compared for temperature behavior

As explained in Figure 4, the heat transfer should happen from the stator to the ambient via brackets. In order to enhance the heat transfer, the PCM should convert the phase from solid to liquid faster so that the heat energy used for the phase change will be lower and the heat capacity will also be lower, enhancing the conduction within the solid and then to the ambient, which we can see in Figure 14 in terms of the heat capacity of the PCM. Based on the above results, the best-performing PCM configuration is 6 mm. The lower the thickness, the better the phase change of the PCM. An increase in the thickness of the PCM will not improve the heat transfer and is very ineffective, which means more heat energy will be stored within the system, which will increase the temperature of the brackets.

8. 3. Study of PCM Configurations with Cyclic Load Conditions

Among the three different PCM configurations studied based on the results, the best-performing configuration is 6 mm. It is better in terms of heat transfer and latent heat. The simulations were performed for a period of 100 s to understand the best performing configurations. The aim of these studies is to understand PCM configurations that can effectively manage thermal loads. Now the cyclic load condition will be applied to study the heat transfer behavior and the volume change of PCM materials, as shown in Figure 17. The cyclic load conditions, PCMs are subjected to repetitive heating and cooling cycles. Understanding the behavior of PCMs under these conditions is crucial for optimizing their performance and ensuring efficient thermal management. The cyclic load conditions were performed for a period of 400 s, as shown in Figure 18, which is an ideal condition for motor operation to understand the behavior (20). The heat transfer varies between 1 and 9 kW.

Figures 18 and 19 show that under cyclic load conditions, the PCM almost remains solid until 100 s. For every additional 25 s, the phase change of an additional

10 % is noticed, which is a linear progression until 260 s. Almost 50 % of phase change happens at 200 s, and in another 120 s, complete phase change happens, which is faster when compared with the first phase of the melting process, which lasts until 200 s.

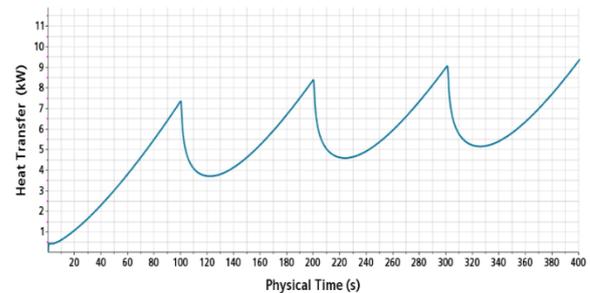


Figure 17. The 6 mm configuration applied with cyclic load conditions

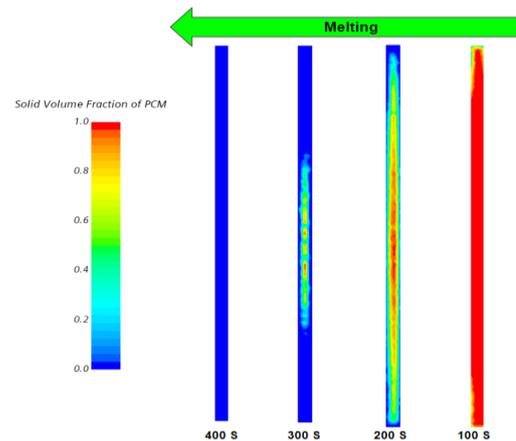


Figure 18. The volume fraction of 6 mm PCM configurations with a function of time

TABLE 4. Compassion with previous study

Bracket Temperature (°C) refereed from [18]	Bracket Temperature (°C) refereed from [19]	Bracket Temperature (°C) from present work
Liquid cooling method 100 -120 (°C)	PCM cooling method 160 (°C)	Liquid & PCM cooling method 80 (°C)

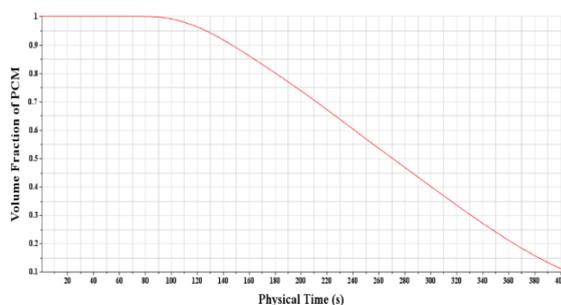


Figure 19. The volume fraction of 6 mm PCM configurations

The phase change of OM35 begins at 45 °C, where only a 1% change is observed. A bracket temperature of 76 °C causes a 50% phase change. The complete phase change happens at a temperature of 90 °C in 360 s. The temperature starts at 45 °C and rises to 100 °C after 400 seconds, as shown in Figure 20.

The PCM unwrap is to show the thermal behavior of the complete PCM from the inlet to outlet location. The bottom wall of the PCM is hotter when compared to the top wall. Under cyclic load conditions, unsymmetrical temperature distribution is observed, which results in the non-uniform melting of PCM, as shown in Figure 21.

The primary coolant unwrap shows a higher temperature at the center of the coolant passage. During cyclic load conditions, the highest temperature is observed in the center of the passage, where we have the outlet, and the coolant at a higher temperature. The non-uniform cooling observed in the PCM is due to heat transfer from the primary coolant, as shown in Figure 22.

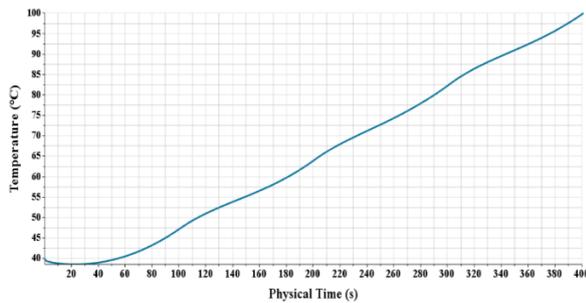


Figure 20. The thermal behaviour of 6 mm PCM configurations

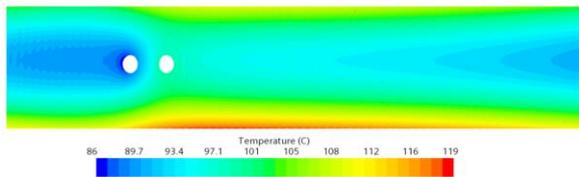


Figure 21. PCM unwrap for 6 mm configurations

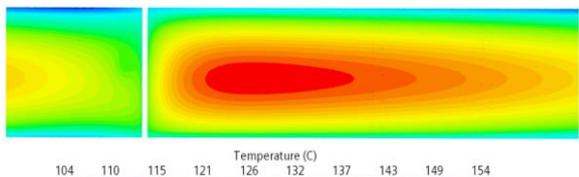


Figure 22. Primary Coolant unwrap for 6 mm configurations

9. CONCLUSION

The concept of the primary and secondary cooling systems of an e-machine works well and helps to keep the operating temperature within limits. The novel invention helps to reduce the bracket temperature from 120 °C to 80 °C, which is a significant improvement compared to other cooling techniques and is beneficial for the mild to hybrid categories of vehicle motor cooling.

In this study, the thickness of PCM, which acts as a secondary coolant, was studied, and it was found that the thickness of PCM will impact the temperature of brackets. Under constant load conditions, it took nearly 100 s to completely melt the PCM. The thicknesses of 6 mm, 8 mm, and 10 mm were studied, and it was found

that the 6 mm thickness shows better results in terms of heat transfer and uniform cooling of brackets.

Under cyclic load conditions, the PCM of 6 mm configurations helps us maintain a uniform temperature of the brackets. Further studies could improve heat transfer by adding the baffles to the cooling channel.

10. REFERENCES

1. Lee K-H, Cha H-R, Kim Y-B. Development of an interior permanent magnet motor through rotor cooling for electric vehicles. *Applied Thermal Engineering*. 2016;95:348-56. 10.1016/j.applthermaleng.2015.11.022
2. Bellettre J, Sartre V, Biais F, Lallemand A. Transient state study of electric motor heating and phase change solid-liquid cooling. *Applied Thermal Engineering*. 1997;17(1):17-31. 10.1016/1359-4311(96)00026-9
3. Huai Y, Melnik RV, Thogersen PB. Computational analysis of temperature rise phenomena in electric induction motors. *Applied Thermal Engineering*. 2003;23(7):779-95. 10.1016/S1359-4311(03)00013-9
4. Boglietti A, Cavagnino A, Staton D, Shanel M, Mueller M, Mejuto C. Evolution and modern approaches for thermal analysis of electrical machines. *IEEE Transactions on industrial electronics*. 2009;56(3):871-82. 10.1109/TIE.2008.2011622
5. Hosain ML, Fdhila RB. Air-gap heat transfer in rotating electrical machines: a parametric study. *Energy Procedia*. 2017;142:4176-81. 10.1016/j.egypro.2017.12.343
6. Yabiku R, Fialho R, Teran L, Ramos ME, Kawasaki N. Use of thermal network on determining the temperature distribution inside electric motors in steady-state and dynamic conditions. *IEEE Transactions on Industry Applications*. 2010;46(5):1787-95. 10.1109/TIA.2010.2057398
7. Cavazzuti M, Gaspari G, Pasquale S, Stalio E. Thermal management of a Formula E electric motor: Analysis and optimization. *Applied Thermal Engineering*. 2019;157:113733. 10.1016/j.applthermaleng.2019.113733
8. Kuria J, Hwang P. Investigation of thermal performance of electric vehicle BLDC motor. *International journal of mechanical engineering*. 2012;1(1):1-17.
9. SC K, W K, MS K. Cooling performance of 25 kW in-wheel motor for electric vehicles. 2013. 10.1007/s12239-013-0060-9
10. Kim DG, Kim SC. An analysis study for thermal design of ISG (Integrated Starter & Generator) for hybrid electric vehicle. *Transactions of the Korean Society of Automotive Engineers*. 2013;21(4):120-7. 10.7467/KSAE.2013.21.5.145
11. SC K. Thermal Performance of Motor and Inverter in an Integrated Starter Generator System for a Hybrid Electric Vehicle. 2013. 10.3390/en6116102
12. Nategh S, Huang Z, Krings A, Wallmark O, Leksell M. Thermal modeling of directly cooled electric machines using lumped parameter and limited CFD analysis. *IEEE Transactions on Energy Conversion*. 2013;28(4):979-90. 10.1109/TEC.2013.2283089
13. Zhou K, Pries J, Hofmann H. Computationally efficient 3-D finite-element-based dynamic thermal models of electric machines. *IEEE Transactions on Transportation Electrification*. 2015;1(2):138-49. 10.1109/TTE.2015.2456429
14. DH L, SC K. Thermal performance of oil spray cooling system for in-wheel motor in electric vehicles. 2014. 10.1016/j.applthermaleng.2013.11.057
15. Huang J, Naini SS, Miller R, Rizzo D, Sebeck K, Shurin S, et al. A hybrid electric vehicle motor cooling system—design, model,

- and control. IEEE Transactions on Vehicular Technology. 2019;68(5):4467-78. 10.1109/TVT.2019.2902135
16. Pechanek R, Bouzek L, editors. Analyzing of two types water cooling electric motors using computational fluid dynamics. 2012 15th International Power Electronics and Motion Control Conference (EPE/PEMC); 2012: IEEE. 10.1109/EPEPEMC.2012.6397424
 17. Li H. Cooling of a permanent magnet electric motor with a centrifugal impeller. International journal of heat and mass transfer. 2010;53(4):797-810. 10.1016/j.ijheatmasstransfer.2009.09.022
 18. Jayakumar S, Sreekanth M. A novel cooling method for electric vehicle motors—a CFD study.
 19. Selvan J, Manavalla S. Numerical Analysis of E-Machine Cooling Using Phase Change Material. Energies. 2022;15(15):5594. 10.3390/en15155594
 20. Durham MO, Lockerd CR. Effect of cyclic loading on motor efficiency. IEEE Transactions on industry applications. 1988;24(6):1153-9. 10.1109/28.17491
 21. Pradeep R, Venugopal T. Experimental study of lithium-ion battery cooling using mixture of phase change materials. International Journal of Electric and Hybrid Vehicles. 2020;12(2):168-83. 10.1504/IJEHV.2020.10027333
 22. Kumar S, Hassan SB, Sharma K, Baheta A. Heat Transfer Coefficients Investigation for TiO₂ Based Nanofluids. International Journal of Engineering, Transactions A: Basics. 2019;32(10):1491-6. 10.5829/ije.2019.32.10a.19

COPYRIGHTS

©2024 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

هوا، مایع و روغن معمولاً برای خنک کردن موتورهای وسایل نقلیه الکتریکی استفاده می شود. مواد تغییر فاز (PCM) جدا از قطعات الکترونیکی به طور گسترده مورد استفاده قرار نگرفتند. به طور کلی، خنک کننده هایی مانند هوا، مایع و روغن به طور جداگانه به عنوان خنک کننده مستقل برای موتور استفاده می شدند. در این تحقیق دو کانال خنک کننده به خنک کننده موتور اضافه شد. مایع به عنوان خنک کننده اولیه استفاده می شود، در حالی که PCM به عنوان خنک کننده ثانویه استفاده می شود. این روش جدید خنک سازی به حفظ دمای بَرَاکت در زیر حد مجاز کمک می کند، حتی اگر خنک کننده مایع در پایین ترین نقطه کار خود کار کند. در این مطالعه، تمرکز اولیه روی کانال خنک کننده PCM با نَگه داشتن مایع خنک کننده مایع در یک شرایط عملیاتی خاص برای مطالعه جزئیات PCM بود. ضخامت PCM بر خنک کننده موتور تأثیر داشت. سه کانال PCM مختلف با ضخامت های ۶ میلی متر، ۸ میلی متر و ۱۰ میلی متر مورد مطالعه قرار گرفتند. بهترین موارد با ضخامت PCM 6 میلی متر شناسایی شد که از نظر بهبود انتقال حرارت ۶ درصد مشاهده شد.



Analytical Modelling of a Six-Phase Surface Mounted Permanent Magnet Synchronous Motor

T. Truong Cong^a, T. Nguyen Vu^{a,b}, D. Bui Minh^{a,b}, H. Vo Thanh^c, V. Dang Quoc^{*a,b}

^a *Laboratory of High-performance Electric Machines (HiPems), Viet Nam*

^b *School of Electrical and Electronic Engineering, Ha Noi University of Science and Technology, Viet Nam*

^c *Faculty of Electrical and Electronic Engineering, University of Transport and Communications, Ha Noi, Viet Nam*

PAPER INFO

Paper history:

Received 21 November 2023

Received in revised form 24 December 2023

Accepted 06 January 2024

Keywords:

Multiple Phase Surface Permanent Magnet Synchronous Motor

Back Electromotive Force

Cogging Torque

Torque Ripple

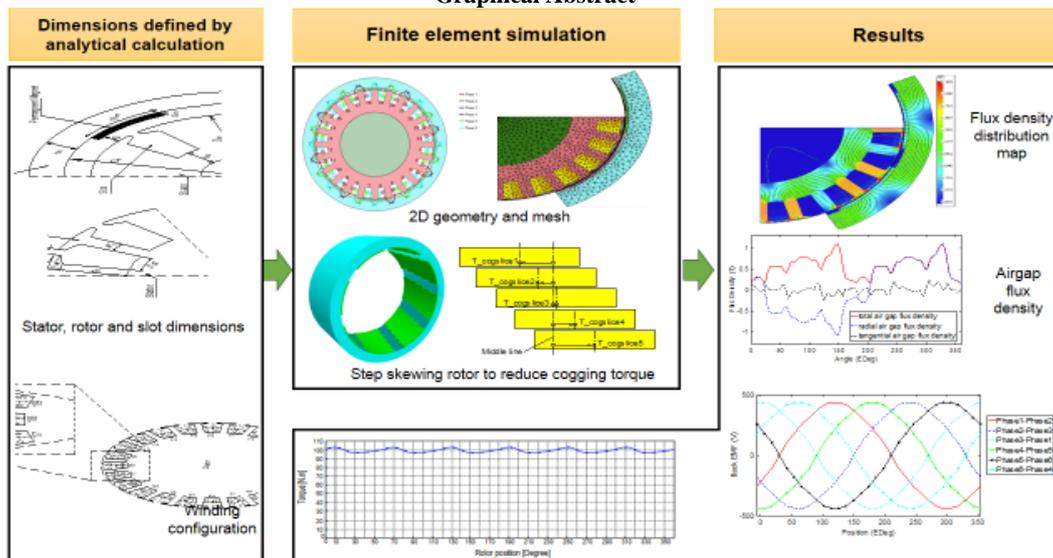
Analytical Model

ABSTRACT

A multi-phase permanent magnet synchronous motors (PMSM) has applied popularly in the field of industry (e.g. trucks, ship propulsion, mining, etc) due to its high torque, efficiency and reliable operation. So far, many researchers have studied the multi-phase PMSM (e.g. a three-phase PMSM, a six-phase PMSM) for electric vehicle applications. But, there are still significant limitations in the quantity of research on the six-phase PMSMs. Particularly, when researching this type of motor, authors mainly have provided specifications of the six-phase PMSMs and then conducted experiments on these machines without giving the detailed formulations to analytically compute and design dimensions and electromagnetic parameters. In this research, an analytic model is first developed to determine the main parameters of a six-phase surface-mounted PMSM (SPMSM). The finite element method (FEM) is then introduced to simulate and compute electromagnetic parameters, such as the current waveform, back electromotive force (EMF), flux density distribution, output torque, cogging torque, torque ripple and harmonic components. The development of proposed methods is applied on a practical problem of a six-phase SPMSM of 7.5kW.

doi: 10.5829/ije.2024.37.07a.07

Graphical Abstract



*Corresponding Author Email: vuong.dangquoc@hust.edu.vn (V. Dang Quoc)

Please cite this article as: Truong Cong T, Nguyen Vu T, Bui Minh D, Vo Thanh H, Dang Quoc V. Analytical Modelling of a Six-Phase Surface Mounted Permanent Magnet Synchronous Motor. International Journal of Engineering, Transactions A: Basics. 2024;37(07):1274-83.

1. INTRODUCTION

A six-phase permanent magnet synchronous motor (PMSM) has been recently used widely for electric vehicle applications. Because these motors have the high torque, power density, efficiency, wide speed range, reliability, durability and safety (1-5). Therefore, their extensive applications have been applied popularly in the fields of industrial, medical, aerospace and military sectors. Specifically, they are used in applications of electric vehicles, electric tractions, collaborative robots, drones electrical drive systems. So far, many researchers, designers and manufacturers have also studied three phase PMSMs. However, compared to the six phase PMSMs, these motors are still limitations for the torque and efficiency (6). In order to improve the performance and reliability of these machines, a six-phase PMSM has been proposed to use in applications demanding the high torque, efficiency and reliability. A novel configuration for a six-phase direct-drive PMSM called the 60° phase-belt toroidal winding configuration (60° -TW) was presented (1) via the finite element method (FEM) to compute the magnetic field distribution, back electromotive force (EMF), cogging torque, output torque and efficiency. In this paper, unlike conventional methods, each coil in the 60° -TW was here wound in the same direction on the stator yoke. Patel et al. (2), proposed a novel winding arrangement for the six-phase PMSM featuring 18 slots and 8 poles. This innovative configuration served to eliminate undesirable space harmonics within the stator magnetomotive force (MMF). Consequently, it leads to enhancements in power/torque density and efficiency while concurrently reducing eddy current losses in the rotor PMs and copper losses in the end windings. In addition, to improve the availability of the drive train for electric vehicles (EV) applications, this paper presented the concept of designing a six-phase PMSM as two distinct three-phase windings. A comprehensive investigation was carried out to explore various possible phase shifts between these two sets of three-phase windings, accounting for their slot-pole combinations and winding arrangements. The optimal phase shift was then selected through an analysis of harmonic distributions and their impact on the machine performance. Scuiller et al. (3) introduced a design approach tailored for multi-phase PMSMs powered by pulse-width modulation (PWM) voltage source inverters. Initially, the potential for enhancing the torque density through harmonic utilization is presented. Then, the distinctive challenges stemming from the PWM-based power supply of multi-phase machines in the design process are addressed. Islam et al. (4) conducted a performance comparison of a five-phase external rotor PM assisted synchronous reluctance motor with two distinct winding configurations. In this context, a five-phase winding configuration proposed

enhancements in power density, fault tolerance capabilities and the mitigation of torque pulsations, in contrast to conventional three-phase windings. Additionally, the incorporation of an external rotor structure contributes to further increases in power density. In, a novel six-phase PMSM with an innovative toroidal winding (NTW) configuration was also presented by Jin et al. (5) to investigate electromagnetic parameters such as the back EMF, cogging torque, torque ripple, output torque, losses and magnetic field distribution. In this paper, each coil of the NTW is uniformly wound onto the stator yoke in the same direction. The obtained results enhanced the low-speed performance and increased the output torque. They were also compared to the traditional six-phase PMSM. Del Pizzo et al. (6) explored two potential electric propulsion motor solutions for unmanned aerial vehicles. It involves a comparison of the sizes, weights, and certain characteristics between a three-phase PMSM and a six-phase motor achieved through a suitable rewinding of the armature, while maintaining fixed stator and rotor magnetic circuits. Won et al. (7) presented an innovative electric truck application featuring a six-phase fractional-slot concentrated winding PMSM. This machine design comprised a dual three-phase winding, spaced 75 degrees apart. A mathematical model for the six-phase interior PMSM (IPMSM) using the -d and-q axis theory was conducted (8). This model was subsequently utilized to deduce the precise interrelationships among different unitized machine parameters, aiming to achieve optimal performance in both inverter control (IC) and traction scenarios. Cheng et al. (9) analyzed a modeling of a six-phase surface mounted PMSM (SPMSM) based on the equivalent magnetic circuit with the magnetic behavior and electrical characteristics.

Despite many papers have researched on six-phase PMSMs as discussed above. However, there are still significant limitations in the quantity of research on these machines. Particularly, when researching these types of motors, authors mainly have provided specifications of the six-phase PMSM and then conducted experiments on these motors without giving the detailed analytical formulations for computing and analyzing their electromagnetic parameters.

In this research, an analytical design is proposed for a six phase SPMSM to determine required dimensions and electromagnetic parameters as well. Then, the FEM is introduced to verify the analytical model via the simulation of the current waveform, back EMF, flux density distribution, output torque, cogging torque, torque ripple and harmonic components.

2. MODEL OF A SIX PHASE PMSM

A structure of the six-phase SPMSM with the 60° toroidal winding (TW) is depicted in Figure 1.

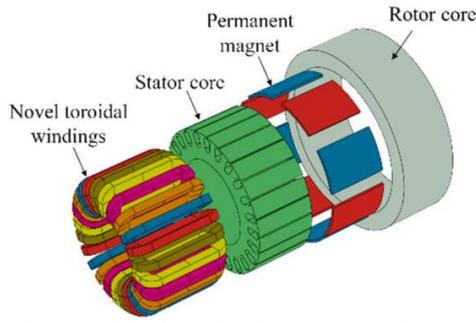


Figure 1. Modeling of the six-phase PMSM with 60°-TW (5)

This design consists of an outer rotor core, an inner stator core and the distinctive 60°-TW. Unlike the conventional winding approach, each coil is wound onto the stator yoke in the 60°-TW (5). The structural characteristics of the 60°-TW is illustrated in Figure 2. The arrangement of stator windings for the six-phase PMSM with 60°-TW is illustrated in Figure 3.

3. ANALYTICAL DESIGN

In this part, a six-phase outer rotor SPMSM of 7.5kW with a 60°-TW configuration is analytically designed. The input parameters of this machine are presented in Table 1.

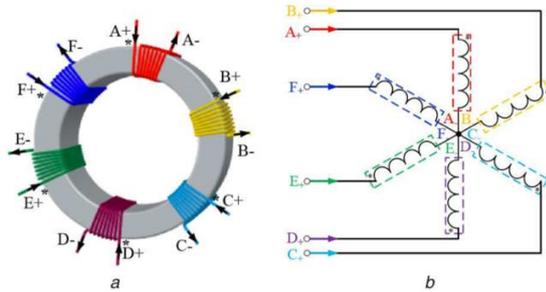


Figure 2. Structure of simplified stator of the six-phase PMSM with 60°-TW (5)

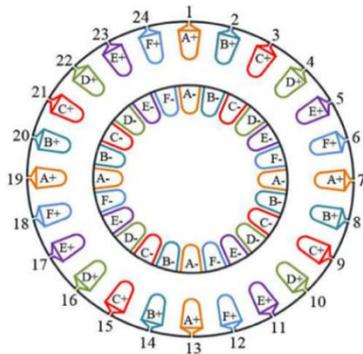


Figure 3. Arrangement of stator windings (5)

TABLE 1. Main parameters of outer rotor SPMSM

Parameters	Value	Unit
Continuous power	7.5	kW
Phase terminal voltage	200	V
Number of phases	6	phase
Number of slots	24	slot
Number of pole pair	4	pole pair
Rated torque	95.5	N.m

In a design process, the determination of parameters for PM is extremely an important part of the SPMSM as it produces the magnetic field in the air gap.

The strength of magnetic field due to the PM in the SPMSM consists of the width, length, thickness and the pole embrace. Figure 4 shows the demagnetization curves of polarization (J) and magnetic flux density (B), where the PM of NdFeB N38SH with the remanence of 1.26 T (at 20°C) and the normal working point of 0.9 T are used in this study. The main dimensions of the magnetic core are presented in Figure 5. The magnetic flux density in the air gap (B_g) is defined as:

$$B_g = \frac{4}{\pi} \sin(\alpha) B_m, \quad (1)$$

where α is the half coverage angle defined in electrical degree and B_m is the magnetic field density due to the PM. The PM thickness is defined then as:

$$d_m = \frac{\mu_m g_{eff}}{B_r A \sin(\alpha) - B_g \pi} \quad (2)$$

where μ_m , g_{eff} and B_r are respectively the permeability of PM, length of effective air gap and remanence of PM. The g_{eff} is defined via the air gap length (g) and Carter's factor (k_c), i.e (10),

$$g_{eff} = k_c \cdot g \quad (3)$$

for

$$k_c = \frac{\tau_s}{\tau_s - \gamma g} \quad (4)$$

where τ_s is the slot pitch and can be defined as:

$$\tau_s = \frac{\pi(D_{ir} - 2g)}{Z} \quad (5)$$

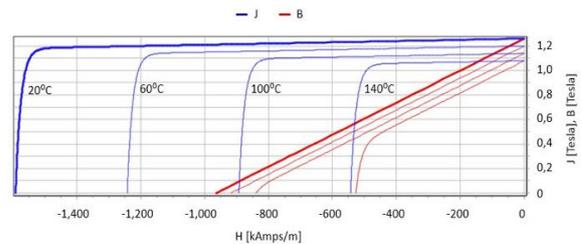


Figure 4. Demagnetization curves for N38SH (11)

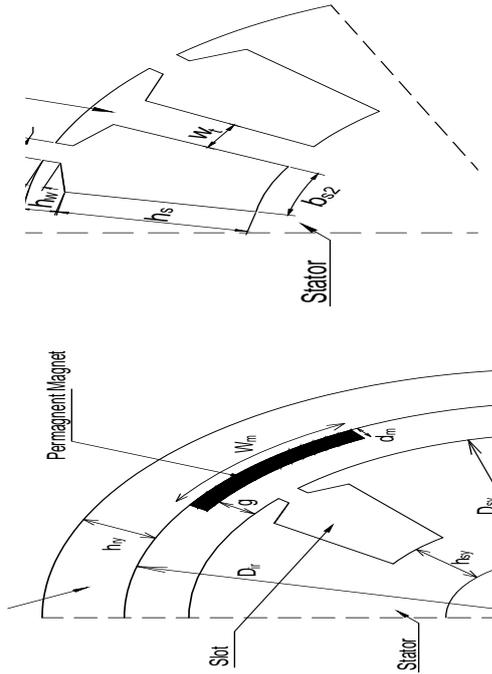


Figure 5. Main dimensions of the magnetic core (top) and slot (bottom)

where D_{ir} is the inner diameter of rotor and Z is the number of slots. The factor if PM motor (γ) is computed via the below expression (10):

$$\gamma = \frac{2b_s}{\pi g} \left[\arctan\left(\frac{b_{so}}{2(L+g)}\right) - \frac{L+g}{b_{so}} \ln \sqrt{1 + \left(\frac{b_{so}}{2(L+g)}\right)^2} \right] \quad (6)$$

where b_{so} is the width of slot opening and L is the length of PM, which is equal to the length of the both rotor and stator. The width of PM (w_m) can be now defined:

$$w_m = \frac{\alpha D_{ir}}{p} \quad (7)$$

where p is the number of pole pair. The volume of the armature part being an essential parameter in determining the D_{ir} and L can be calculated as (10):

$$\frac{\pi}{4} D_{ir}^2 L = \frac{M_n k_{safe}}{2\sigma_m} \quad (8)$$

where M_n is the rated torque of the motor, k_{safe} is the safe factor ($k_{safe} = 2 \div 3$), σ_m is the value of the shear stress of the PM (for material (NdFeB), it can be $\sigma_m = 20 - 50$ kPa). It should be noted that the relation between the D_{ir} and L is presented through the shaping coefficient (k_{shape}), that is:

$$k_{shape} = \frac{L}{D_{ir}} \quad (9)$$

From the Equation 9, the value of D_{ir} and L can be computed.

In addition, the height of stator yoke (h_{sy}) and rotor

yoke (h_{ry}) can be respectively determined as:

$$h_{sy} = \frac{B_m w_m}{2B_{sy}}, \quad h_{ry} = \frac{B_m w_m}{2B_{ry}} \quad (10a-b)$$

where the fields B_{sy} and B_{ry} are respectively the flux densities at the stator and rotor yokes given in Table 2. The width of tooth (w_t) is now defined:

$$w_t = \frac{2p B_m w_m}{Z B_t} \quad (11)$$

where B_t is the tooth flux density given in Table 2.

The number of conductors (N_c) per coil is given as:

$$N_c = \frac{U_{phase}}{2\pi\sqrt{2} f q k_w B_g \cos\delta D_{ir} L} \quad (12)$$

where U_{phase} , f , q , k_w and δ represent the phase voltage, frequency, number of slot per pole per phase, winding factor and torque angle, respectively. The torque angle for SPMSM is usually designed in the range of $15 \div 30$ degrees. In this study, it is chosen as 20 degrees (2). The slot area can be calculated as:

$$A_{slot} = \frac{4N_c A_{Cu}}{k_{fill}} \quad (13)$$

where A_{Cu} is the copper area of the conductor and k_{fill} is the slot filling factor.

As presented in Figure 5, the slot top width (b_{s1}), slot bottom width (b_{s2}), slot height (h_s) can be calculated as the below expressions:

$$b_{s1} = \frac{\pi(D_{os} - 2h_{so} - 2h_w)}{Z} - w_t \quad (14)$$

$$b_{s2} = \sqrt{\frac{b_1^2 - 4\pi \times A_{slot}}{Z}} \quad (15)$$

$$h_s = \frac{2A_{slot}}{b_1 + b_2} \quad (16)$$

where h_{so} and h_w are respectively the height and wedge of the slot opening.

Based on the analytical calculation process above, the required dimensions of a six-phase outer rotor SPMSM of 7.5 kW are given in Table 3.

4. ANALYSIS OF NO AND FULL LOAD OPERATIONS

In this part, the machine is operated under no-load conditions considered as a valuable means of assessing the motor magnetic circuit, a crucial element in motor

TABLE 2. Value of magnetic Densities of PM machine (5)

Position	Flux density (T)
Stator yoke	1.0 – 1.5
Rotor yoke	1.0 – 1.5
Tooth	1.6 – 2.0

TABLE 3. Main dimensions of a six-phase SPMSM of 7.5kW

Parameters	Value	Unit
D_{ir}	224	mm
L	112	mm
h_{ry}	19	mm
g	1	mm
w_m	63,53	mm
d_m	2,5	mm
w_t	10,6	mm
h_s	22,9	mm
b_{s1}	17	mm
b_{s2}	11	mm
b_{so}	5	mm
h_{so}	1	mm
h_w	2	mm
h_{sy}	19	mm
N_c	44	turn
n_c	224	

design. This paper focuses on studying the electromagnetic parameters such as the magnetic flux density distribution, back EMF, output torque, cogging torque and torque ripple. Based on the required parameters already given in Table 3, a 2-D model of the proposed motor is considered for both no and full load conditions.

The no-load back-EMF depending on several factors (such as winding factors, number of turns per phase, magnetic flux density in air gap, frequency) is defined as follows (2).

$$E_0 = \sqrt{2} \sin\left(\frac{y\pi}{\tau 2}\right) \frac{1}{q} \left| \sum_{m=1}^{N_c} e^{-j\theta_m} \right| p B N_c \tau f L \quad (17)$$

where y is the coil pitch, τ is the pole pitch, f is the frequency (Hz), q is the slot number of single phase per pole and θ_m is the electrical angle between adjacent slots. The term j is expressed as the current direction (with $i = 1$ for the positive current direction and $i = -1$ for the negative current direction).

The cogging torque is a type of torque appearing on the teeth that can lead to the vibration and noise in SPMSM (12, 13). When using the SPMSM in variable speed drive applications, if the frequency of torque fluctuations aligns with the mechanical resonance frequency of the stator or rotor, it can amplify the vibration and noise originating from the cogging torque. Thus, the calculation of the cogging torque is very importance in the design and production of high-performance SPMSMs.

The expression for the cogging torque (T_{cog}) is computed via the following equations (11, 14-17).

$$T_{cog}(\theta) = \frac{2LB_g^2 Z p}{\pi \mu_0 N_L} (R_{in}^2 - R_{out}^2) T_k \quad (18)$$

$$T_k = \sum_{k=1}^{\infty} \frac{K_{sk}}{k} \sin\left(k N_L \frac{b_0}{2}\right) \sin\left(k N_L \frac{\alpha_p}{2p}\right) \sin\left(k N_L \left(\theta - \frac{\alpha_s}{2}\right)\right) \quad (19)$$

$$K_{sk} = \frac{2 \sin\left(\frac{k N_L \alpha_s}{2}\right)}{k N_L \alpha_s} \quad (20)$$

where

- B_σ is the maximum magnetic flux density in air gap,
- N_L is the lowest common multiple of N_s and $2p$,
- μ_0 is the permeability of air,
- R_{in} is the inner radius of the air gap,
- R_{out} is the outer radius of the air gap,
- b_0 is the slot opening,
- α_p indicates the pole-arc to pole-pitch ratio,
- α_s is the skewing angle,
- K_{sk} is the skew factor.

5. SIMULATION RESULTS

Based on the required dimensions obtained from the analytical model given in Table 3, the FEM is introduced to compute and analyse the electromagnetic parameters of the proposed motor. The first step is considered with no skewing PM to see the waveform of the back EMF, then a skewing PM technique is presented to improve this draw back.

The 2-D geometry and mesh are presented in Figures 6 and 7, respectively. Winding configurations of a six-phase outer rotor SPMSM is presented in Figure 8. The direction of currents in the six-phase winding is pointed out in Table 4.

The distribution of back EMF waveform and output torque of the six-phase outer rotor SPMSM are shown in Figures 9 and 10, respectively. It can be seen that in

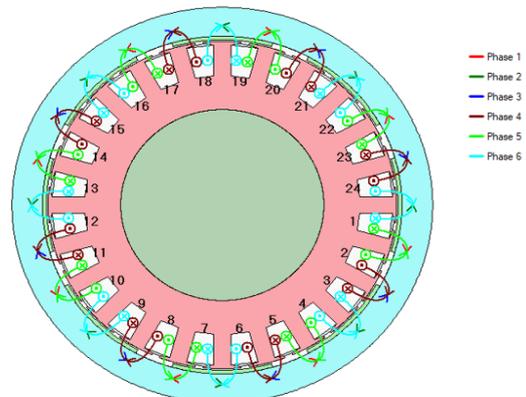


Figure 6. Geometry of the proposed motor in 2D

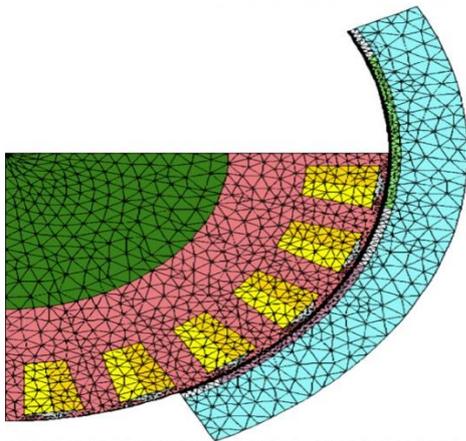


Figure 7. 2D-Mesh on a quarter view of model

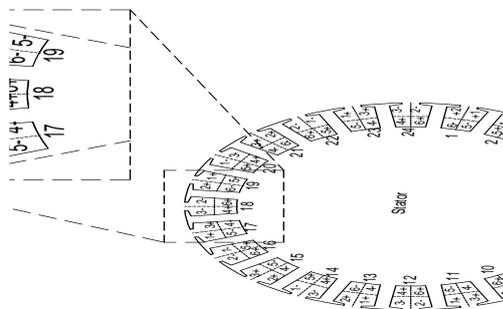


Figure 8. Winding configurations of a six-phase outer rotor SPMSM

TABLE 4. Direction of currents

Slots	Current directions	Slots	Current directions	Slots	Current directions
1	2+ 1+	2	1- 3-	3	3+ 2+
	6- 5-		5+ 4+		4- 6-
4	2- 1-	5	1+ 3+	6	3- 2-
	6+ 5+		5- 4-		4+ 6+
7	2+ 1+	8	1- 3-	9	3+ 2+
	6- 5-		5+ 4+		4- 6-
10	2- 1-	11	1+ 3+	12	3- 2-
	6+ 5+		5- 4-		4+ 6+
13	2+ 1+	14	1- 3-	15	3+ 2+
	6- 5-		5+ 4+		4- 6-
16	2- 1-	17	1+ 3+	18	3- 2-
	6+ 5+		5- 4-		4+ 6+
19	2+ 1+	20	1- 3-	21	3+ 2+
	6- 5-		5+ 4+		4- 6-
22	2- 1-	23	1+ 3+	24	3- 2-
	6+ 5+		5- 4-		4+ 6+



Figure 9. Back EMF waveform without using the skewing technique

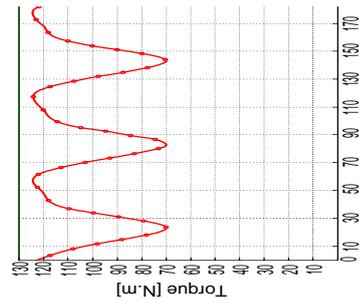


Figure 10. Output torque waveform without using the skewing technique

Figure 9, the waveform is still non sinusoidal due to high harmonic components. Thus, to make sure that the back EMF waveform is sinusoidal, the skewing technique for the PM with different angles is proposed as in Figure 11. Here, the PM is divided into five segments with different angles as given in Table 5. The minimal cogging torque with the use of skewing PM technique is presented in Figure 12. It should be noted that when the skew angle is chosen, the PM skew angle is zero for a symmetric case (see Table 5). These angles are chosen randomly to show how well the skewing technique could bring. However, these angles can be used in an optimization process to obtain the best result with the minimum torque ripple.

The map of flux density distribution with the skewing PM technique is shown in Figure 13. It can be seen that

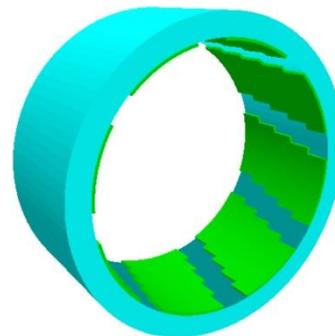


Figure 11. PM with the skewing technique

TABLE 5. Skewing angle of the PM with five slices

Segments	Angles
1	-6
2	-3
3	0
4	3
5	6

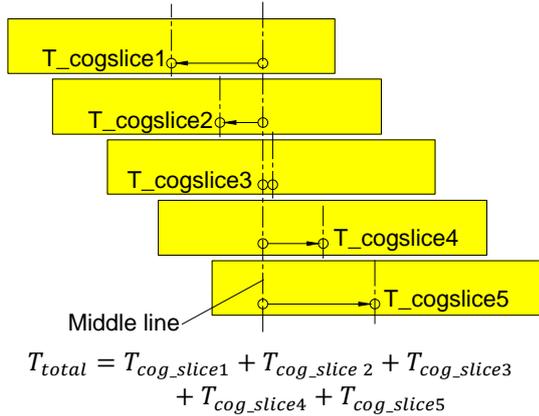


Figure 12. Cogging torque with skewing PM technique

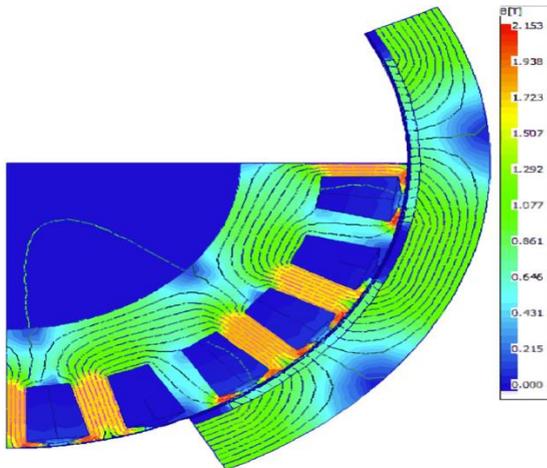


Figure 13. Flux density distribution with the skewing PM technique

the maximum value is 2.153 T, which is acceptable. It should be also noted that the higher value of flux density concentrates on the teeth and the corner of tooth tips due to the small area while the other parts of the core have the smaller value of flux density. The distribution of flux density waveform consisting of both the radial and tangential fluxes in the air gap is presented in Figure 14.

The harmonic components for this field is analyzed as shown in Figure 15. The flux linkage in no load and full load mode waveform and their harmonics order are also

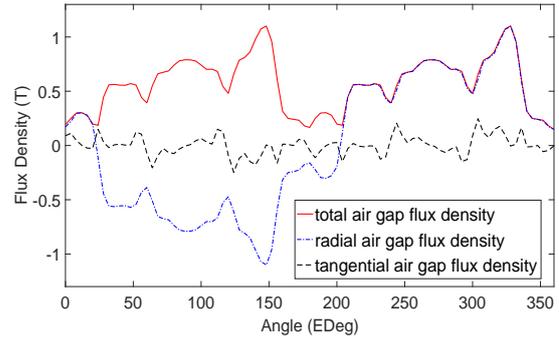


Figure 14. Waveform of air gap flux density

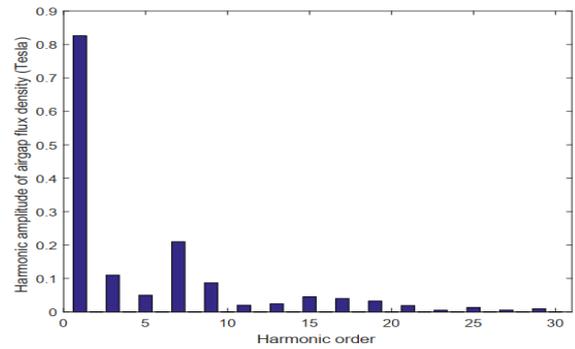


Figure 15. Harmonic components of the air gap flux density

presented in Figures 16 and 17. The back EMF waveform after using the skewing technique is shown in Figure 18. It can be seen that it is almost sinusoidal, which is the expected output result. The harmonic order of the line-line back EMF with the harmonic distortion of 2.785% is shown in Figure 19. The output torque waveform is pointed out in Figure 20. Its torque ripple is shown in Figure 21, with the value of under 3.5%.

The torque ripple holds significant importance in the design. One of the main reasons appearing the torque ripple is the cogging torque that is presented in Figure 22. This outcome signifies the motor's stable and smooth operation, a crucial aspect to be attained in the overall design.

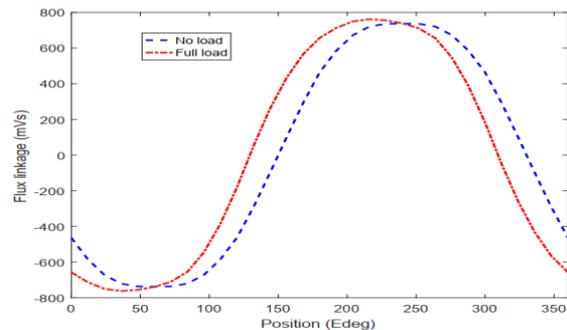


Figure 16. Flux linkage waveform

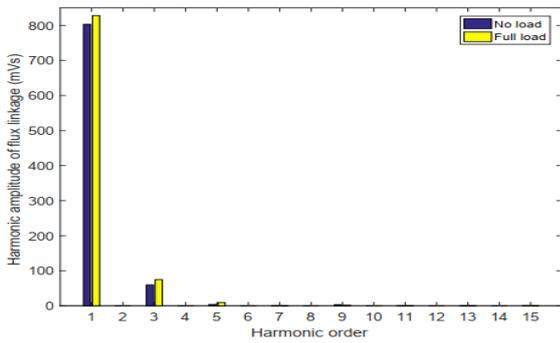


Figure 17. Harmonics order of the flux linkage under no load and full load mode

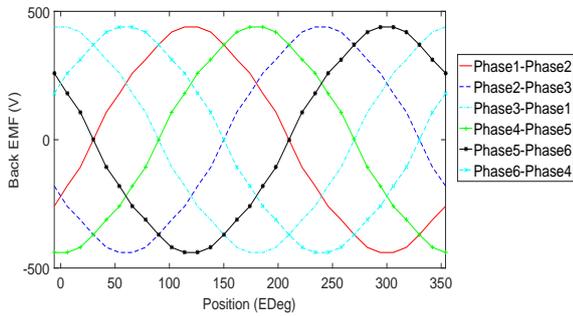


Figure 18. Back EMF waveform using the skewing technique

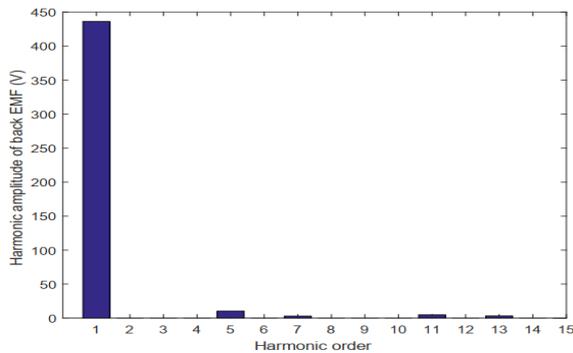


Figure 19. Harmonic components of the back EMF

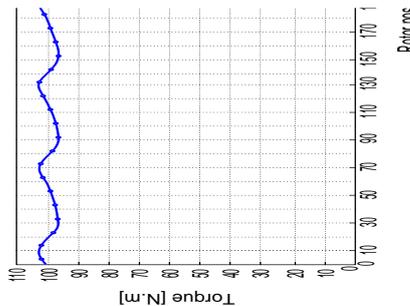


Figure 20. Output torque waveform

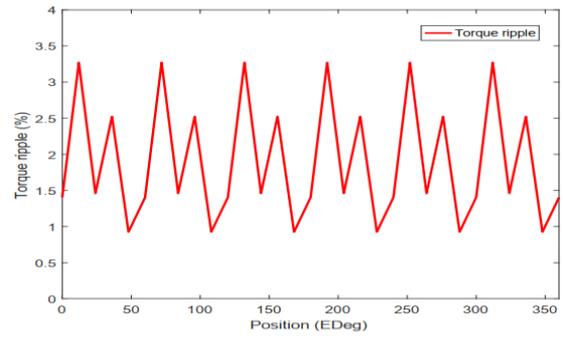


Figure 21. Torque ripple waveform

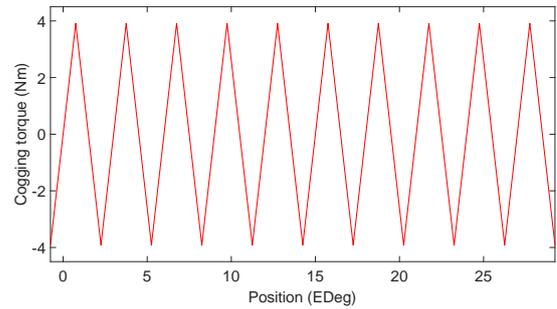


Figure 22. Cogging torque waveform

6. CONCLUSION

In this paper, the required parameters of the 7.5 kW six-phase outer rotor SPMSM using the PM material of NdFeB (N38) have been successfully obtained by the analytic model. The FEM has been also applied to verify and simulate the electromagnetic parameters of the proposed motor, such as the waveform of back EMF, output torque torque ripple and cogging torque by using the skewing PM technique. The magnetic flux density in air gap, harmonic components of the air gap flux density and waveform of flux linkage under no and full load operations have been also successfully presented. The obtained results can be served as useful reference for designers, researchers and manufactures to go on completing the prototype of design for the 7.5 kW six-phase outer rotor SPMSM. This is also a foundation for many subsequent studies, including potential research areas such as optimizing design calculations using optimization methods like genetic algorithms, swarm optimization, etc.

The developed method could be extended for calculating the network from equivalent resistances for each part of the motor. This process helps determine the waveforms of important parameters during the motor's operation, such as air gap flux density, dynamic reactance waveform, output torque waveform, as well as tooth torque waveform. Subsequently, various approaches can be proposed to improve the waveforms of these motor

parameters, aiming to enhance the overall optimization of the motor.

7. AUTHOR CONTRIBUTIONS

Authors Trinh Truong Cong and Vuong Dang Quoc conducted the research and simulation. Authors Thanh Nguyen Vu, Ha Vo Thanh and Dinh Bui Minh analyzed results. Author Vuong Dang Quoc wrote the paper; all authors had approved the final manuscript.

8. ACKNOWLEDGEMENT

This research is funded by Hanoi University of Science and Technology (HUST) under project number T2023-PC-043.

The authors also gratefully acknowledges Quy Nhon University, created favorable conditions for the authors to use the copyright-supported Ansys software program to compute and simulate the practical problem in this research. This software is the package belonging to ANSYS Electronics Desktop V19. R1.

9. REFERENCES

- Wei Y, Si J, Cheng Z, Xu S, Dong L, Liang J. Design and characteristic analysis of a six-phase direct-drive permanent magnet synchronous motor with 60° phase-belt toroidal winding configuration for electric vehicle. *IET Electric Power Applications*. 2020;14(13):2659-66. 10.1049/iet-epa.2020.0083
- Patel VI, Wang J, Wang W, Chen X. Six-phase fractional-slot-per-pole-per-phase permanent-magnet machines with low space harmonics for electric vehicle application. *IEEE Transactions on Industry Applications*. 2014;50(4):2554-63. 10.1109/TIA.2014.2301871
- Scuiller F, Semail E, Charpentier J-F, Letellier P. Multi-criteria-based design approach of multi-phase permanent magnet low-speed synchronous machines. *IET Electric Power Applications*. 2009;3(2):102-10. 110. 10.1049/iet-epa:20080003
- Islam MZ, Bonthu SSR, Choi S, editors. Comparison of two different winding topologies for external-rotor five-phase PM-assisted synchronous reluctance motor in vehicle applications. 2017 IEEE International Electric Machines and Drives Conference (IEMDC); 2017: IEEE. 10.1109/IEMDC.2017.8002399
- Jin F, Si J, Cheng Z, Su P, Dong L, Qi G, editors. Analysis of a six-phase direct-drive permanent magnet synchronous motor with novel toroidal windings. 2019 IEEE Vehicle Power and Propulsion Conference (VPPC); 2019: IEEE.
- Del Pizzo A, Di Noia LP, Di Tommaso AO, Miceli R, Rizzo R, editors. Comparison between 3-ph and 6-ph PMSM drives for the electric propulsion of unmanned aerial vehicles. 2021 IEEE 15th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG); 2021: IEEE. 10.1109/CPE-POWERENG50821.2021.9501082
- Won H, Hong Y-K, Platt J, Choi M, Bryant B, Choi S, editors. Six-phase fractional-slot concentrated winding ferrite spoke-type permanent magnet synchronous motor for electric truck. 2021 IEEE International Electric Machines & Drives Conference (IEMDC); 2021: IEEE.
- Iyer LV, Lai C, Dhulipati H, Mukundan S, Mukherjee K, Kar N. Investigation of a Six-Phase Interior Permanent Magnet Synchronous Machine for Integrated Charging and Propulsion in EVs. *SAE International Journal of Alternative Powertrains*. 2018;7(2):103-16.
- Cheng L, Sui Y, Zheng P, Yin Z, Wang C. Influence of stator MMF harmonics on the utilization of reluctance torque in six-phase PMA-SynRM with FSCW. *Energies*. 2018;11(1):108.
- Won H, Hong Y-K, Lee W, Choi M. Roles of coercivity and remanent flux density of permanent magnet in interior permanent magnet synchronous motor (IPMSM) performance for electric vehicle applications. *AIP Advances*. 2018;8(5).
- Zhu Z, Howe D. Influence of design parameters on cogging torque in permanent magnet machines. *IEEE Transactions on energy conversion*. 2000;15(4):407-12.
- Gao C, Gao M, Si J, Hu Y, Gan C. A novel direct-drive permanent magnet synchronous motor with toroidal windings. *Energies*. 2019;12(3):432.
- Heins G, Ionel DM, Thiele M. Winding factors and magnetic fields in permanent-magnet brushless machines with concentrated windings and modular stator cores. *IEEE Transactions on Industry Applications*. 2015;51(4):2924-32.
- Potgieter JH, Kamper MJ. Double PM-rotor, toothed, toroidal-winding wind generator: A comparison with conventional winding direct-drive PM wind generators over a wide power range. *IEEE Transactions on Industry Applications*. 2016;52(4):2881-91.
- Madhavan R, Fernandes BG. Axial flux segmented SRM with a higher number of rotor segments for electric vehicles. *IEEE Transactions on Energy Conversion*. 2013;28(1):203-13.
- Zhu L, Jiang S, Zhu Z, Chan C. Analytical methods for minimizing cogging torque in permanent-magnet machines. *IEEE transactions on magnetics*. 2009;45(4):2023-31.
- Won H, Hong Y-K, Choi M, Platt J, Bryant B, Choi S, et al. Novel design of six-phase spoke-type ferrite permanent magnet motor for electric truck application. *Energies*. 2022;15(6):1997. <https://doi.org/10.3390/en15061997>

COPYRIGHTS

©2024 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**

موتورهای سنکرون آهنربای دائم چند فاز (PMSM) به دلیل گشتاور بالا، راندمان و عملکرد قابل اعتماد، در زمینه صنعت (مانند کامیون ها، نیروی محرکه کشتی، معدن و غیره) کاربرد زیادی دارد. تاکنون، بسیاری از محققان PMSM چند فاز (به عنوان مثال، PMSM سه فاز، PMSM شش فاز) را برای کاربردهای خودروهای الکتریکی مورد مطالعه قرار داده اند. اما، هنوز محدودیت های قابل توجهی در کمیت تحقیقات در مورد PMSM های شش فاز وجود دارد. به ویژه، هنگام تحقیق در مورد این نوع موتور، نویسندگان عمدتاً مشخصات PMSM های شش فاز را ارائه کرده و سپس آزمایشاتی را بر روی این ماشین ها بدون ارائه فرمول های دقیق برای محاسبه تحلیلی و طراحی ابعاد و پارامترهای الکترومغناطیسی انجام داده اند. در این تحقیق، ابتدا یک مدل تحلیلی برای تعیین پارامترهای اصلی یک PMSM شش فاز روی سطح (SPMSM) توسعه داده شد. سپس روش اجزای محدود (FEM) برای شبیه سازی و محاسبه پارامترهای الکترومغناطیسی، مانند شکل موج جریان، نیروی الکتروموتور برگشتی (EMF)، توزیع چگالی شار، گشتاور خروجی، گشتاور چرخشی، ریپل گشتاور و اجزای هارمونیک معرفی می شود. توسعه روش های پیشنهادی بر روی یک مشکل عملی SPMSM شش فاز ۷.۵ کیلووات اعمال می شود.



Investigating Hostile Post Detection in Gujarati: A Machine Learning Approach

B. J. Rameshbhai*^a, K. Rana^b

^a Department of Computer Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

^b Department of Computer Engineering, Sarvajani College of Engineering and Technology, Gujarat Technological University, Ahmedabad, Gujarat, India

PAPER INFO

Paper history:

Received 03 September 2023

Received in revised form 13 December 2023

Accepted 13 January 2024

Keywords:

Hostile Text Detection

Machine Learning

Hate Text Detection

Text Classification

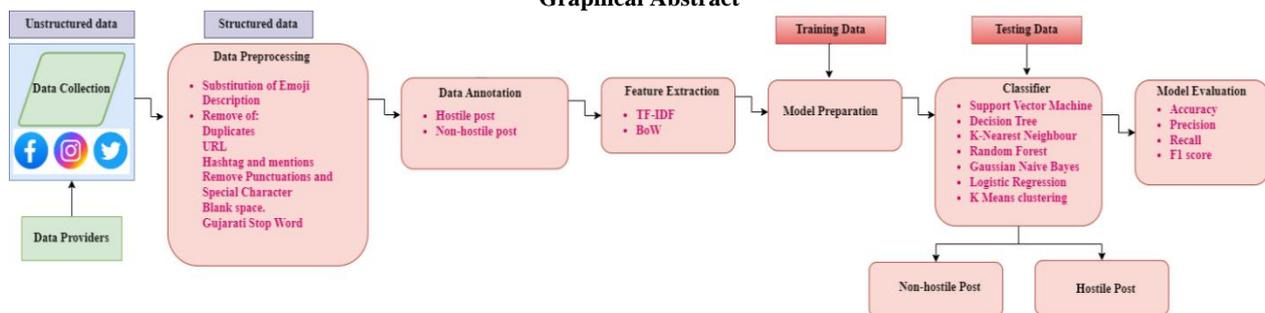
Gujarati Text Dataset

ABSTRACT

Hostile post on social media is a crucial issue for individuals, governments and organizations. There is a critical need for an automated system that can investigate and identify hostile posts from large-scale data. In India, Gujarati is the sixth most spoken language. In this work, we have constructed a major hostile post dataset in the Gujarati language. The data are collected from Twitter, Instagram and Facebook. Our dataset consists of 1,51,000 distinct comments having 10,000 manually annotated posts. These posts are labeled into the Hostile and Non-Hostile categories. We have used the dataset in two ways: (i) Original Gujarati Text Data and (ii) English data translated from Gujarati text. We have also checked the performance of pre-processing and without pre-processing data by removing extra symbols and substituting emoji descriptions in the text. We have conducted experiments using machine learning models based on supervised learning such as Support Vector Machine, Decision Tree, Random Forest, Gaussian Naive-Bayes, Logistic Regression, K-Nearest Neighbor and unsupervised learning based model such as k-means clustering. We have evaluated performance of these models for Bag-of-Words and TF-IDF feature extraction methods. It is observed that classification using TF-IDF features is efficient. Among these methods Logistic regression outperforms with an Accuracy of 0.68 and F1-score of 0.67. The purpose of this research is to create a benchmark dataset and provide baseline results for detecting hostile posts in Gujarati Language.

doi: 10.5829/ije.2024.37.07a.08

Graphical Abstract



1. INTRODUCTION

Gujarati is the sixth most spoken language in India, spoken by 56 million people¹. During covid19 pandemic the use of social media platforms such as Twitter,

Facebook, Instagram, linkedin, Reddit and YouTube has increased drastically (1, 2). Social media platforms are now used for many purposes such as education, politics, entertainment, business, charity work etc. It affects people of all age groups. On social media, people share

* Corresponding Author Institutional Email:
jagrutiboda10@gmail.com (B. J. Rameshbhai)

¹ https://en.wikipedia.org/wiki/Gujarati_language

their opinion, suggestions and emotions very openly (3, 4). Many times abusive language, the spreading of a violent message and swear words are also used. Hostile contents affect the mental health of people and promote violence thus spreading negativity among the users. The hostile post refers to abusive language mentioned in the post that targets individuals, organizations, groups of people, communities, religious, races, gender etc. (5). According to the survey, there is an increase in the content of Indian language hostile posts on social media, like Hindi (1, 2, 6-9), Marathi (7, 10-12), Bengali (8), Khasi (9), Punjabi, Gujarati and English language (9). Very less research work has been done in the low-resource Indian languages (5). Hence, we require a system that can automatically detect hostile posts written in these languages (10). Social media have a vast number of people who write their posts in Gujarati language. It is essential to have a system for Gujarati language. The state-of-the-art hostile text detection methods are available for English language. To enhance the research in low resource Indian languages, we have studied various methods which can detect hostile posts in Hindi (1, 2, 6-8, 13, 14), Marathi (7, 11, 12), Bengali (12), Saudi (4), Roman Urdu (15, 16), Tamil (17) and code-mixed language (7, 18-22). Figure 1 shows basic approaches used for hostile post-detection. Mainly these approaches are divided into two major categories: (i) Machine Learning based (23-35) and (ii) Deep Learning based (5, 6, 27-29, 34-38). Machine learning-based approach is categorized into two subparts:(i) Supervised machine learning based and (ii) Unsupervised machine learning based (27, 30-32, 37). Support vector machine (SVM) (39), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbour (KNN), Gaussian Naive Bayes (GNB) and Logistic Regression (LR) algorithms are used

for supervised learning (24). k-means clustering algorithm is used for unsupervised learning (31, 32).

To improve accuracy, researchers use deep learning-based approaches to detect hostile post in high-resource languages. The deep learning-based approach is divided into main two categories :(i) Encoder-Decoder based (9) and (ii) Transformer Based (10, 19). The Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long short-term memory (LSTM) and Bidirectional Long short-term memory (BiLSTM) (29) [32] based approaches come under the first category and Bidirectional Encoder Representations from Transformers (BERT), Romanian BERT Model (RoBERT), Multilingual BERT etc. come under the transformer based approach (3, 33, 40). To the best of our knowledge, we are the first in detecting Gujarati hostility post. We prepared a 10,000 manually labeled dataset that is valuable work in Gujarati Language Processing. This research is significant as it provides a baseline result to detect hate text in Gujarati. The proposed approach has the potential to make a more inclusive and respectful environment. The objective of our paper is to create a benchmark Gujarati hostility detection text dataset, prepare a systematic literature review and detect hostile posts using seven machine learning classifiers SVM, DT, RF, KNN, GNB, LR and K means clustering.

1. 1. Motivation People normally use their native language frequently for sharing opinions, suggestions and ideas on social media. As per research literature, Gujarati text processing has not been done for a large scale. We aim to detect hostile posts in the Gujarati language. To the best of our knowledge, machine

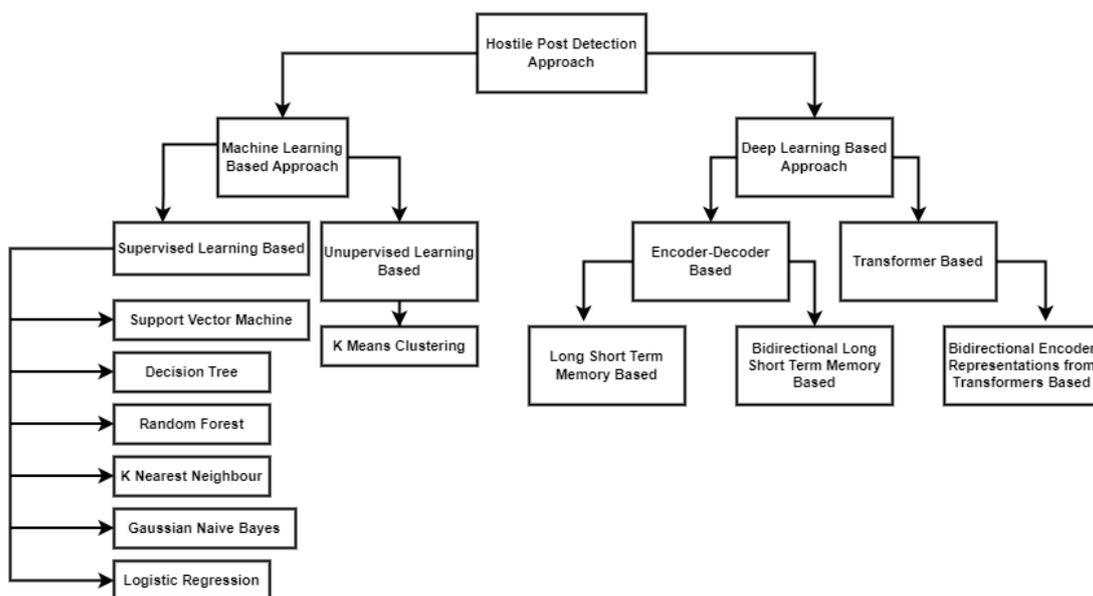


Figure 1. Existing hostile post detection approaches

learning techniques are being used for hostile post detection in the Gujarati language for the first time.

1. 2. Contribution Following are the contributions of this research work.

- There is no standard dataset available for Gujarati Language Processing (GLP). Therefore, we have created a big dataset for Gujarati language processing, consisting of 1,51,000 posts scraped from the social media platform Twitter, Instagram, and Facebook. Total 10,000 posts are manually labeled with two categories: (i) Non-Hostile and (ii) Hostile.
- The dataset is used in two ways:(i) Gujarati Text Data collected from social media (ii)Translation-based approach - the whole dataset is translated into English using Google API.
- Usefulness of emoji's in Gujarati Text Understanding
- Compared the performance of two feature extraction methods: Bag-of-Words (BoW), Term frequency-inverse document frequency (TF-IDF).
- Compared the performance of machine learning models: SVM, DT, RF, KNN, GNB, LR and k-means clustering.
- Listed major challenges that are identified for Gujarati text processing.

The remaining paper is arranged as follows: Review of existing hostile post detection is discussed in section 2. Dataset construction is discussed in section 3. The method that is used for data preprocessing, feature extraction and model development is described in section 4. Sections 5 and 6 discuss the experimental results. The error produced by our model and challenges are described in section 7. Finally, section 8 concludes our work and discusses future direction.

2. RELATED WORK

The usage of social media drastically increases day by day by people of various educational backgrounds and cultures. The inappropriate content on social media spreads negativity and damages the hygiene of the web. Detecting the hostile post for a low-resource Gujarati language is challenging because of the lack of sufficient labelled data and very less contribution from the researchers (5). We studied various approaches used for hostile post detection in different languages. Bhatnagar et al. (1) has developed an automated system that can automatically detect hostile posts in Hindi language. The proposed novel approach is used for multi-label classification. The model can also distinguish hostile posts and offensive speech. The deep learning-based model outperformed Hindi language hostile post-detection. Velankar et al. (7) presented the Marathi

Language Dataset L3CubeMahaHate that contains 25000 various samples that are classified into 4 classes. Deep learning-based various models such as CNN, LSTM, BiLSTM and transformer-based BERT models like IndicBERT, mBERT, and A Robustly Optimized BERT Pretraining Approach (RoBERTa) are used on their dataset. The dataset is also evaluated on monolingual Marathi BERT models like MahaBERT, MahaALBERT and MahaRoBERTa. The transformer approach provides a result for 4 class classification. Banerjee et al. (40) explored various transformer approaches such as mBERT, XMR-large, XLMR-based etc. These models were developed for hostility detection in Indo-Aryan and English language. The model performed excellently for English multi-class classification. The model classifies text into hate, offensive and profane categories. They tested the model for English, Hindi, Marathi and code-mixed language. They concluded that XLM-Roberta-large model outperforms. Warjri et al. (41) introduced a fake news detection approach for the Khasi language for the first time. The fake news collected from social media articles and posts. They have manually annotated 116 news and applied three machine learning classifiers RF, DT and LR. Among these classifiers DT provides the highest accuracy. Bhardwaj et al. (14) proposed a novel deep learning based HostileNet architecture. HostileNet used the concept of transformer based approach BERT. The author added hand-crafted features such as lexicon, emoticon, and hashtag embeddings with Hindi BERT to improve the accuracy for hate text detection. They used publicly available dataset CONSTRAINT-2021 dataset for coarse-grained that means binary label such as hostile and non-hostile classification and fine-grained that means multi label such as fake, hate, offensive, defamation classification. Luo et al. [26] have discussed the supervised machine learning-based approach for text classification. They created their own data that consist of different categories of data such as women, sports, literature etc. Supervised Machine Learning based classifiers SVM, KNN, NB and LR are used to classify data. SVM outperforms their data. Finally, they conclude that the classification algorithm accuracy depends on the type and size of the dataset. Felber (25) used a machine learning model for COVID-19 Fake News Detection. He focused on various text features such as n-grams, readability, emotional tone and punctuation. These text features are used for text understanding. SVM with linear Kernel, Random Forest, Logistic regression, Naive Bayes and Multilayer Perceptron are used to identify fake news detection. SVM provides the best performance for fake news detection. Fahad et al. (26) proposed a novel approach for finding bad intended news using machine learning. They aimed to set the best accuracy for detecting fake news. Most of the methods try to work on specific article domains but it is challenging to work for diverse news domains. They are using various textual

characteristics and machine learning classifiers on publicly available real world dataset to detect the fake news. The SVM and LR provide best results to detect fake news. Aluru et al. (34) performed a deep survey for hate speech detection in multilingual languages. The 16 different data sets and 9 languages were covered for hate speech detection. The deep learning-based model was developed for multilingual hate speech classification. The experimental result shows that LASER + LR (Language-Agnostic Sentence Representations + Logistic Regression model) based approach outperforms for low resource language. The transformer-based approach BERT is more effective for high-resource language. Akram et al. (32) tried a novel Deep Auto-Encoder Based Linguistics approach for Urdu News headlines clustering. For the first time a clustering model for Urdu News clustering was used. The result analysis exhibited that Urdu news headlines were easily categorized. The deep learning-based text clustering and k-means clustering algorithm was implemented for news headline clustering. The Deep learning-based approach outperformed k-means clustering approach. Deep Literature study shows that the majority of the researchers focused on high resource language. They applied Deep learning and transformer based models on existing dataset to improve the detection accuracy. From recent study, it is clear that no standard datasets available in Gujarati language and none have set a baseline result for hate text detection in Gujarati language. After

discussing all the existing work, we have found that the system can perform the following task for hostile post detection. Figure 2 shows the flow of process for hostile post detection. The detailed explanation of each phase is discussed in the subsequent sections:

3. DATASET CONSTRUCTION

We created a new dataset that covers various types of hostile posts such as racist, religious, political, educational and festival in Gujarati text. To obtain the hostile post, we have prepared a list of good and bad keywords in Gujarati with the help of two Gujarati language experts. Both experts have completed their postgraduate studies in Gujarati. These keywords are frequently used by social media users to spread hostile posts. We identified the events that occurred in Gujarat over the last five years between 2017 to 2022. Based on that we have prepared a list of 95 keywords to retrieve comments from social media. Table 1 shows a few examples of terms used for data retrieval. We collected data from the most widely used three social media platforms Twitter, Instagram and Facebook. We have used a web scrapper tool Apify to collect comments using various 95 keywords such as 'Gujarat', 'Patidar', 'Mandi', 'Sports', 'Janta', 'Corona', 'Mataji' and many more. We identified swear words that are frequently used in Gujarati language. Swear words are also used to retrieve hateful comments from social media posts. All the comments are written exclusively in Gujarati.

Twitter Apify library gives a maximum 3300 comments for a particular word search. We have collected our data without any bias. Many comments are written in a mixed language such as Gujarati-Hindi and Gujarati-English. These mixed-language comments are eliminated and Gujarati comments are selected manually. Each comment is collected using Python code and extracted

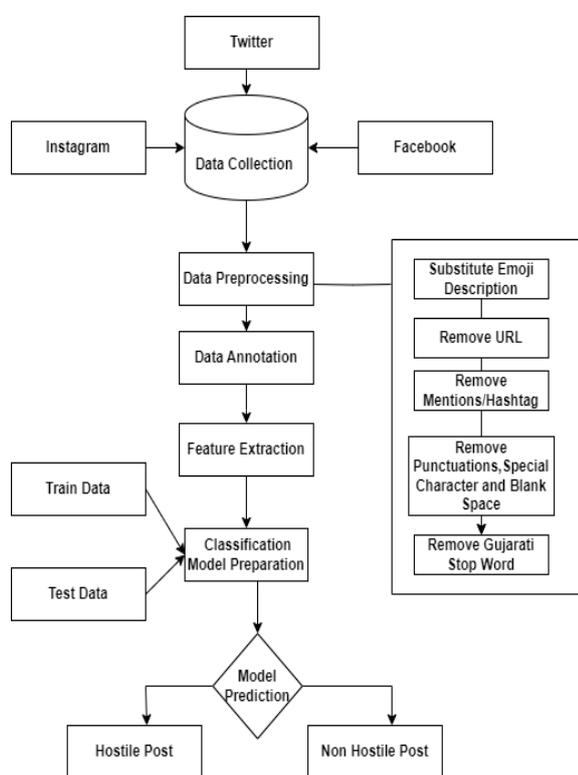


Figure 2. System flow

TABLE 1. Examples of terms used for data retrieval

Sr. No.	Search word	Total number of posts
1	પાટીદાર આંદોલન	3300
2	ગરબા	1176
3	દેશભક્તિ	3300
4	નવરાત્રી	3300
5	ગુજરાત સરકાર	507
6	મહાભારત	3300
7	માતાજી	3300
8	મુખ્યમંત્રી	3332
9	રામાયણ	3300
10	કોરોના	1020

data is collected in a CSV file. The sample of collected data is shown in Table 2.

3. 1. Data Annotation The data are manually labelled with basic two categories:

- Hostile Post: It contains hateful text and swear words that targets some gender, religious, organization, government and individual person.
- Non-Hostile Post: It does not contain harmful text which is normally neutral and positive content.

4. PROPOSED METHODOLOGY

4. 1. Pre-Processing While manually checking the data, we noticed that the dataset consists of emoji's, blank spaces, links, special symbols etc. Preprocessing is necessary step to obtain better performance for text classification problem. In natural language processing, first step is to preprocess the data. This step cleans the data and prepares comments as inputs to the classification model. We used appropriate regular expressions to clean the data. We used scikit-learn for preprocess the data.

TABLE 2. Sample of collected data

Sr. No.	Comments
1	@AAPGujarat આતંકવાદી નું સમર્થન કરતી આપ સરકાર પાકિસ્તાન ભણવા ગયેલા કાશ્મીરના યુવાનો આતંકવાદી બની ગયા, એમાંથી૧૭નાં મોત
2	#National News #Kashmir #Pakistan #Terrorist #MidDay News #MidDay Gujarati https://t.co/YheFEs4R4O
3	મારી શાયરી પસંદ કરવા બદલ હું તમારો હૃદય પૂર્વક આભારી છું.શું હું તમને ટેગ કરી શકું છું ?🙏🙏@panktinipanktio
4	રશિયા યુકેન સમજૂતી...ફોડો ફટાકડા,વગાડો ઢોલ. https://t.co/vIIKOoUO4a
5	@NobatDaily અલગાવ નહિ આતંકવાદી કહો. સાંતી દુતો થી ફાટતી લાગે છે

TABLE 3. Dataset Description

Dataset Characteristics	Collected Data
Total number of posts	1,51,000
After pre-processing dataset size	1,21,000
Annotated Data	10,000
Hostile Label Data	5000
Non-Hostile Label Data	5000
Minimum number of words in one post	1
Maximum number of words in one post	252
Minimum number of character in one post	2
Maximum number of character in one post	1444
Annotated data with Emoji	1200

- Removal of Duplicate Sentence: The duplicate data are removed from the dataset.
- Uniform Resource Locator (URL): The URL or link is removed using regular expression.
- Mentions and Hashtag: The mentions and hashtags are deleted from the data.
- Punctuations: The Special characters and punctuation marks are removed from the data.
- Blank Space: The extra spaces are also removed from the comment.
- English Hindi and Coded Mixed Sentence: Sometimes user writes comments in Hindi-Gujarati or Gujarati-English mixed language. We have removed such comments manually from the dataset.
- Gujarati Stop Words: We have prepared a list of Gujarati stopwords so that these stopwords can be removed from the dataset.
- Emojis: People normally write comments with emojis. We have replaced emoji's with appropriate description. The flow of emoji description substitution is shown in Figure 3. Table 4 illustrates all the pre-processing tasks with example.

4. 2. Feature Extraction Using Bag-of-Words and TF-IDF Model

The pre-processed sentences are converted into numeric data using the feature extraction method. Out of many feature extraction methods, we have used BoW and TF-IDF (28). During Literature study, we identified that these two feature extraction techniques are widely used to extract the feature and convert text data into numeric data.

- Bag-of-Words (36): Bag-of-words (BoW) method is used to extract the text feature to prepare a model. Bag-of-Words consist of: (i) A measure of the presence of words in the document (ii) A Vocabulary of words from the document.
- TF-IDF (27, 28, 37): One of the most effective feature extraction methods is TF-IDF. TF-IDF stands for Term Frequency – Inverse Document Frequency. TF-IDF method converts words into numeric data based on the frequency of words in the document and the importance of the word. TF is Term frequency that refers to the total number of times a given term appears in the document against the total number of words in the document. The IDF is the inverse document frequency that measures how much information a word provides. It measures the weight of a given word in the entire document.

4. 3. Model Preparation We have implemented machine learning-based techniques for hostile post detection on our dataset. We have used supervised learning based and unsupervised learning based approaches.

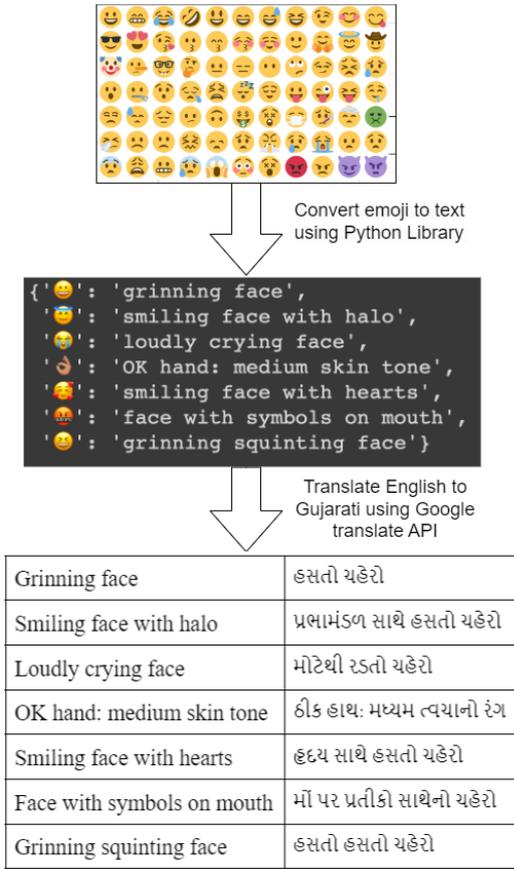


Figure 3. Flow of emoji description substitution

TABLE 4. Pre-processing task

Preprocessing step	Before preprocessing	After preprocessing
Substitute emoji description	મોઘવારી, બેરોજગારી, શિક્ષણનાં ખાનગીકરણથી બચીએ, હવે પરિવર્તન લાવીએ. 😊😊	મોઘવારી, બેરોજગારી, શિક્ષણનાં ખાનગીકરણથી બચીએ, હવે પરિવર્તન લાવીએ. હસતો ચહેરો ખુશખુશાલ. લાલી ચહેરો
Remove URL	ભારત માતા કી જય' બોલનાર દેશપ્રેમી અને ન બોલનાર દેશદ્રોહી સાબિત થતાં નથી - https://t.co/bw1A62wF6q	ભારત માતા કી જય' બોલનાર દેશપ્રેમી અને ન બોલનાર દેશદ્રોહી સાબિત થતાં નથી
Remove hashtag	પાટીદાર આંદોલન ને આશ્રય લઇને ગુજરાત રાજનીતિ માં ભૂસકો મારવો એ લાલચી @HardikPatel_ રેલી માં તો આને કહ્યું તું કે તે રાજનીતિ માં નહિ આવે. #Hardik #Fakepatidarleader	પાટીદાર આંદોલન ને આશ્રય લઇને ગુજરાત રાજનીતિ માં ભૂસકો મારવો એ લાલચી @HardikPatel_ રેલી માં તો આને કહ્યું તું કે તે રાજનીતિ માં નહિ આવે.

Remove mentions	આવિ ગયા લુખાઓ ધમકી ઓ આપવા આખી ફોજ ને છુટી મુકી દીધી છે . ભાઇ લોગ હવેતો માનસો ને કે આપીયાવ કરતા યમયા સારા હતા કમસે કમ સંસ્કાર હતા . @AAPGujarat @VISHAL_DAVE_ @Gopal_Italia	આવિ ગયા લુખાઓ ધમકી ઓ આપવા આખી ફોજ ને છુટી મુકી દીધી છે . ભાઇ લોગ હવેતો માનસો ને કે આપીયાવ કરતા યમયા સારા હતા કમસે કમ સંસ્કાર હતા
Remove blank space	કર્ણાટકના બે પૂર્વ મુખ્યમંત્રી સહિત 64 લોકોને જાનથી મારી નાખવાની ધમકી મળી	કર્ણાટકના બે પૂર્વ મુખ્યમંત્રી સહિત 64 લોકોને જાનથી મારી નાખવાની ધમકી મળી
Remove punctuations	રાજસ્થાનના પ્રતાપગઢ જિલ્લામાં બે કોમ વચ્ચે અથડામણ ફાટી નીકળી, ત્રણનાં મરણ, દને ઈજા: &કફરૂ લાગુ કરી દેવાયો!.....:*	રાજસ્થાનના પ્રતાપગઢ જિલ્લામાં બે કોમ વચ્ચે અથડામણ ફાટી નીકળી ત્રણનાં મરણ, દને ઈજા કફરૂ લાગુ કરી દેવાયો
Remove code mixed text	ધમકી આપવાની ફરિયાદ મામલે સુનાવણી કાનૂન સે ઉપર કોઈ નહીં !!	ધમકી આપવાની ફરિયાદ મામલે સુનાવણી
Remove stop word	ગજબ દેશ છે મારો, ખેતીપ્રધાન હોવા છતાંય દાળ આયાત કરી રહ્યો છે અને જીવદયા પ્રેમી હોવા છતાંય માંસની નિકાસ કરે છે	ગજબ દેશ છે મારો, ખેતીપ્રધાન હોવા છતાંય દાળ આયાત કરી રહ્યો છે જીવદયા પ્રેમી હોવા છતાંય માંસની નિકાસ કરે

- Supervised learning based approach (23): It requires labeled data while preparing the model. During the training, the model learns features and understands the text. It then classifies using classification algorithms such as Support vector Machine, Decision Tree, Random forest, K Nearest Neighbour, Gaussian Naive Bayes, Logistic Regression.
- Unsupervised learning based Approach (30, 32): For the data without labels, we use an unsupervised learning-based approach. It learns from input text features and groups them with a similar pattern. We have used the k-means clustering algorithm for creating clusters with similar features. We have used 70:30 ratio for training and testing data as shown in Table 5.

5. EXPERIMENTS

There are various machine learning algorithms used for text classification. We have performed various experiments to evaluate the machine learning model for Gujarati hostile post-detection. Amongst different

Machine learning algorithms, selection of a particular algorithm for our work was decided based on experimentation. To understand the behavior of different machine learning algorithms we have experimented using Support Vector Machine, Decision Tree, Random forest, K Nearest Neighbour, Gaussian Naive Bayes, Logistic Regression and k-means clustering algorithms. Posts are classified into hostile or non-hostile classes. In this section we describe the dataset statistics used in our experiments, experiment setup, machine learning methods used for training the model, hyperparameters values and evaluation metrics.

5. 1. Dataset Splitting The dataset contains a total of 10,000 comments to be classified as hostile or non-hostile. We have 5000 hostile and 5000 non hostile posts. We have used 70%-30% for training (7000 instances) and testing (3000 instances). For the supervised machine learning approach, we have used labeled data having text posts and the corresponding label.

5. 2. Experimental Results

5. 2. 1. Implementation Details We implemented a machine learning model using python library scikit-learn, pandas and numpy. We executed the experiments on Google Collaboratory, which provides a free Jupyter notebook environment. Table 6 shows the hyperparameter value of machine learning classifier that we have used for model tuning.

TABLE 5. Statistics of dataset splitting

Dataset	Number of Data
Training Data	7000
Testing Data	3000
Total Data	10000

TABLE 6. Hyperparameter value of classifier

Machine learning classifier	Hyperparameter
SVM	C: 1.0
	Kernel: 'rbf'
	Degree: 3
	Gamma: 'scale'
	Coef0: 0
	Shrinking: True
	Probability: False
	Criterion: 'gini'
	Splitter: 'best'
	Max depth: None
Decision Tree	Min samples split: 2
	Min samples leaf: 1
	Max features: 'auto'
	Random state: None

KNN	C: 1.0
	Kernel: 'rbf'
	Degree: 3
	Gamma: 'Scale'
	Coef0:0
Random Forest	Shrinking: True
	Probability: False
	N Estimators:100
	Criterion: 'gini'
	Max depth: None
Gaussian Naive Bayes	Min samples split: 2
	Max features:'auto'
	Bootstrap: True
	Priors: None
	Var_smoothing: 1e-9
Logistic Regression	Penalty:'l2'
	C:1.0
	Solver:'lbfgs'
	Max_iter:100
	Multi_class:'auto'
K-means clustering	Random_state:None
	Fit_intercept:True
	n_clusters: 2
	n_init: 10
	max_iter:100
	tol: 1e-4
	random_state:None
	algorithm:auto

5. 2. 2. Model Evaluation The performance of the model is evaluated using Accuracy, Macro Precision, Recall and Macro F1-score.

- **Accuracy:** Accuracy is the ratio between correct prediction and Total Number of given samples
- **Macro Precision:** Precision is the ratio between the correctly identified positives samples (true positives) and all identified positives samples. We used macro precision that provides arithmetic mean of all the precision values for the both hostile and Non-hostile classes.
- **Recall:** Recall is the ratio between the true positives and what was actually labeled.
- **Macro F1-score:** The F1 score is calculated as the arithmetic mean of precision and recall. It is used to find the average rate. The macro-averaged F1 score is the mean of all the individual class F1 scores.

In this section, we discuss the result of supervised and unsupervised machine learning models. The Two datasets used are: (i) Gujarati Text data and (ii) Translated English data. The seven supervised and unsupervised machine learning-based classifiers are used. These classifier results are evaluated using Accuracy, Macro Precision, Recall and Macro F1-score

which is shown in Table 7. We have experimented with different scenarios considering Gujarati language and

translated into English language datasets. We have also considered cases without Emoji and with Emoji's

TABLE 7. Machine learning based classifier result

Sr.no	Algorithm used	Dataset	Emoji description	Feature extraction	Accuracy	Macro Precision	Recall	Macro F1-score
1	SVM	Gujarati	No	BoW	0.65	0.65	0.67	0.66
2	Decision Tree	Gujarati	No	BoW	0.65	0.66	0.63	0.65
3	KNN	Gujarati	No	BoW	0.59	0.59	0.63	0.61
4	Random Forest	Gujarati	No	BoW	0.65	0.65	0.68	0.66
5	Gaussian Naive-Bayes	Gujarati	No	BoW	0.65	0.65	0.67	0.66
6	Logistic Regression	Gujarati	No	BoW	0.66	0.66	0.67	0.67
7	K-means clustering	Gujarati	No	BoW	0.51	0.51	0.52	0.52
1	SVM	Gujarati	No	TF-IDF	0.66	0.66	0.67	0.67
2	Decision Tree	Gujarati	No	TF-IDF	0.64	0.65	0.65	0.65
3	KNN	Gujarati	No	TF-IDF	0.58	0.58	0.66	0.62
4	Random Forest	Gujarati	No	TF-IDF	0.65	0.65	0.70	0.67
5	Gaussian Naive-Bayes	Gujarati	No	TF-IDF	0.63	0.69	0.47	0.56
6	Logistic Regression	Gujarati	No	TF-IDF	0.68	0.67	0.67	0.67
7	K-means clustering	Gujarati	No	TF-IDF	0.52	0.52	0.53	0.53
1	SVM	Gujarati	Yes	BoW	0.65	0.65	0.67	0.67
2	Decision Tree	Gujarati	Yes	BoW	0.66	0.67	0.64	0.66
3	KNN	Gujarati	Yes	BoW	0.59	0.59	0.63	0.61
4	Random Forest	Gujarati	Yes	BoW	0.65	0.65	0.68	0.66
5	Gaussian Naive-Bayes	Gujarati	Yes	BoW	0.65	0.65	0.67	0.66
6	Logistic Regression	Gujarati	Yes	BoW	0.66	0.66	0.67	0.67
7	K-means clustering	Gujarati	Yes	BoW	0.51	0.51	0.52	0.52
1	SVM	Gujarati	Yes	TF-IDF	0.66	0.66	0.67	0.67
2	Decision Tree	Gujarati	Yes	TF-IDF	0.65	0.66	0.65	0.65
3	KNN	Gujarati	Yes	TF-IDF	0.58	0.58	0.66	0.62
4	Random Forest	Gujarati	Yes	TF-IDF	0.65	0.65	0.70	0.67
5	Gaussian Naive-Bayes	Gujarati	Yes	TF-IDF	0.63	0.69	0.47	0.56
6	Logistic Regression	Gujarati	Yes	TF-IDF	0.68	0.67	0.67	0.67
7	K-means clustering	Gujarati	Yes	TF-IDF	0.52	0.52	0.53	0.53
1	SVM	English	No	BoW	0.64	0.64	0.68	0.66
2	Decision Tree	English	No	BoW	0.62	0.62	0.62	0.62
3	KNN	English	No	BoW	0.57	0.56	0.70	0.62
4	Random Forest	English	No	BoW	0.64	0.67	0.67	0.65
5	Gaussian Naive-Bayes	English	No	BoW	0.65	0.65	0.67	0.66
6	Logistic Regression	English	No	BoW	0.66	0.66	0.67	0.67
7	K-means clustering	English	No	BoW	0.51	0.51	0.52	0.52
1	SVM	English	No	TF-IDF	0.66	0.65	0.66	0.66
2	Decision Tree	English	No	TF-IDF	0.63	0.63	0.64	0.63
3	KNN	English	No	TF-IDF	0.55	0.53	0.94	0.68
4	Random Forest	English	No	TF-IDF	0.66	0.66	0.69	0.67
5	Gaussian Naive-Bayes	English	No	TF-IDF	0.65	0.65	0.65	0.66
6	Logistic Regression	English	No	TF-IDF	0.66	0.66	0.66	0.67
7	K-means clustering	English	No	TF-IDF	0.52	0.52	0.52	0.53
1	SVM	English	Yes	BoW	0.65	0.65	0.66	0.66
2	Decision Tree	English	Yes	BoW	0.65	0.65	0.67	0.66
3	KNN	English	Yes	BoW	0.58	0.58	0.64	0.61
4	Random Forest	English	Yes	BoW	0.65	0.65	0.68	0.67
5	Gaussian Naive-Bayes	English	Yes	BoW	0.65	0.65	0.67	0.66
6	Logistic Regression	English	Yes	BoW	0.66	0.66	0.67	0.67
7	K-means clustering	English	Yes	BoW	0.51	0.51	0.52	0.52

1	SVM	English	Yes	TF-IDF	0.66	0.66	0.67	0.66
2	Decision Tree	English	Yes	TF-IDF	0.64	0.64	0.66	0.65
3	KNN	English	Yes	TF-IDF	0.59	0.57	0.70	0.63
4	Random Forest	English	Yes	TF-IDF	0.66	0.66	0.69	0.67
5	Gaussian Naive-Bayes	English	Yes	TF-IDF	0.65	0.65	0.67	0.66
6	Logistic Regression	English	Yes	TF-IDF	0.66	0.66	0.67	0.67
7	K-means clustering	English	Yes	TF-IDF	0.52	0.52	0.53	0.53

description to analyze impact of Emoji. Both methods of feature extraction viz. BoW and TF-IDF with Gujarati dataset having Emoji description as well as without Emoji are evaluated in the Figure 4.

The same evaluation is also carried out for English dataset which is shown in Figure 5. Accuracy performance of the machine learning classifier ranges from 0 to 1. Supervised Learning based classifiers provide good results as compared to unsupervised learning. We have identified that the TF-IDF feature extraction method extracts good features as compared to Bag-of-Words method. Out of all machine learning methods which we have evaluated, Logistic Regression gives better results for the hostile post detection. We have also analyzed that Gujarati data gives good result as compared to the translated English data. The reason

could be that the translated English data loses its meaning during conversion because some words are not correctly translated into English. In addition, we have also tried substitution of emoji's appropriate description to understand the sentence in a better way. While analyzing the result we observed that an approach using emoji substitution improves the accuracy. However, the improvement is marginal because only 12% of sentences in our dataset contain emoji.

6. RESULTS AND DISCUSSION

We have developed the model for a large dataset and we could achieve the accuracy of 68%. Logistic regression performs well when we use it for two class classification. We are also working on two classes: hostile and non-hostile. Since we are getting comparatively good results for the Logistic regression method (36-39). We have also compared the result of Bag-of-Words and TF-IDF feature extraction methods. We conclude that TF-IDF is more efficient as compared to Bag-of-Words method because BoW provides the frequency of words in a document whereas TF-IDF provides additional data such as how a word is important in the document. Translated English data could not perform well due to inefficient Gujarati text translation. The result of both data shows that there is no wide difference in accuracy. The Gujarati text data provide baseline result 68% and Translated data provide baseline result 66% using TF-IDF feature extraction method.

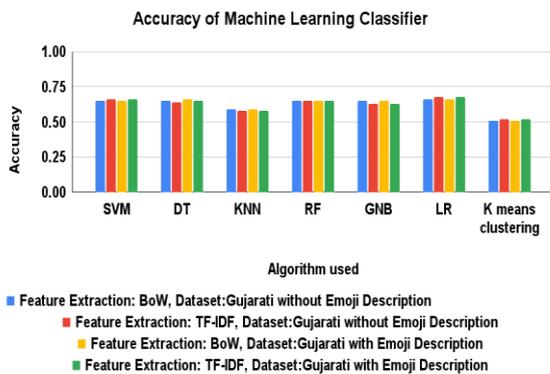


Figure 4. Evaluation of ML approaches for Gujarati data

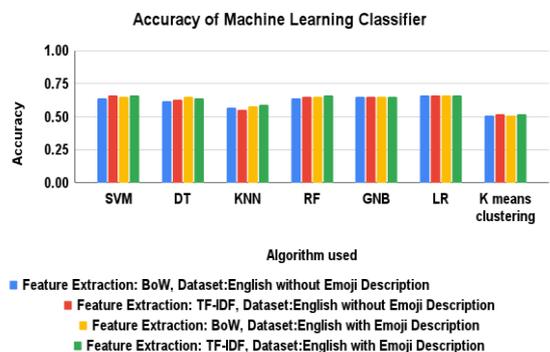


Figure 5. Evaluation of ML approaches for English data

7. ERROR ANALYSIS

In this section, we have thoroughly analyzed the errors produced by our models to understand limitations and challenges of hostile post detection in Gujarati language. We have described all the challenges we have faced. We have illustrated some of the misclassified input sentences and probable reason for misclassification in Table 8. The result does not get optimal due to many challenges. The challenges are mentioned in the next section:

7. 1. Challenges in Hostile Post Detection in Gujarati Language

- Pre-Processing: There is no standard library

TABLE 8. Error Analysis

Sr. No.	Misclassified Input sentence	Reason for Misclassification
1	આતંકવાદી કયા ધર્મ ના હતી	From one sentence we are unable to predict the actual context of the sentence. It required the full content to understand the sentence's meaning.
2	આતંકવાદી ઓ સે મેરા પુરાના નાતા છે	The sentence contains Hindi words but they are written in Gujarati language. These types of sentences are impossible to understand.
3	આ ફોટો જોઈને કટ્ટર હિન્દુ વાદ ના જંડા લઈ ને ફરતા હોય એમને વિચાર કરવો સત્તા માટે જ બધા ખેલ હોય છે જેમાં તમે કુદયા કરો સો	Few comments are based on some images or video, so it is difficult to interpret without seeing the associated image.
4	Wow પેટ્રોલ -ડીઝલમાં ભાવમાં ઘટાડો પેટ્રોલમાં 9 અને ડીઝલમાં 6 રૂપિયાનો ઘટાડો ગેસના સિલિન્ડર પર પણ 200 રૂપિયાની રાહત જાવ જવસા કરો.	The Gujarati and English mixed sentences are completely not removed from the dataset.

available that perfectly performs the text preprocessing steps in the Gujarati language. Therefore, we have to make regular expressions to preprocess the data.

- Data collection: Data collection is also one challenge for us because normally people write comments in code-mixed language. Therefore, it is difficult to collect texts that are purely written in Gujarati text.
- Manually data annotation for testing Data: Machine learning requires labeled data. It is a huge task to manually label the data.
- No standard method is available: High resource languages such as English have huge resources Dataset, wordnet, preprocessing techniques, feature extraction techniques, automatic data annotation techniques and algorithms for text data understanding. For low-resource Indian languages these resources are not available.
- Important data loss: During pre-processing some meaningful data may be lost. While removing mentions, few natural words are removed from the text. For example, “@Gopal_Italia આમ આદમી પાર્ટી માં કોળી સમાજ નુ મહત્વ કેટલું ?” that is a non-hostile post. But the classifier incorrectly classified it as hostile. The reason behind that is that @Gopal_Italia is removed from the sentence during preprocessing. Another example, “21 Days #પરિવર્તન_યાત્રા અંતર્ગત ગુજરાતની તમામ 182

વિધાનસભામાં પહોંચીને જનતાનાં પ્રશ્નો સાંભળીશું.” is also not correctly classified because the English word Days and numeric values are removed in preprocessing.

- Code-mixed text: The data having code-mixed sentences loses its meaning after applying preprocessing. For example, the post, “રશિયાએ ભારતીય student ને બહાર કાઢવા યુદ્ધ રોક્યું હતું આ પણ Fake news હતી” was a mixed Gujarati and English language sentence. The English words ‘student’ and ‘Fake news’ are removed during preprocessing. Therefore, the meaning of the sentence is not correctly identified. Another example, “Aantakvadi ni pream kahani nathi hoti” was a Gujarati post written in English and thus was not correctly classified.

8. CONCLUSION AND FUTURE WORK

Hostile Post detection in the Gujarati language has become a notable problem therefore there is a huge requirement for automation systems for hostile post detection. In this work, we have developed a Gujarati Text dataset that contains more than one lakh posts having hostile posts and non-hostile posts. The data are collected from social media Twitter, Instagram and Facebook in time duration from the year 2017 to 2022. Total 10,000 data are manually labeled with two major categories having equal number of hostile and non-hostile posts. The data contains unwanted symbols, URLs, duplicates, punctuations and emoticons. Therefore, we performed a data cleaning process and removed unwanted data. Emoji’s are often used in the short text for the expression. We have replaced emoji’s with their description in the post. We evaluated supervised and unsupervised machine learning models: Support vector Machine, Decision Tree, Random forest, K Nearest Neighbour, Gaussian Naive Bayes, Logistic Regression, and k-means clustering algorithms. The result shows that the Supervised learning-based algorithm Logistic Regression outperforms. There are various challenges identified during the implementation. In this paper, we tried to overcome these challenges. However, better ways to address these challenges are still possible. We believe that our dataset will be beneficial in Gujarati Language Processing related studies.

9. DATA AVAILABILITY

This dataset is not publicly available currently due to the thesis defense has not been completed yet. If you have any query, contact the corresponding author.

10. REFERENCES

- Bhatnagar V, Kumar P, Bhattacharyya P. Investigating hostile post detection in hindi. *Neurocomputing*. 2022;474:60-81. 10.1016/j.neucom.2021.11.096
- Bhardwaj M, Akhtar MS, Ekbal A, Das A, Chakraborty T. Hostility detection dataset in Hindi. *arXiv preprint arXiv:201103588*. 2020. <https://doi.org/10.48550/arXiv.2011.03588>
- Dowlagar S, Mamidi R. Hasocone@ fire-hasoc2020: Using bert and multilingual bert models for hate speech detection. *arXiv preprint arXiv:210109007*. 2021. <https://doi.org/10.48550/arXiv.2101.09007>
- Alshaalan R, Al-Khalifa, H., editor Hate speech detection in saudi twittersphere: A deep learning approach. In: *Proceedings of the Fifth Arabic Natural Language Processing Workshop*; 2020: IEEE. 10.1109/I2CT51068.2021.9418073
- Joshi R, Karnavat R, Jirapure K, Joshi R, editors. Evaluation of deep learning models for hostility detection in hindi text. 2021 6th International conference for convergence in technology (I2CT); 2021: IEEE. 10.1109/I2CT51068.2021.9418073
- Bhatnagar V, Kumar P, Moghili S, Bhattacharyya P, editors. Divide and conquer: an ensemble approach for hostile post detection in Hindi. *Combating Online Hostile Posts in Regional Languages during Emergency Situation: First International Workshop, CONSTRAINT 2021, Collocated with AAAI 2021, Virtual Event, February 8, 2021, Revised Selected Papers 1*; 2021: Springer.
- Velankar A, Patil H, Gore A, Salunke S, Joshi R. Hate and offensive speech detection in hindi and marathi. *arXiv preprint arXiv:211012200*. 2021. <https://doi.org/10.48550/arXiv.2110.12200>
- Phung TM, Cloos J. An exploratory experiment on Hindi, Bengali hate-speech detection and transfer learning using neural networks. *arXiv preprint arXiv:220101997*. 2022. <https://doi.org/10.48550/arXiv.2201.01997>
- Kamble S, Joshi A. Hate speech detection from code-mixed hindi-english tweets using deep learning models. *arXiv preprint arXiv:181105145*. 2018. <https://doi.org/10.48550/arXiv.1811.05145>
- Velankar A, Patil H, Gore A, Salunke S, Joshi R. L3cube-mahahate: A tweet-based marathi hate speech detection dataset and bert models. *arXiv preprint arXiv:220313778*. 2022. <https://doi.org/10.48550/arXiv.2203.13778>
- Glazkova A, Kadantsev M, Glazkov M. Fine-tuning of pre-trained transformers for hate, offensive, and profane content detection in english and marathi. *arXiv preprint arXiv:211012687*. 2021. <https://doi.org/10.48550/arXiv.2110.12687>
- Chavan T, Patankar S, Kane A, Gokhale O, Joshi R. A twitter bert approach for offensive language detection in marathi. *arXiv preprint arXiv:221210039*. 2022. <https://doi.org/10.48550/arXiv.2212.10039>
- Kamal O, Kumar A, Vaidhya T, editors. Hostility detection in hindi leveraging pre-trained language models. *Combating Online Hostile Posts in Regional Languages during Emergency Situation: First International Workshop, CONSTRAINT 2021, Collocated with AAAI 2021, Virtual Event, February 8, 2021, Revised Selected Papers 1*; 2021: Springer.
- Bhardwaj M, Sundriyal M, Bedi M, Akhtar MS, Chakraborty T. HostileNet: Multilabel Hostile Post Detection in Hindi. *IEEE Transactions on Computational Social Systems*. 2023. 10.1109/TCSS.2023.3244014
- Khan MM, Shahzad K, Malik MK. Hate speech detection in roman urdu. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*. 2021;20(1):1-19. <https://doi.org/10.1145/3414524>
- Haq NU, Ullah M, Khan R, Ahmad A, Almogren A, Hayat B, et al. USAID: an intelligent system for slang and abusive text detection in PERSO-Arabic-scripted Urdu. *Complexity*. 2020;2020:1-7. <https://doi.org/10.1155/2020/6684995>
- Anbukkarasi S, Varadhaganapathy S. Deep learning-based hate speech detection in code-mixed Tamil text. *IETE Journal of Research*. 2023;69(11):7893-8. <https://doi.org/10.1080/03772063.2022.2043786>
- Farooqi ZM, Ghosh S, Shah RR. Leveraging transformers for hate speech detection in conversational code-mixed tweets. *arXiv preprint arXiv:211209986*. 2021. <https://doi.org/10.48550/arXiv.2112.09986>
- Biradar S, Saumya S, editors. *liitdwd@ tamilnlp-acl2022: Transformer-based approach to classify abusive content in dravidian code-mixed text*. *Proceedings of the second workshop on speech and language technologies for Dravidian languages*; 2022. 10.18653/v1/2022.dravidianlangtech-1.16
- Mohapatra SK, Prasad S, Bebart DK, Das TK, Srinivasan K, Hu Y-C. Automatic hate speech detection in english-odia code mixed social media data using machine learning techniques. *Applied Sciences*. 2021;11(18):8575. <https://doi.org/10.3390/app11188575>
- Nayak R, Joshi R. Contextual hate speech detection in code mixed text using transformer based approaches. *arXiv preprint arXiv:211009338*. 2021. <https://doi.org/10.48550/arXiv.2110.09338>
- Sreelakshmi K, Premjith B, Soman K. Detection of hate speech text in Hindi-English code-mixed data. *Procedia Computer Science*. 2020;171:737-44. <https://doi.org/10.1016/j.procs.2020.04.080>
- Luo X. Efficient English text classification using selected machine learning techniques. *Alexandria Engineering Journal*. 2021;60(3):3401-9. <https://doi.org/10.1016/j.aej.2021.02.009>
- Sanoussi MSA, Xiaohua C, Agordzo GK, Guindo ML, Al Omari AM, Issa BM, editors. Detection of hate speech texts using machine learning algorithm. 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC); 2022: IEEE. 10.1109/CCWC54503.2022.9720792
- Felber T. Constraint 2021: Machine learning models for COVID-19 fake news detection shared task. *arXiv preprint arXiv:210103717*. 2021. <https://doi.org/10.48550/arXiv.2101.03717>
- Fahad N, Goh KM, Hossen MI, Shopnil KS, Mitu IJ, Alif MAH, et al. Stand up against bad intended news: An approach to detect fake news using machine learning. *Emerging science journal*. 2023;7(4):1247-59. 10.28991/ESJ-2023-07-04-015
- Defersha N, Tune K. Detection of hate speech text in afan oromo social media using machine learning approach. *Indian J Sci Technol*. 2021;14(31):2567-78. 10.17485/IJST/v14i31.1019
- Badjatiya P, Gupta S, Gupta M, Varma V, editors. Deep learning for hate speech detection in tweets. *Proceedings of the 26th international conference on World Wide Web companion*; 2017.
- Bangyal WH, Qasim R, Rehman NU, Ahmad Z, Dar H, Rukhsar L, et al. Detection of fake news text classification on COVID-19 using deep learning approaches. *Computational and mathematical methods in medicine*. 2021;2021:1-14. <https://doi.org/10.1155/2021/5514220>
- Aggarwal CC, Zhai C. A survey of text classification algorithms. *Mining text data*. 2012:163-222. https://doi.org/10.1007/978-1-4614-3223-4_6
- Allahyari M, Pouriyyeh S, Assefi M, Safaei S, Trippe ED, Gutierrez JB, et al. A brief survey of text mining: Classification, clustering and extraction techniques. *arXiv preprint*

- arXiv:170702919. 2017.
https://doi.org/10.48550/arXiv.1707.02919
32. Akram MW, Salman M, Bashir MF, Salman SMS, Gadekallu TR, Javed AR. A novel deep auto-encoder based linguistics clustering model for social text. *Transactions on Asian and Low-Resource Language Information Processing*. 2022. https://doi.org/10.1145/3527838
33. Qasim R, Bangyal WH, Alqarni MA, Almazroi AA. A fine-tuned BERT-based transfer learning approach for text classification. *Journal of healthcare engineering*. 2022;2022. https://doi.org/10.1155/2022/3498123
34. Aluru SS, Mathew B, Saha P, Mukherjee A. Deep learning models for multilingual hate speech detection. arXiv preprint arXiv:200406465. 2020. https://doi.org/10.48550/arXiv.2004.06465
35. Hassan SU, Ahamed J, Ahmad K. Analytics of machine learning-based algorithms for text classification. *Sustainable operations and computers*. 2022;3:238-48. https://doi.org/10.1016/j.susoc.2022.03.001
36. Indrawan G, Setiawan H, Gunadi A. Multi-class svm classification comparison for health service satisfaction survey data in bahasa. *HighTech and Innovation Journal*. 2022;3(4):425-42. 10.28991/HIJ-2022-03-04-05
37. Balamurugan V, Vedanarayanan V, Sahaya Anselin Nisha A, Narmadha R, Amirthalakshmi T. Multi-label Text Categorization using Error-correcting Output Coding with Weighted Probability. *International Journal of Engineering, Transactions B: Applications*. 2022;35(8):1516-23. 10.5829/ije.2022.35.08b.08
38. Dorrani Z. Traffic Scene Analysis and Classification using Deep Learning. *International Journal of Engineering, Transactions C: Aspects*. 2024;37(3):496-502. 10.5829/IJE.2024.37.03C.06
39. Zare F, Mahmoudi-Nasr P. Feature Engineering Methods in Intrusion Detection System: A Performance Evaluation. *International Journal of Engineering, Transactions A: Basics*. 2023;36(7):1343-53. 10.5829/ije.2023.36.07a.15
40. Banerjee S, Sarkar M, Agrawal N, Saha P, Das M. Exploring transformer based models to identify hate speech and offensive content in english and indo-aryan languages. arXiv preprint arXiv:211113974. 2021. https://doi.org/10.48550/arXiv.2111.13974
41. Warjri S, Pakray P, Lyngdoh SA, Maji AK, editors. Fake news detection using social media data for Khasi language. 2023 *International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC)*; 2023: IEEE. 10.1109/ISACC56298.2023.10083518

COPYRIGHTS

©2024 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**

پست خصمانه در رسانه های اجتماعی یک موضوع حیاتی برای افراد، دولت ها و سازمان ها است. نیاز اساسی به یک سیستم خودکار وجود دارد که بتواند پست های متخاصم را از داده های مقیاس بزرگ بررسی و شناسایی کند. در هند، گجراتی ششمین زبان پرگویش است. در این کار، ما یک مجموعه داده پست خصمانه اصلی به زبان گجراتی ساخته ایم. داده ها از تویتر، اینستاگرام و فیس بوک جمع آوری شده است. مجموعه داده ما شامل ۱۵۱۰۰۰ نظر مجزا است که ۱۰۰۰۰ پست مشروح دستی دارند. این پست ها در دسته های خصمانه و غیر خصمانه برچسب گذاری شده اند. ما از مجموعه داده به دو صورت استفاده کرده ایم: (i) داده های متن اصلی گجراتی) و (ii) داده های انگلیسی ترجمه شده از متن گجراتی). همچنین عملکرد پیش پردازش و بدون پیش پردازش داده ها را با حذف نمادهای اضافی و جایگزینی توضیحات ایموجی در متن بررسی کرده ایم. ما آزمایش هایی را با استفاده از مدل های یادگیری ماشین مبتنی بر یادگیری بدون نظارت مانند خوشه بندی k-means انجام داده ایم. ما عملکرد این مدل ها را برای روش های استخراج ویژگی لجستیک، k-نزدیک ترین همسایه و مدل مبتنی بر یادگیری بدون نظارت مانند خوشه بندی k-means انجام داده ایم. ما عملکرد این مدل ها را برای روش های استخراج ویژگی Bag-of-Words و TF-IDF ارزیابی کرده ایم. مشاهده می شود که طبقه بندی با استفاده از ویژگی های TF-IDF کارآمد است. در میان این روش ها رگرسیون لجستیک با دقت ۰.۶۸ و FI-score عملکرد بهتری دارد. هدف از این تحقیق ایجاد یک مجموعه داده معیار و ارائه نتایج پایه برای تشخیص پست های خصمانه در زبان گجراتی است.



Effect of Training Data Ratio and Normalizing on Fatigue Lifetime Prediction of Aluminum Alloys with Machine Learning

M. Matin, M. Azadi*

Faculty of Mechanical Engineering, Semnan University, Semnan, Iran

PAPER INFO

Paper history:

Received 01 December 2023

Received in revised form 20 December 2023

Accepted 31 December 2023

Keywords:

Machine Learning

Fatigue Lifetime

Extreme Gradient Boosting

Aluminum Alloys

Normalization Techniques

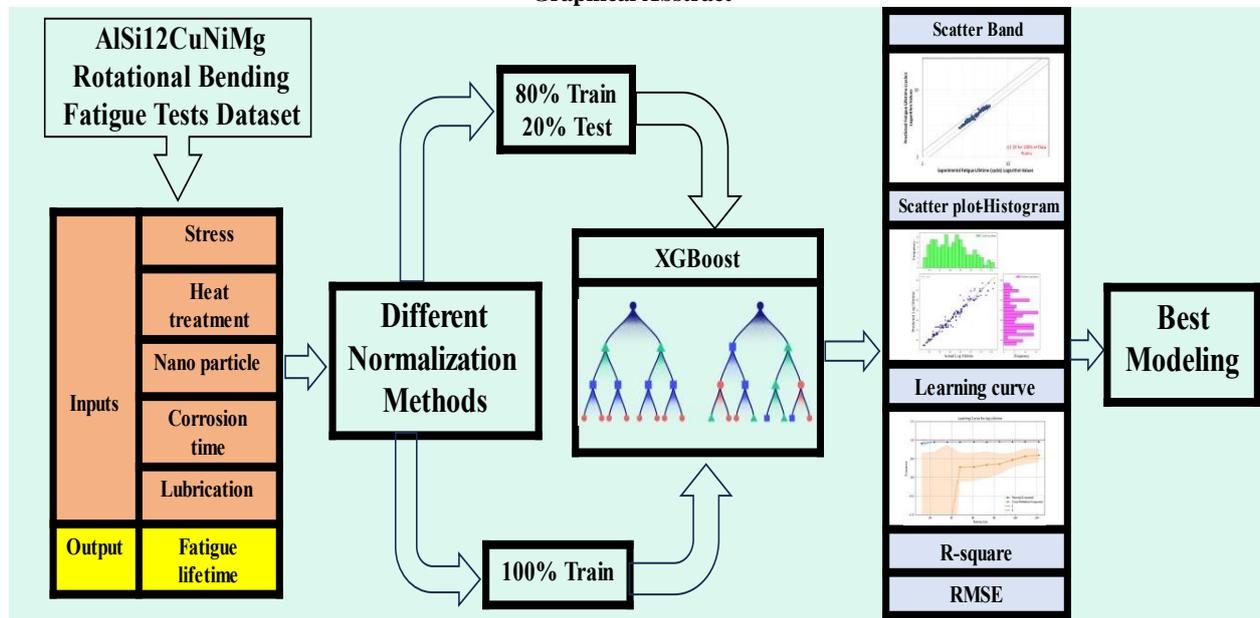
Training Data Percentage

ABSTRACT

It is critical to evaluate the estimation of the fatigue lifetimes for the piston aluminum alloys, particularly in the automotive industry. This paper investigates the effect of different normalization methods on the performance of the fatigue lifetime estimation using Extreme Gradient Boosting (XGBoost), as a supervised machine learning method. For this purpose, the dataset used in this study includes various physical and experimental inputs related to an aluminum alloy and the corresponding fatigue lifetime outputs. Furthermore, before fitting the XGBoost model, different fatigue lifetime preprocessing methods were utilized and evaluated using metrics such as Root Mean Square Error (RMSE), Determination Coefficient (R^2), and Scatter Band (SB). The results indicate that modeling fatigue lifetime with logarithmic values as a preprocessing method excels when XGBoost is trained with 100% of the data. However, other normalization methods demonstrate superior accuracy in estimating test data with a 20% test and 80% train set split.

doi: 10.5829/ije.2024.37.07a.09

Graphical Abstract



*Corresponding Author Email: m_azadi@semnan.ac.ir (M. Azadi)

1. INTRODUCTION

Fatigue loading and the resulting crack propagation are significant issues in the industry (1, 2). Researchers have attempted to increase the fatigue lifetime of industrial equipment and materials by employing various methods, including geometry-related technologies such as autofrettage (3), and manufacturing procedures such as welding (4). As a result, numerous researchers have attempted to improve the fatigue behavior of materials, geometries and estimate their fatigue lifetime. This representation is valid in the realm of alloy and metal fabrication. The following paragraph will address this claim.

First, the mechanical behavior of metals and alloys will be discussed, with a focus on aluminum alloys in the literature. Azadi et al. (5) investigated how the dwell time, thermomechanical loading factor, and maximum temperature affected the thermomechanical fatigue behavior of aluminum alloys. Furthermore, research efforts have expanded to investigate the effects of various aging heat treatments on the hardness of aluminum alloy (6). Akhtar et al. (7) investigated the optimal heat treatment temperature for aluminum alloy, revealing that 175°C was the most effective. Azadi and Parast (8) demonstrated in their work that the fatigue lifetime of aluminum alloys depends on the variation of the stress, fretting force, heat treatment condition, nanoparticles, and lubrication. Subsequently, following a thorough examination of the literature on the mechanical properties of materials, particularly aluminum alloys, it is clear that artificial intelligence (AI) methods could be helpful in estimating fatigue lifetime, providing motivation for further research.

After discussing the dependence of fatigue lifetimes on various variables and motivating using AI methods in the fatigue lifetime estimation in the second step, the literature provides the following insights.

Choi (9) estimated the steel fatigue lifetime of various chemical compositions using five different machine-learning methods. Based on his findings, XGBoost had the best accuracy with a determination coefficient (R^2) of 98.50%. Moreover, due to the substantial variability in the fatigue lifetime of alloys, particularly in the case of aluminum piston alloys, researchers are compelled to predict their fatigue lifetime behaviors and explore alternative innovative methods for estimation against destructive testing (10).

Matin and Azadi (11) conducted a Shapley Additive Explanations (SHAP) value-based analysis with XGBoost, employing five different machine learning models to predict the fatigue lifetime of aluminum alloys. Among those models, XGBoost emerged as the most effective, demonstrating compatibility with their dataset when trained on all available data.

This study motivates the researchers to find the best

model for predicting the fatigue lifetime of aluminum alloys based on testing data and preprocessing methods. Furthermore, because of the potential of this work on the possibility of increasing fatigue lifetime, the literature provides the following paragraph for discussion: firstly, the normalization and scaling approaches, as preprocessing steps, transform raw data into a standardized format. For this aim, several researches have been conducted to suggest normalization methods, such as Min-Max normalization (12), Manhattan normalization (MN) (13), Euclidean normalization (EN) (14), and maximum absolute normalization (MAN) (15). Furthermore, the variability in fatigue properties arises from different sources, such as experimental conditions, material properties, and testing equipment. As a result, normalizing fatigue and its characteristics may provide valuable approaches for improving fatigue modeling (16, 17). Second, numerous studies show that the amount of training data and the ratio of training to testing data affect the accuracy of machine learning models. Additionally, these factors play a significant role in improving machine learning prediction modeling. Medar et al. (18) conducted six trials using various training ratios (ranging from 2/12 to 12/12) to calculate the mean absolute error. Ramezan et al. (19) studied the impacts of different quantities of training data on diverse supervised machine-learning classification methods.

The present study demonstrates the impact of normalization approaches on enhancing the fatigue lifetime prediction of AlSi12CuNiMg aluminum alloy, in the engine piston application. The prediction relies on stress levels in the rotary bending fatigue tests, heat treatment conditions of manufactured standard test specimens, the fretting (wear) force during the test, the corrosion time subjected to test specimens, and the existence of the lubrication during the test, all serving as inputs for the XGBoost machine learning model.

The novelty of this work lies in finding the best modeling process for estimating the fatigue lifetime in different training and testing sets through data splitting by incorporating various normalization approaches. The main application of this work is estimating the fatigue lifetime of aluminum alloys with highly accurate methods under multiple inputs, which traditional methods such as S-N curves do not provide. Moreover, by examination of test data on estimation of the fatigue lifetime, the proficiency of this method is demonstrated in reducing destructive tests like rotary bending fatigue tests for analysis.

2. RESEARCH METHODS

2.1. Experimental Dataset

This segment is based on the experimental dataset described by Azadi and Parast (8). The goal was to evaluate how different ISO

1143 standard specimens perform in corrosion fatigue, pure fatigue, and fretting fatigue tests with varying inputs in this work (8). Furthermore, these specimens were manufactured using a commercially known alloy called AlSi12CuNiMg, which has widespread application in the automotive manufacturing industry. Each sample in this dataset contains six distinct variables that serve as features (all of the variables were considered integers). The variables included "stress" ranging from 90 MPa to 120 MPa; "fretting force" with values ranging from 0 N to 20 N; the presence of "lubrication" denoted by a value of 0 for non-existing and 1 for existing; the percentage of "nano-particles" in the manufacturing of specimens, ranging from 0 to 1%; the corrosion time of manufactured specimens in H₂SO₄, ranging from 0 hours to 200 hours; and the existence of T6 heat-treatment, with a value of 0 for non-existing and 1 for existing heat-treatment. Moreover, the target is the fatigue lifetime, ranging from 500 cycles to 1,398,100 cycles.

2. 2. Modeling Techniques

XGBoost is an expandable tree-boosting technique, offering significant potency and speed for machine learning tasks. Equation 1 denotes the objective involving minimizing regularization (20).

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{j=1}^K \Omega(f_j) \tag{1}$$

where $l(y_i, \hat{y}_i)$ represents the difference between the estimated and actual values, $\Omega(f_j)$ and K are identified as regularization factors and the cumulative count of trees, respectively.

Various normalization and scaling methods are utilized to define a specific data range to enhance the performance and accuracy of the machine learning model. Some of these methods are as follows (21, 22):

MAN: In this normalization approach, X_n is the normalized variable, X represents the unprocessed variable, and $|X|_1$ represents the Manhattan norm. Equation 2 exemplifies this technique (22).

$$X_n = \frac{X}{|X|_1} \tag{2}$$

$$|X|_1 = \sqrt{X_1^2 + X_2^2 + \dots + X_n^2}$$

EN: Within this normalization strategy, X_n illustrates the normalized variable, X represents the unprocessed variable, and $|X|_2$ represents the Euclidean norm. Equation 3 exemplifies this technique (22).

$$X_n = \frac{X}{|X|_2} \tag{3}$$

$$|X|_2 = |X_1| + |X_2| + |X_3| + \dots + |X_n|$$

MAN: In this method, X_n is the normalized variable, X represents the unprocessed variable, and $MAX(|X|)$ indicates the maximum absolute value of that particular variable among all the samples. Equation 4 illustrates this method (21).

$$X_n = \frac{X}{Max(|X|)} \tag{4}$$

Modified min-max normalization (MMN): This method is closely based on Min-Max normalization. The only difference is that when calculating the logarithm value within this normalization approach, it becomes necessary to eliminate zero values from the normalized variables. Equation 5 defines this method with the following equation. Moreover, X_n represents the normalized variable, X is the unprocessed variable, X_{min} represents the minimum value among all unprocessed variables, and X_{max} denotes the maximum value among all unprocessed variables (21).

$$X_n = \frac{X - X_{min}}{X_{Max} - X_{min}} \text{ and } X_n \neq 0 \tag{5}$$

The performance and accuracy of estimation, regression, and machine learning modeling can be evaluated using metrics like R^2 and $RMSE$. Additionally, the SB offers a valuable way to illustrate a factor that covers the complete range of actual lifetimes versus the estimated fatigue lifetime (23, 24). The R^2 metric, ranging from 0 to 1, is commonly employed to assess the efficacy of a machine learning approach. A higher R^2 indicates better prediction of the target variable. Similarly, SB plots represent the actual and predicted values on the axes on a logarithmic scale. The line with the equation of $y=x$ signifies a high accuracy in modeling (experimental and predicted values), and SB values are the slopes of the lines that enclose the data points. A lower SB indicates high accuracy and less scattering of the data from the $y=x$ line (23, 24). Figure 1 provides a detailed representation of the SB plot, as depicted.

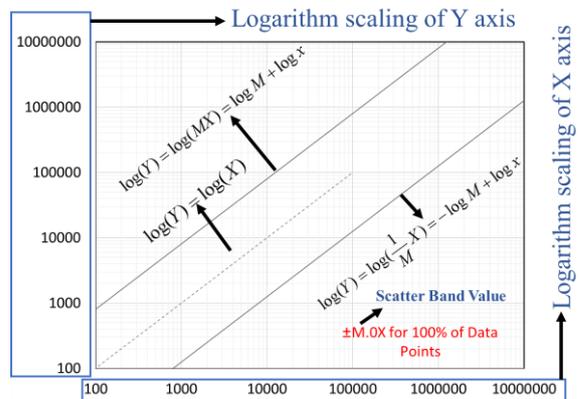


Figure 1. The detailed depiction of the SB plot

3. RESULTS AND DISCUSSION

Table 1 presents the comparative results of diverse normalization methods, employing XGBoost as the machine learning algorithm. It displays R^2 and $RMSE$ values for normalized fatigue lifetime and its logarithm, using the entire dataset without splitting it into separate training and test data. Additionally, a study was conducted to predict the fatigue lifetime under various conditions using XGBoost and 100% of the data as training data (24). Their results represented the R^2 value of 97.66%, which closely aligned with the findings of the present study, where the R^2 value was 95.66%. Moreover, the results presented in this table demonstrate that MMN achieves the highest modeling accuracy among the other normalization techniques. However, it is worth noting that the fatigue lifetime without normalization was fitted remarkably well using XGBoost. Moreover, according to alternate researches, EN had a lower performance versus MMN, which, within this study, is also demonstrated to have a lower performance (25).

Singh and Singh (26) examined how normalization methods influenced 21 widely recognized datasets, including Iris, Australian, Breast Cancer, and others. They assessed their model's accuracy and found that, notably, in some datasets, normalization without feature selection and weighting yielded inadequate results.

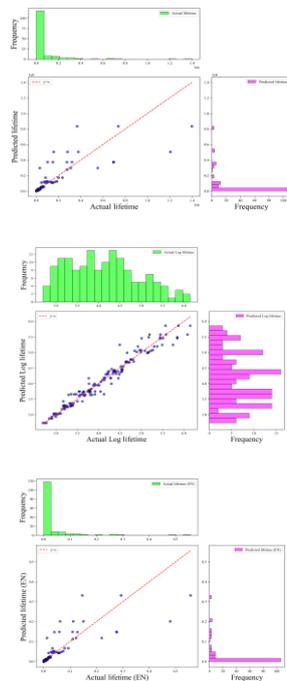
Notably, in some cases, this method resulted in lower accuracy. The study also revealed a fluctuating accuracy pattern across the 21 different datasets. This pattern highlighted that, in general, max absolute normalization (MAN), followed by min-max normalization, and unnormalized methods achieved higher accuracy, as indicated by their respective performances (26). Notably, the current study shows that when considering 100% of the training data, modified MMN outperforms both unnormalized and MAN methods.

Figure 2 illustrates the histograms, scatter plots, and SB for preprocessed fatigue lifetime data and its estimation.

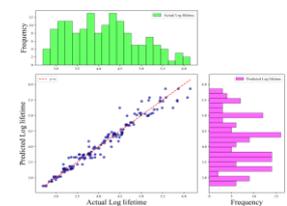
TABLE 1. The values of R^2 and $RMSE$ for different normalization methods using 100% of data as training data

Normalization method	Lifetime		Logarithm of lifetime	
	R^2	$RMSE$	R^2	$RMSE$
No Normalization	68.10	108624.986	95.66	0.170
MN	67.97	0.009	95.66	0.170
EN	68.09	0.043	95.66	0.170
MMN	68.10	0.077	95.99	0.182
MAN	68.10	0.077	95.66	0.170

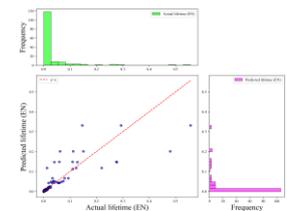
Note: The bold values signify the superior accomplishments



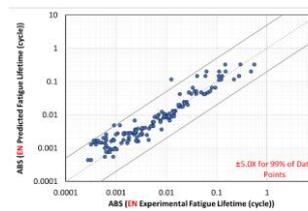
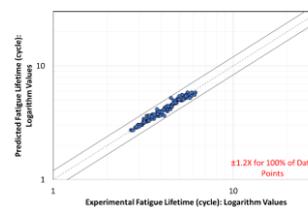
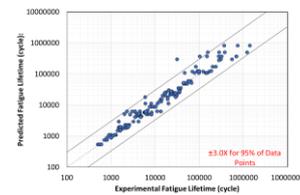
(a)



(b)



(c)



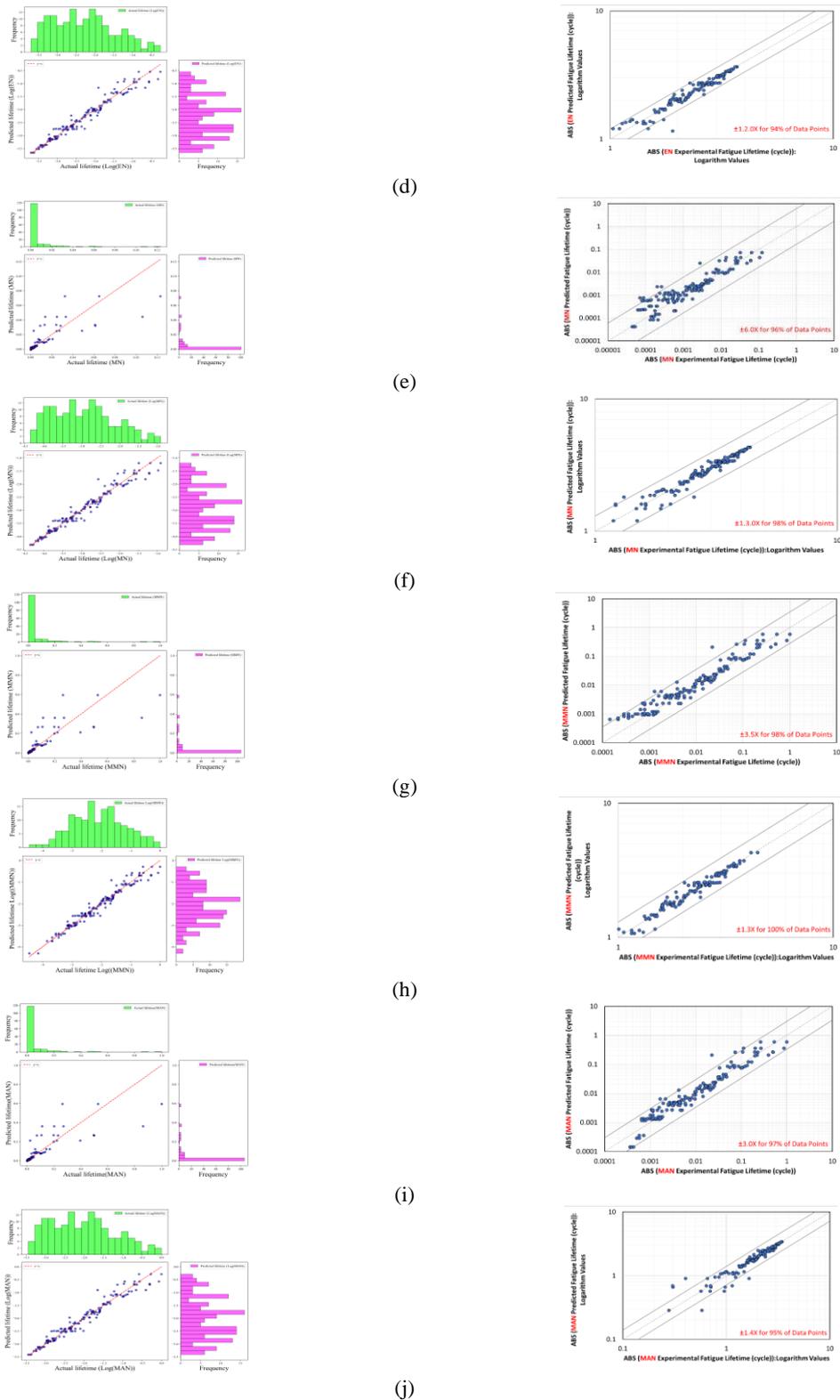


Figure 2. The histogram, scatter plot, and SB for preprocessing fatigue lifetime values and estimating them with 100% of the training data for different models: (a) No normalization of fatigue lifetime modeling, (b) logarithm value of fatigue lifetime modeling, (c) EN of fatigue lifetime modeling, (d) logarithm value of EN fatigue lifetime modeling, (e) MN of fatigue lifetime modeling, (f) logarithm value of MN fatigue lifetime modeling, (g) MMN of fatigue lifetime modeling, (h) logarithm value of MMN fatigue lifetime modeling, (i) MAN of fatigue lifetime modeling, and (j) logarithm value of MAN fatigue lifetime modeling

TABLE 2. The mean values of R^2 and $RMSE$ for different normalization methods using 80-20% of the data as random training and testing

Normalization method	Lifetime		Logarithm of lifetime	
	Mean R^2	Mean $RMSE$	Mean R^2	Mean $RMSE$
No Normalization	13.95	176235.565	88.47	0.279
MN	19.37	0.014	88.67	0.276
EN	18.18	0.068	88.63	0.277
MMN	18.10	0.122	88.72	0.300
MAN	18.11	0.122	88.61	0.277

Note: The bold values signify the superior accomplishments.

Each plot on the left and right sides of the figure corresponds to a specific preprocessing method using 100% of the data for training. The figure presents the distribution of fatigue lifetime data following each preprocessing approach and its associated estimation. Moreover, a study depicts the SB plot for this experimental dataset, using 100% of the data as training data to predict the logarithmic values of fatigue lifetime, which range between 2.69897 and 6.14554 (8). This representation is included in Figure 2(b) of this work.

Several methods exist for splitting data into training and test sets. Surono et al. (27) employed a convolutional neural network to achieve an 80%-20% ratio for training and testing in various machine learning methods to estimate lung disease. In a different study, Kurdthongmee (28) investigated the impact of varying the number of training data points from 100 to 250 to determine the optimal amount for estimating parawood pith. For a more specific focus on fatigue lifetime estimation, Choi (9) reported a 30%-70% test and train ratio using XGBoost for estimating steel fatigue lifetime based on stress. The achieved R^2 for testing fatigue lifetime was 98.03%. He et al. (29) used stress and fatigue lifetime datasets in their work. They explored the effects of three different test-train ratios on estimating the fatigue lifetime of three commercial steels, concluding that a 10%-90% training ratio with artificial neural networks and random forest yielded the best results.

Table 2 in this study compares the results of various normalization methods using XGBoost as the machine learning algorithm. It calculates the average R^2 and $RMSE$ over 20 iterations after randomly splitting the data into 80% training and 20% testing sets. The results demonstrate that, unlike when using 100% of the training data, when using 80% of the data for training and calculating mean R^2 for test data, all of the normalization methods improved their accuracy in modeling both the preprocessed fatigue lifetime and its logarithm.

This section of the study illustrates that when using the same algorithms and an equal number of training samples, the accuracy varies between repetitions that use different training samples, which is consistent with the other study (19). Furthermore, a comparison of the results in Tables 1 and 2 reveals that the number of training data samples has a significant impact on the metrics values. This observation is consistent with the findings of other research studies (18, 30).

Figure 3 illustrates histograms, scatter plots, and SB for preprocessed fatigue lifetime data and its estimation. Each plot on the left and right sides of the figure corresponds to a specific preprocessing method. These methods trained on 80% of the data and generated plots for a specific random state, which was consistent across all modeling. Moreover, the graphs relate to the test data. The figure presents the distribution of fatigue lifetime data following each preprocessing approach and its associated estimation. Furthermore, it illustrates a subset of five logarithmic normalization method approaches among all preprocessing techniques, as the remaining techniques lack satisfactory accuracy.

According to SB plots in Figures 2 and 3, coverages of the data by SB lines are different. These SB values in these figures correspond to optimized positions of SB lines for high data coverage with a low value of SB. Therefore, Table 3 represents SB values for different data coverages of SB lines to compare SB values effortlessly. Moreover, it indicates that SB values for fatigue lifetime and its logarithm, with no requirement for preprocessing techniques, are the best achievements.

A separate study conducted experiments using 27 different specimens subjected to varying stress levels (31). To compare the SB values presented in Table 3 for estimating the fatigue lifetime in this study with those from the other research (31). Those specimens were utilized to develop an experimental equation for the fatigue lifetime prediction, and the SB value was reported with a margin of ± 2 for their modeling of the fatigue lifetime in cast iron crankshafts (31). In contrast, the present study used the entire dataset for training and obtained an SB value of ± 10 , implying that its modeling may not be as accurate as other research. Notably, the current dataset is five times larger than the previous one. Even after accounting for 90% coverage of SB lines, the SB value remained constant at ± 2 , demonstrating the strength and reliability of the proposed approach even when applied to over 100 data points.

The estimated fatigue lifetimes in low-cycle fatigue were not valid when the logarithmic transformer was not applied to the fatigue lifetime and its normalized values in this study. The estimated fatigue lifetimes of aluminum alloys deviated significantly from the physics of the lifetime, resulting in some negative estimated lifetimes. However, using a logarithmic transformer

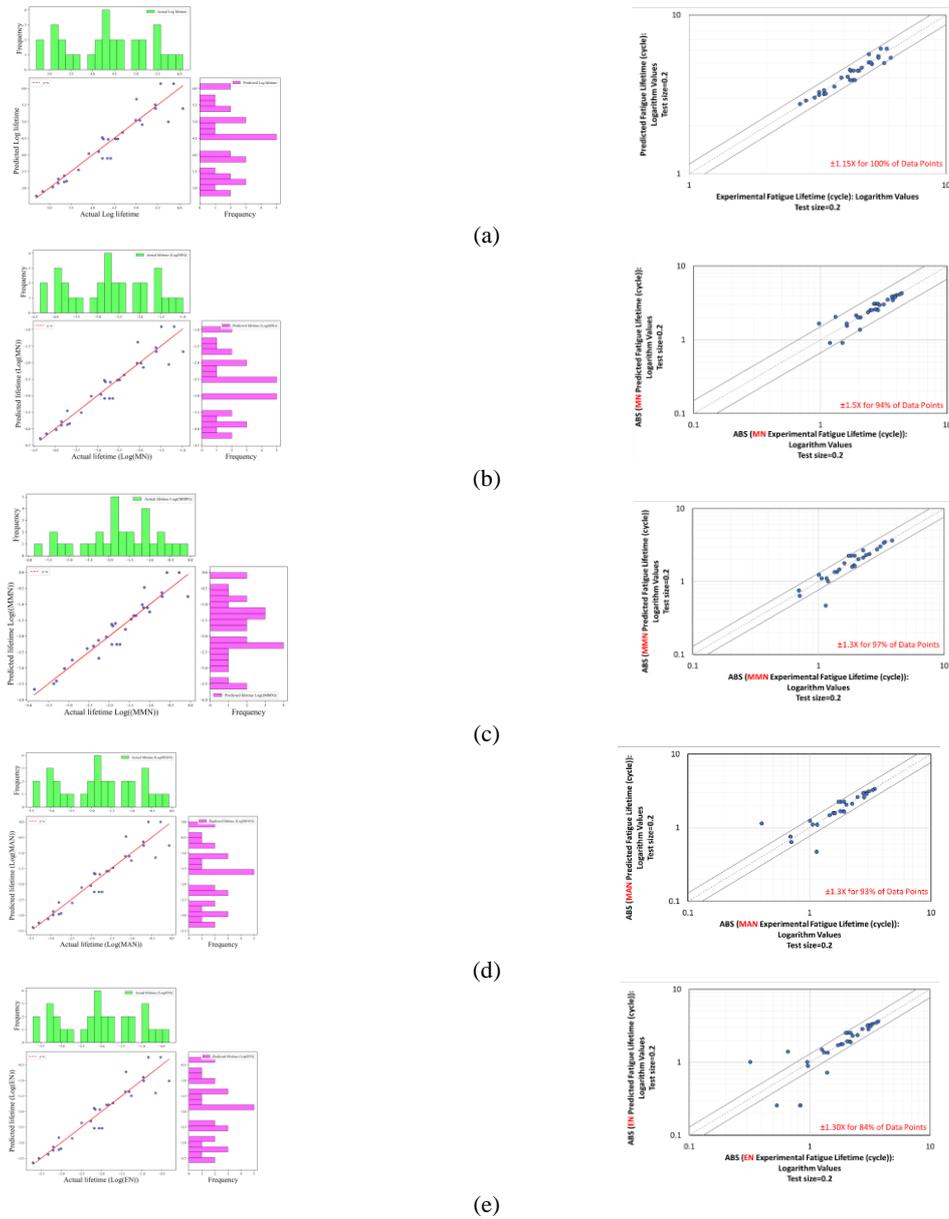


Figure 3. The histogram, scatter plot, and *SB* for same random state testing preprocessing fatigue lifetime values and estimating them with 80-20% of the data as training and testing for different models: (a) Logarithm value of fatigue lifetime modeling, (b) logarithm value of MN fatigue lifetime modeling, (c) logarithm value of MMN fatigue lifetime modeling, (d) logarithm value of MAN fatigue lifetime modeling, and (e) Logarithm value of EN fatigue lifetime modeling

proved to be a wise decision since it effectively controlled the pattern of the fatigue lifetime, transitioning from low-cycle fatigue lifetimes to high-cycle fatigue lifetimes. Furthermore, no physical laws or empirical equations related to the fatigue lifetime were used to obtain these predicted values.

A major challenge in training or fine-tuning machine learning models is calculating the number of observations required for the optimum performance. Although having more training observations leads to

better model performance, in theory, the procedure of gathering more data is typically time-consuming, expensive, or even impossible (32). Similarly, in the rotary bending fatigue tests, the preparation of specimens incurs costs for the manufacturer, and the tests themselves are both destructive and time-consuming. In such cases, learning curves can be used to determine whether the number of samples used is enough. Figure 4 represents learning curves for fatigue lifetime modeling and its logarithm value modeling. The training R^2 lines

TABLE 3. SB values for all methods showcasing different data coverage of SB lines

Method	Test size (%)	Number of scatter data	SB			
			100% data	95% data	90% data	85% data
Life time	0	147	10.0	3.0	2.0	1.8
Log lifetime	0	147	1.2	1.1	1.1	1.1
MAN	0	147	9.0	2.8	2.2	1.9
Log MAN	0	147	1.3	1.2	1.1	1.1
MMN	0	146	9.0	3.4	3.0	2.5
Log MMN	0	146	1.3	1.2	1.1	1.1
EN	0	147	9.0	4.3	3.5	2.9
Log EN	0	147	1.6	1.2	1.1	1.1
MN	0	147	12.0	6.0	4.0	3.5
Log MN	0	147	1.4	1.2	1.1	1.1
Log lifetime	20	30	1.1	1.1	1.1	1.1
Log MAN	20	30	2.8	1.4	1.3	1.3
Log MMN	20	30	2.4	1.3	1.2	1.2
Log EN	20	30	3.2	2.1	2.0	1.3
Log MN	20	30	1.7	1.5	1.4	1.3

are roughly equal to 1, indicating that the models are catching the patterns in the training data.

Additionally, the R^2 values for cross-validation (CV) are approaching convergence with the R^2 values from the training set, as depicted in the learning curves. Generally, one of the reasons for the proximity of the training R^2 line to 1 is the overfitting of the model. However, the increasing number of training data indicates that the model requires more samples to achieve better performance and accuracy (33, 34). In this study, the accuracy of the test data, especially for logarithmic modeling, was notable and the models did not demonstrate insensitivity to small changes in the training data. Therefore, based on Figure 4, to achieve a better performance in predicting outcomes, surpassing the representations of the models presented in this study, it is recommended to expand the dataset to include additional samples.

4. CONCLUSIONS

This study focused on using various normalization methods for fatigue lifetime and its logarithm as preprocessing tasks, employing extreme gradient boosting (XGBoost) for prediction. The results included scatter bands, metrics (R^2 and $RMSE$), histograms, and scatter plots for 100% and 80% of the data employed as training. The predicted values of preprocessed fatigue lifetimes in various XGBoost models depend on specific conditions from the rotary bending fatigue tests and standard manufactured specimens used in the same test. The subsequent results were generated through this process:

- The best scatter band (SB) achievement for all normalization approaches, using 100% of the data for training data modeling and a random 20% of data for testing data modeling, was the logarithm of fatigue lifetime with SB values of ± 1.2 and ± 1.1 , respectively. Therefore, it demonstrates the most accurate model among the other models in estimating.
- When using 100% of the data as training, all techniques of the normalization and unnormalized approaches had nearly identical values (with a difference of less than 0.5%). The SB values, on the other hand, varied. They ranged from ± 9.0 to ± 12.0 for non-logarithmic models and from ± 1.2 to ± 1.6 for logarithmic models. Therefore, it demonstrates the normalization method had no significant impact on 100% training ratio modelling.
- Calculating the mean values of R^2 for 80% of random training modeling illustrates that non-logarithmic normalization methods exhibited better accuracy (at least 4% higher than unnormalized fatigue lifetime modeling), whereas, for logarithmic

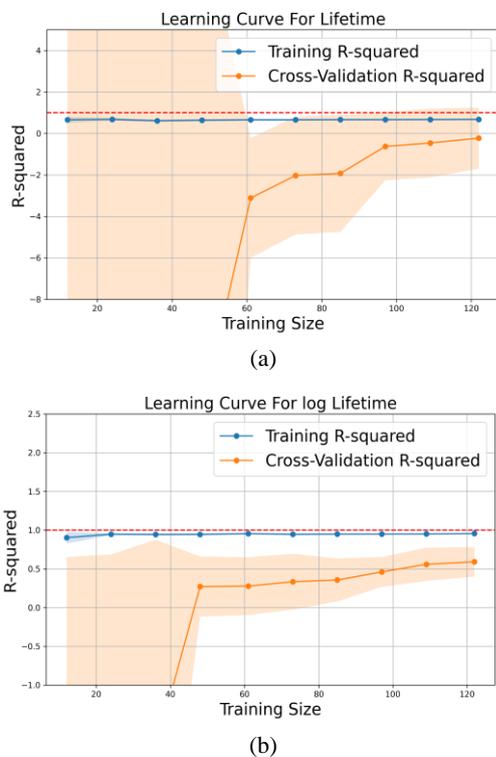


Figure 4. Learning curves for (a) fatigue lifetime modeling and (b) logarithm of fatigue lifetime modeling

modeling, the values were closely aligned (ranging from 88.47% to 88.72%).

- Learning curves show that as the amount of training data increases, the prediction of testing data improves. As a result, additional samples should be added to the dataset to improve the accuracy of all models.

The primary limitation of the current work occurs when the logarithmic transformer is not applied to the normalized fatigue lifetime and the fatigue lifetime without normalization in XGBoost modeling. The predicted values for the fatigue lifetime and its normalization differed from the physics of the fatigue lifetime, particularly for low-cycle fatigue lifetimes, with an incorrect negative estimation. However, the logarithmic models were accurate, and the predictions agreed with the physics of the fatigue lifetime. Furthermore, for the future research aimed at improving fatigue lifetime modeling, it is suggested that an activation function be used in the output layer of a neural network. Additionally, investigating the deep learning frameworks that allow users to directly impose constraints on the layer may be beneficial, particularly for datasets with binary and categorical variables. This method is especially useful for datasets with few target values, such as the one used in this study to represent aluminum alloys. For evaluating the fatigue lifetimes of aluminum alloys in the automotive industry, these methods provide highly accurate alternatives to destructive tests such as rotary bending fatigue tests.

5. REFERENCES

1. Farrahi G, Faghidian S, Smith D. Reconstruction of residual stresses in autofrettaged thick-walled tubes from limited measurements. *International Journal of Pressure Vessels and Piping*. 2009;86(11):777-84. 10.1016/j.ijpvp.2009.03.010
2. Ismaiel A. Wind turbine blade dynamics simulation under the effect of atmospheric turbulence. *Emerging Science Journal*. 2023;7(1):162-76. 10.28991/ESJ-2023-07-01-012
3. Ali Faghidian S. Analytical approach for inverse reconstruction of eigenstrains and residual stresses in autofrettaged spherical pressure vessels. *Journal of Pressure Vessel Technology*. 2017;139(4):041202. 10.1115/1.4035980
4. Farrahi G, Faghidian S, Smith D. An inverse method for reconstruction of the residual stress field in welded plates. 2010. 10.1115/1.4001268
5. Azadi M, Farrahi G, Winter G, Eichseder W. The effect of various parameters on out-of-phase thermo-mechanical fatigue lifetime of A356.0 cast aluminum alloy. *International Journal of Engineering, Transactions C: Aspects*. 2013;26(12):1461-70. 10.5829/idosi.ije.2013.26.12c.06
6. Azadi M, Rezanezhad S, Zolfaghari M. Effects of various ageing heat treatments on microstructural features and hardness of piston aluminum alloy. *International Journal of Engineering, Transactions A: Basics*. 2019;32(1):92-8. 10.5829/ije.2019.32.01a.12
7. Akhtar M, Qamar SZ, Muhammad M, Nadeem A. Optimum heat treatment of aluminum alloy used in manufacturing of automotive piston components. *Materials and Manufacturing Processes*. 2018;33(16):1874-80. 10.1080/10426914.2018.1512128
8. Azadi M, Parast MSA. Data analysis of high-cycle fatigue testing on piston aluminum-silicon alloys under various conditions: Wear, lubrication, corrosion, nano-particles, heat-treating, and stress. *Data in brief*. 2022;41:107984. 10.1016/j.dib.2022.107984
9. Choi D-K. Data-driven materials modeling with XGBoost algorithm and statistical inference analysis for prediction of fatigue strength of steels. *International Journal of Precision Engineering and Manufacturing*. 2019;20:129-38. 10.1007/s12541-019-00048-6
10. Matin M, Azadi M. A Novel Machine Learning-Based Model for Predicting of Transient Fatigue Lifetime in Piston Aluminum Alloys. Available at SSRN 4598611. 10.2139/ssrn.4598611
11. Matin M, Azadi M. Machine learning-based modeling for estimating bending fatigue lifetimes in AlSi12CuNiMg aluminum alloy of engine pistons under different inputs: Fretting force, corrosion time, lubrication, heat-treating, nano-particles, and stress. *Corrosion Time, Lubrication, Heat-Treating, Nano-Particles, and Stress*. 2023. 10.2139/ssrn.4549376
12. Munkhdalai L, Munkhdalai T, Park KH, Lee HG, Li M, Ryu KH. Mixture of activation functions with extended min-max normalization for forex market prediction. *IEEE Access*. 2019;7:183680-91. 10.1109/ACCESS.2019.2959789
13. Zhou W, Liu A, Wu L, Chen X. A L1 normalization enhanced dynamic window method for SSVEP-based BCIs. *Journal of Neuroscience Methods*. 2022;380:109688. 10.1016/j.jneumeth.2022.109688
14. Dai Z, Chen W, Huang X, Li B, Zhu L, He L, et al., editors. Cnn descriptor improvement based on l2-normalization and feature pooling for patch classification. 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO); 2018: IEEE. 10.1109/ROBIO.2018.8665330
15. Gómez-Escalonilla V, Martínez-Santos P, Martín-Loeches M. Preprocessing approaches in machine-learning-based groundwater potential mapping: an application to the Koulikoro and Bamako regions, Mali. *Hydrology and Earth System Sciences*. 2022;26(2):221-43. 10.5194/hess-26-221-2022
16. Lv S, Liu C, Chen D, Zheng J, You Z, You L. Normalization of fatigue characteristics for asphalt mixtures under different stress states. *Construction and Building Materials*. 2018;177:33-42. 10.1016/j.conbuildmat.2018.05.109
17. Lv S, Wang P, Fan X, Cabrera MB, Hu L, Peng X, et al. Normalized comparative study on fatigue characteristics of different pavement materials. *Construction and Building Materials*. 2021;271:121907. 10.1016/j.conbuildmat.2020.121907
18. Medar R, Rajpurohit VS, Rashmi B, editors. Impact of training and testing data splits on accuracy of time series forecasting in machine learning. 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA); 2017: IEEE. 10.1109/ICCUBEA.2017.8463779
19. Ramezan CA, Warner TA, Maxwell AE, Price BS. Effects of training set size on supervised machine-learning land-cover classification of large-area high-resolution remotely sensed data. *Remote Sensing*. 2021;13(3):368. 10.3390/rs13030368
20. Meng Y, Yang N, Qian Z, Zhang G. What makes an online review more helpful: an interpretation framework using XGBoost and SHAP values. *Journal of Theoretical and Applied Electronic Commerce Research*. 2020;16(3):466-90. 10.3390/jtaer16030029
21. Md AQ, Kulkarni S, Joshua CJ, Vaichole T, Mohan S, Iwendi C. Enhanced preprocessing approach using ensemble machine learning algorithms for detecting liver disease. *Biomedicine*. 2023;11(2):581. 10.3390/biomedicine11020581

22. Chiu W-Y, Chen B-S. Mobile location estimation in urban areas using mixed Manhattan/Euclidean norm and convex optimization. *IEEE transactions on Wireless Communications*. 2009;8(1):414-23. 10.1109/T-WC.2009.080156
23. Azadi M, Shahsavand A, Parast MSA. Analyzing experimental data from reciprocating wear testing on piston aluminum alloys, with and without clay nano-particle reinforcement. *Data in Brief*. 2022;45:108766. 10.1016/j.dib.2022.108766
24. Nasiri H, Azadi M, Dadashi A. Interpretable extreme gradient boosting machine learning model for fatigue lifetimes in 3D-printed polylactic acid biomaterials. Available at SSRN 4364418. 2023. 10.2139/ssrn.4364418
25. Eesa AS, Arabo WK. A normalization methods for backpropagation: a comparative study. *Science Journal of University of Zakho*. 2017;5(4):319-23. 10.25271/2017.5.4.381
26. Singh D, Singh B. Investigating the impact of data normalization on classification performance. *Applied Soft Computing*. 2020;97:105524. 10.1016/j.asoc.2019.105524
27. Surono S, Afitian MYF, Setyawan A, Arofah DKE, Thobirin A. Comparison of CNN Classification Model using Machine Learning with Bayesian Optimizer. *HighTech and Innovation Journal*. 2023;4(3):531-42. 10.28991/HIJ-2023-04-03-05
28. Kurdthongmee W. Comprehensive Evaluation of Deep Neural Network Architectures for Parawood Pith Estimation. *HighTech and Innovation Journal*. 2023;4(3):543-59. 10.28991/HIJ-2023-04-03-06
29. He L, Wang Z, Akebono H, Sugeta A. Machine learning-based predictions of fatigue life and fatigue limit for steels. *Journal of Materials Science & Technology*. 2021;90:9-19. 10.1016/j.jmst.2021.02.021
30. Althnian A, AlSaeed D, Al-Baity H, Samha A, Dris AB, Alzakari N, et al. Impact of dataset size on classification performance: an empirical evaluation in the medical domain. *Applied Sciences*. 2021;11(2):796. 10.3390/app11020796
31. Khameneh MJ, Azadi M. Evaluation of high-cycle bending fatigue and fracture behaviors in EN-GJS700-2 ductile cast iron of crankshafts. *Engineering Failure Analysis*. 2018;85:189-200. 10.1016/j.engfailanal.2017.12.017
32. Cruz F, Castelli M. Learning Curves Prediction for a Transformers-Based Model. Available at SSRN 4305463. 2023. 10.28991/ESJ-2023-07-05-03
33. Zhang G, Shi Y, Yin P, Liu F, Fang Y, Li X, et al. A machine learning model based on ultrasound image features to assess the risk of sentinel lymph node metastasis in breast cancer patients: Applications of scikit-learn and SHAP. *Frontiers in Oncology*. 2022;12:944569. 10.3389/fonc.2022.944569
34. Giola C, Danti P, Magnani S. Learning curves: A novel approach for robustness improvement of load forecasting. *Engineering Proceedings*. 2021;5(1):38. 10.3390/engproc2021005038

COPYRIGHTS

©2024 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

ارزیابی برآورد طول عمر خستگی برای آلیاژهای آلومینیوم، به ویژه در صنعت خودرو، امری حیاتی است. این مقاله تاثیر روش‌های عادی سازی مختلف را بر عملکرد تخمین طول عمر خستگی با استفاده از تقویت گرادیان شدید (XGBoost) به عنوان یک روش یادگیری ماشین نظارت شده بررسی می‌کند. به این منظور در این مطالعه مجموعه داده ای شامل ورودی های فیزیکی و تجربی متنوع مربوط به آلیاژ آلومینیوم در کنار عمر خسته به عنوان خروجی استفاده شده است. علاوه بر این، قبل از برازش مدل XGBoost، روش‌های مختلف پیش پردازش طول عمر خستگی اعمال می‌شوند و سپس با استفاده از معیارهایی مانند خطای میانگین مربعات ریشه (RMSE)، ضریب تعیین (R^2)، و باند پراکندگی (SB) ارزیابی شده اند. نتایج به دست آمده نشان می‌دهد که مدل سازی طول عمر خستگی با مقادیر لگاریتمی به عنوان یک روش پیش پردازش زمانی که XGBoost با ۱۰۰ درصد داده‌ها آموزش داده می‌شود، برتری دارد. با این حال، سایر روش‌های نرمال سازی دقت بالاتری را در تخمین داده‌های آزمایش با تقسیم ۲۰٪ تست و ۸۰٪ مجموعه داده یادگیری نشان می‌دهند.



Enhancing Wind Power Conversion System Control Under Wind Constraints Using Single Hidden Layer Neural Network

A. Mazari^{*a}, H. Ait Abbas^b, K. Laroussi^a, B. Naceri^c

^a Laboratory of Applied and Automation and Industrial Diagnostic (LAADI), University of Djelfa, Djelfa, Algeria

^b Laboratory of Electrical and Automatic Systems Engineering (LGSEA), University of Bouira, Bouira, Algeria

^c Laboratory of Identification, Commande, Control and Communication (LI3CUB), University of Biskra, Algeria

PAPER INFO

Paper history:

Received 27 November 2023

Received in revised form 25 December 2023

Accepted 10 January 2024

Keywords:

Wind Power Generation System

Cascaded Doubly Fed Induction Generator

Proportional Integral Derivative

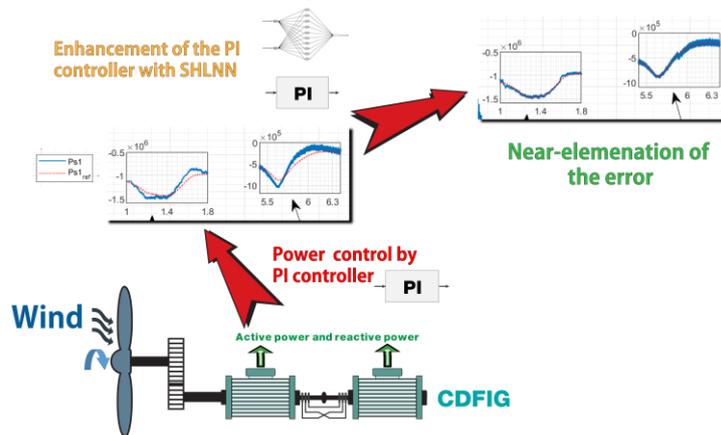
Single Hidden Layer Neural Network

ABSTRACT

In the realm of wind power generation, cascaded doubly fed induction generators (CDFIG) play a pivotal role. However, the classical proportional integral derivative (PID) controllers used within such systems often struggle with instability and inaccuracies arising from wind variability. This study proposes an enhancement to overcome these limitations by incorporating a single hidden layer neural network (SHLNN) into the wind power conversion systems (WPCS). The SHLNN aims to complement the PID controller by addressing its shortcomings in handling nonlinearities and uncertainties. This integration exploits the adaptive nature and low computational demand of SHLNNs, utilizing historical wind speed and power data to form a more resilient control strategy. Through Matlab/Simulink simulations, this approach is rigorously compared against traditional PID control methods. The results demonstrate a marked improvement in performance, highlighting the SHLNN's capacity to contend with the intrinsic variabilities of wind patterns. This contribution is significant as it offers a sophisticated yet computationally efficient solution to enhance CDFIG-based WPCS, ensuring more stable and accurate energy production.

doi: 10.5829/ije.2024.37.07a.10

Graphical Abstract



NOMENCLATURE

C_p	Power coefficient	w_{ij}	Weights between the second and the third layer
F_{dq}, k_{dq}	Adaptation gains	x_k	Input of neural network

* Corresponding Author Email: ali.mazari@univ-djelfa.dz (A. Mazari)

Please cite this article as: Mazari A, Ait Abbas H, Laroussi K, Naceri B. Enhancing Wind Power Conversion System Control Under Wind Constraints Using Single Hidden Layer Neural Network. International Journal of Engineering, Transactions A: Basics. 2024;37(07):1306-16.

L_{m1}, L_{m2}	Mutual inductances [H]	y_{nn}	Output of neural network
L_{r1}, L_{r2}	Inductances of the first and the second rotor [H]	μ_d, μ_q	Neural network inputs in d-q axis
L_{s1}, L_{s2}	Inductances of the first and the second stator [H]	G	Gearbox's gain
P_{s1}	Active power of the first generator [W]	C	Constante
P_t	Aerodynamic power [W]	I	Unit matrix
Q_{s1}	Reactive power of the first generator [W]	J	Equation of jacobly
R_{s1}	First stator resistance [ohm]	V	Weights matrix for the hidden layer
R_{s2}	Second stator resistance [ohm]	W	Weights matrix
R_t	The blade length [m]	e	Error
T_{em}	Electromagnetic torque [N.m]	f	Friction of the wind turbine [N.m/(rad/s)]
T_g	Generator side torque [N.m]	j	Inertia of the wind turbine [kg.m ²]
T_t	Aerodynamic torque [N.m]	Greek Symbols	
\hat{V}	Estimate matrix for V	\tilde{i}_{dq}	Dynamic error
V_0	Primary value of V	Δ_d, Δ_q	Target terms in d-q axis
V_s	Grid voltage	Ω_{mec}	Mechanical speed of the generator [rad/s]
\hat{W}	Estimate matrix for the W	Ω_t	Rotational speed of the wind turbine [rad/s]
W_0	Primary value of W	τ_r	Response time [s]
W_{k+1}, W_k	Results of calculated weights by Gauss-Newton	$\varphi_{dr1}, \varphi_{qr1}$	First rotor fluxes in d-q axis [WB]
i_{dr}, i_{qr}	Rotor currents in d-q axis [A]	$\varphi_{dr2}, \varphi_{qr2}$	second rotor fluxes in d-q axis [WB]
i_{ds1}, i_{qs1}	First stator currents in d-q axis [A]	$\varphi_{ds1}, \varphi_{qs1}$	First stator fluxes in d-q axis [WB]
i_{ds2}, i_{qs2}	Second stator currents in d-q axis [A]	$\varphi_{ds2}, \varphi_{qs2}$	Second stator fluxes in d-q axis [WB]
k_i, k_p	PI regulator's gains	$\varphi_j(\cdot)$	Activation function
p_1, p_2	Number of poles of the first and the second generator	ω_{r1}, ω_{r2}	Angular frequency of the first and the second rotor
s_1, s_2	Slips of the first and the second generator	ω_{s1}, ω_{s2}	Angular frequency of the first and the second stator
u_Δ	Adaptive term	ϑ_{vj}	Biases between the first and the second layer
u_{ad}, u_{aq}	Adaptive terms in d-q axis	ϑ_{wt}	Biases between the second and the third layer
v_{ds1}, v_{qs1}	First stator voltages in d-q axis [V]	β	Pitch angle [°]
v_{ds2}, v_{qs2}	Second stator voltages in d-q axis [V]	λ	Velocity ratio
v_{jk}	Weights between the first and the second layer	μ	Positive coefficient
v_w	Wind velocity [m/s]	ρ	Air density [kg/m ³]

1. INTRODUCTION

As societies increasingly depend on electrical energy, there has been a notable rise in the utilization of renewable energy sources like wind power (1). This trend is primarily driven by the urgent need to lessen reliance on fossil fuels and mitigate the environmental impact caused by traditional energy sources (2).

Within the realm of wind power, one promising technology is the cascaded doubly fed induction generator (CDFIG) system (3). This innovative system combines two doubly fed induction generators (DFIGs) while preserving their original design (3, 4). It involves coupling two wound rotor induction generators at their rotor windings, both mounted on a shared shaft (3, 5, 6). Compared to traditional wind power systems, the CDFIG system offers several advantages, including enhanced reliability and reduced maintenance costs (7). One key benefit of employing CDFIG system is its ability to operate at variable speeds (4, 6). This is particularly valuable as it effectively stabilizes the generator voltage frequency when fluctuations occur in the rotor speed (6).

There has been a growing trend in recent years to integrate intelligent techniques in WPCS (8). Artificial

neural networks (ANNs) are widely recognized as a fundamental, powerful, and efficient approach for creating control strategies in situations where systems are only partially understood and have complex mathematical models (9). ANNs possess the capability to learn and approximate the dynamics of nonlinear systems by utilizing available input and output information (10). They are commonly applied in the fields of identification, adaptive control, and observation (11).

The literature encompasses various control techniques employed in wind power conversion systems (WPCS) utilizing the cascaded doubly fed induction generator (CDFIG). Notable examples include direct torque and flux control (12), sliding mode control (SMC) as described by Maafa et al. (7) and Zahedi et al. (13), This control technique frequently encounters undesired chattering phenomena (14). However, this phenomena can be decreased by updating the traditional SMC into the terminal SMC by Abdolhadi et al. (15) and the super-twisting SMC by Yan and Cheng (16) but it leads to more complexity of the controller (17). The predictive torque control of CDFIG highlighted by Bayhan et al. (18) which focuses on directly controlling the torque or flux, and the predictive voltage control also by Abdolrahimi

and Arab Khaburi (19). Dauksha and Iwanski (20) an indirect torque control algorithm for a CDFIG that allows it to work with an unbalanced power grid using different strategies. The development of an intelligent LAMDA-based controller for three-phase induction motors has shown improved precision and disturbance response by Morales et al. (21). Additionally, studies on the behavior of wind turbine blades in different turbulence levels have underscored the necessity for durable designs in renewable energy technologies (22). In finance, the integration of time-series analysis into neural network models has significantly enhanced stock price prediction accuracy (23). Our contribution aims to combine the strengths of the CDFIG with intelligent technique in the WPCS by integrating SHLNN which provides low computing complexity and good adaptation abilities to the control system.

This research identifies a gap in the existing literature regarding a control system that effectively manages these nonlinearities and uncertainties without introducing additional complexity. To address this gap, the article proposes a control strategy that integrates a single hidden layer neural network (SHLNN) with a PI controller. The SHLNN is designed to complement the PI controller by enhancing its adaptive capabilities while maintaining a low computational footprint.

The paper presents a detailed exposition of the mathematical model of the WPGS encompassing the CDFIG and the wind turbine models, the controller of the active power and reactive power of CDFIG designed including field-oriented control with PI controller in the first case then the implementation of the SHLNN alongside with a PI controller in the second case. Subsequently, applying the control methods in the WPGS and discussing the simulation results by comparing both cases. The paper concludes with a summary of the findings and their implications for future research in the wind power generation systems.

2. MATHEMATICAL MODEL OF THE WPCS

2.1. Model of The CDFIG The two doubly fed induction generators are connected mechanically and electrically, their stator windings are electrically isolated from each other, their rotors are connected to the same shaft, and the rotor circuits are linked with a reverse phase sequence (24). These generators are referred to the first generator as the control generator (CG) and the second as the power generator (PG) (24), as shown in Figure 1:

By employing a commonly used DFIG model and assuming symmetrical electrical circuits for both the first generator and the second generator across all phases (20), the rotor windings of the two generators are interconnected in opposite connection, this linkage gives

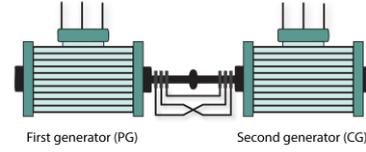


Figure 1. The cascaded doubly fed induction generator

the following model in d-q axis (13, 25):

$$v_{ds1} = R_{s1}i_{ds1} + L_{s1} \frac{di_{ds1}}{dt} + L_{m1} \frac{di_{dr}}{dt} - \omega_{s1}L_{s1}i_{qs1} - \omega_{s1}L_{m1}i_{qr} \quad (1)$$

$$v_{qs1} = R_{s1}i_{qs1} + L_{s1} \frac{di_{qs1}}{dt} + L_{m1} \frac{di_{qr}}{dt} + \omega_{s1}L_{s1}i_{ds1} + \omega_{s1}L_{m1}i_{dr} \quad (2)$$

$$0 = R_r i_{dr} + L_r \frac{di_{dr}}{dt} + L_{m1} \frac{di_{ds1}}{dt} - \omega_{g1}L_r i_{qr} - \omega_{g1}L_{m1}i_{qs1} - L_{m2} \frac{di_{ds2}}{dt} + \omega_{g1}L_{m2}i_{qs2} \quad (3)$$

$$0 = R_r i_{qr} + L_r \frac{di_{qr}}{dt} + L_{m1} \frac{di_{qs1}}{dt} + \omega_{g1}L_r i_{dr} + \omega_{g1}L_{m1}i_{ds1} - L_{m2} \frac{di_{qs2}}{dt} - \omega_{g1}L_{m2}i_{ds2} \quad (4)$$

$$v_{ds2} = R_{s2}i_{ds2} + L_{s2} \frac{di_{ds2}}{dt} - L_{m2} \frac{di_{dr}}{dt} - \omega_{s2}L_{s2}i_{qs2} + \omega_{s2}L_{m2}i_{qr} \quad (5)$$

$$v_{qs2} = R_{s2}i_{qs2} + L_{s2} \frac{di_{qs2}}{dt} - L_{m2} \frac{di_{qr}}{dt} + \omega_{s2}L_{s2}i_{ds2} - \omega_{s2}L_{m2}i_{dr} \quad (6)$$

The fluxes (25):

$$\varphi_{ds1} = L_{s1}i_{ds1} + L_{m1}i_{dr1} \quad (7)$$

$$\varphi_{qs1} = L_{s1}i_{qs1} + L_{m1}i_{qr1} \quad (8)$$

$$\varphi_{dr1} = L_{r1}i_{dr1} + L_{m1}i_{ds1} \quad (9)$$

$$\varphi_{qr1} = L_{r1}i_{qr1} + L_{m1}i_{qs1} \quad (10)$$

$$\varphi_{ds2} = L_{s2}i_{ds2} + L_{m2}i_{dr2} \quad (11)$$

$$\varphi_{qs2} = L_{s2}i_{qs2} + L_{m2}i_{qr2} \quad (12)$$

$$\varphi_{dr2} = L_{r2}i_{dr2} + L_{m2}i_{ds2} \quad (13)$$

$$\varphi_{qr2} = L_{r2}i_{qr2} + L_{m2}i_{qs2} \quad (14)$$

The frequency response of both stators and rotors is provided as follows (26):

$$\omega_{g1} = \omega_{g2} = \omega_{s1} - \omega_{r1} \quad (15)$$

$$\omega_{s2} = \omega_{s1} - \omega_{r1} - \omega_{r2} \quad (16)$$

$$\omega_{r1} = p_1\Omega \quad (17)$$

The total electromagnetic torque (T_{em}) of the cascade is determined by the addition of the torques generated by the two individual generators (25):

$$T_{em} = p_1 L_{m1} (i_{dr} i_{qs1} - i_{ds1} i_{qr}) + p_2 L_{m2} (i_{dr} i_{qs2} - i_{ds2} i_{qr}) \quad (18)$$

2. 2. Modeling of The Wind Turbine The wind turbine harnesses the kinetic energy of the wind, resulting in the rotation of its blades and the conversion of this energy into mechanical energy, which drives the rotor of the generator. The aerodynamic power (P_t) and torque (T_t) associated with this process are represented by the following expression (27):

$$P_t = \frac{1}{2} C_p(\lambda, \beta) \rho R_t \pi v_w^3 \quad (19)$$

$$T_t = \frac{P_t}{\Omega_t} \quad (20)$$

The power coefficient is given by Kelkoul and Boumediene (28):

$$C_p = (0.3 - 0.0167\beta) \sin\left(\frac{\pi(\lambda+0.1)}{10-0.3\beta}\right) - 0.00184(\lambda - 3)\beta \quad (21)$$

The gearbox serves to connect the generator and the wind turbine, amplifying the rotational speed transmitted from the wind turbine. The relationship between the rotational speed of the generator and that of the wind turbine is defined as follows (27, 28):

$$\Omega_t = \frac{\Omega_{mec}}{G} \quad (22)$$

The fundamental equation that governs the dynamics of speed and torques can be expressed as follows (27, 28):

$$j \frac{d\Omega_{mec}}{dt} + f\Omega_{mec} = T_g - T_{em} \quad (23)$$

3. CONTROL OF THE CDFIG

3. 1. Field Oriented Control It is a naturalistic separation between the magnitude associated with the flux (the excitation current) and that related to the torque (the armature current), which makes the delay of torque response decrease and increase the speed control rang. Hence, a higher efficiency over a long-term steady state load (9).

The initial stator flux φ_{s1} is aligned with the d-axis. When traversing along the q-axis, the flux of the first stator remains consistently at zero (5).

$$\varphi_{ds1} = \varphi_{s1} \quad (24)$$

$$\varphi_{qs1} = 0 \quad (25)$$

By considering the resistance of the first stator neglected, the voltages can be simplified as follows:

$$v_{ds1} = 0 \quad (26)$$

$$v_{qs1} = V_s = \omega_{s1} \varphi_{s1} \quad (27)$$

The voltages of the second stator in relation with the currents, the first stator currents in relation with those of the second stator. Hence, the active and reactive powers are described by Maafa et al. (25):

$$v_{ds2} = R_{s2} i_{ds2} + (L_{s2} - C \cdot L_{m2}) \frac{di_{ds2}}{dt} - s \cdot \omega_{s1} (L_{s2} - C \cdot L_{m2}) i_{qs2} \quad (28)$$

$$v_{qs2} = R_{s2} i_{qs2} + (L_{s2} - C \cdot L_{m2}) \frac{di_{qs2}}{dt} + s \cdot \omega_{s1} (L_{s2} - C \cdot L_{m2}) i_{ds2} + C \cdot s \frac{L_{m1} V_s}{L_{s1}} \quad (29)$$

$$i_{ds1} = \frac{V_s}{\omega_{s1} L_{s1}} \left(1 + \frac{C \cdot L_{m1}^2}{L_{s1} \cdot L_{m2}}\right) - C \cdot \frac{L_{m1}}{L_{s1}} i_{ds2} \quad (30)$$

$$i_{qs1} = -C \cdot \frac{L_{m1}}{L_{s1}} i_{qs2} \quad (31)$$

$$P_{s1} = -C \cdot V_s \frac{L_{m1}}{L_{s1}} i_{qs2} \quad (32)$$

$$Q_{s1} = \frac{V_s^2}{\omega_{s1} L_{s1}} \left(1 + \frac{C \cdot L_{m1}^2}{L_{s1} \cdot L_{m2}}\right) - C \cdot V_s \frac{L_{m1}}{L_{s1}} i_{ds2} \quad (33)$$

$$C = \frac{L_{m2}}{L_{r1} + L_{r2} - \frac{L_{m1}^2}{L_{s1}}} \quad (34)$$

$$s = s_1 \cdot s_2 = \frac{\omega_{s1} - \Omega(p_1 + p_2)}{\omega_{s1}} \quad (35)$$

The first stator's active and reactive powers (P_{s1} , Q_{s1}) will be the output signals from the CDFIG that will be controlled to achieve the tracking reliability of energy in the WPCS.

3. 2. PI Controller The speed control loop, PI regulator structure is illustrated in Figure 2. The design is based on the dynamic equation governing the behavior of the rotating components:

The open-loop transfer function is expressed as follows:

$$\frac{Y}{Y_{ref}} = \frac{P + \frac{k_i}{p} \frac{C \cdot L_{m1} V_s}{L_{s1} (L_{s2} - C \cdot L_{m2})}}{\frac{P}{k_p} \frac{R_{s2}}{P + (L_{s2} - C \cdot L_{m2})}} \quad (36)$$

Several approaches are used to calculate the parameters k_i and k_p (29), by using the poles compensation approach, the closed-loop transfer function is obtained

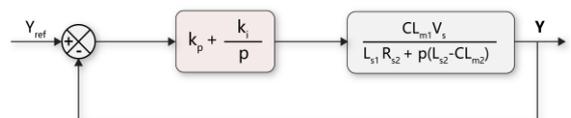


Figure 2. PI regulator structure

as follows:

$$\frac{Y}{Y_{ref}} = \frac{1}{1 + p \frac{L_{s1}(L_{s2} - C.L_{m2})}{k_p.C.L_{m1}.V_s}} \quad (37)$$

$$\frac{k_i}{k_p} = \frac{R_{s2}}{(L_{s2} - C.L_{m2})} \quad (38)$$

The gains of PI regulator are distinguished by:

$$k_p = \frac{L_{s1}(L_{s2} - C.L_{m2})}{\tau_r.C.L_{m1}.V_s} \quad (39)$$

$$k_i = k_p \frac{R_{s2}}{(L_{s2} - C.L_{m2})} = \frac{L_{s1}.R_{s2}}{\tau_r.C.L_{m1}.V_s} \quad (40)$$

$$\tau_r = \frac{1}{1 + p \frac{L_{s1}(L_{s2} - C.L_{m2})}{k_p.C.L_{m1}.V_s}} \quad (41)$$

3. 3. Adaptive Control Architecture In order to benefit the advantages of neural networks (NN) in power control of the CDFIG, we will incorporate two blocks of SHLNN alongside PI controllers. The first block will be utilized in the d-axis, associated with reactive power Q, and the second block in the q-axis, associated with active power P. The purpose of integrating these components is to enable the calculation of adaptive terms that actively influence our system, with the goal of minimizing tracking errors.

By implementing the SHLNN blocks on the d-axis and q-axis, we enhance the capability of our power control system. The SHLNN utilizes its neural network architecture to learn and adapt to varying operating conditions. It analyzes input data, such as voltage and current measurements, and produces adaptive terms u_{ad} and u_{aq} , that can be used to adjust the control signals for the CDFIG.

Inputs of NN blocks are:

- For the d-axis: $[V_{ds2}; Q_{s1}]$
- For the q-axis: $[V_{qs2}; P_{s1}]$

The outputs are Δ_d for d-axis and Δ_q for q-axis, which are considered as targets.

$$\frac{di_{ds2}}{dt} = v_{ds2} + \Delta_d \quad (42)$$

$$\frac{di_{qs2}}{dt} = v_{qs2} + \Delta_q \quad (43)$$

From Equations 5 and 6 we have:

$$\frac{di_{ds2}}{dt} = \frac{v_{ds2}}{L_{s2}} + \frac{1}{L_{s2}} \left(-R_{s2}i_{ds2} + L_{m2} \frac{di_{dr}}{dt} + \omega_{s2}L_{s2}i_{qs2} - \omega_{s2}L_{m2}i_{qr} \right) \quad (44)$$

$$\frac{di_{qs2}}{dt} = \frac{v_{qs2}}{L_{s2}} + \frac{1}{L_{s2}} \left(-R_{s2}i_{qs2} + L_{m2} \frac{di_{qr}}{dt} - \omega_{s2}L_{s2}i_{ds2} + \omega_{s2}L_{m2}i_{dr} \right) \quad (45)$$

From Equation 37 The outputs can be distinguished as follows:

$$\Delta_d = \frac{1}{L_{s2}} \left(-R_{s2}i_{ds2} + L_{m2} \frac{di_{dr}}{dt} + \omega_{s2}L_{s2}i_{qs2} - \omega_{s2}L_{m2}i_{qr} \right) \quad (46)$$

$$\Delta_q = \frac{1}{L_{s2}} \left(-R_{s2}i_{qs2} + L_{m2} \frac{di_{qr}}{dt} - \omega_{s2}L_{s2}i_{ds2} + \omega_{s2}L_{m2}i_{dr} \right) \quad (47)$$

Figure 3 shows the adaptive controller architecture of the CDFIG:

3. 4. Neural network Approximation Within a feedforward neural network, information exclusively goes in a singular direction, progressing from the input layer to the output layer, without any reverse propagation (30). In the context of controlling nonlinear systems, neural networks (NNs) are commonly utilized as adaptive components to offset the impact of uncertainties present within the system. These uncertainties can include unaccounted dynamics, parameters that vary with time, and errors in modeling. NNs are favored for this purpose due to their remarkable capability to approximate continuous mappings of real values, provided that an appropriate NN architecture is employed (11), the general structure of NN is shown in Figure 4.

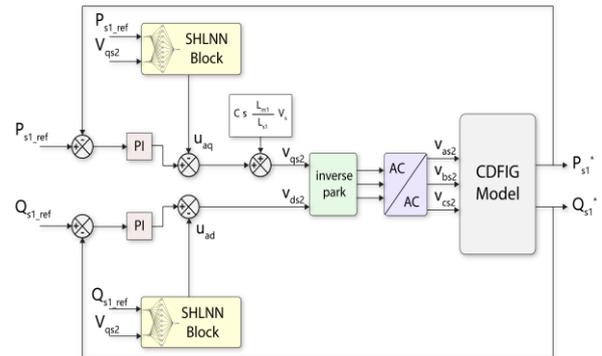


Figure 3. The adaptive control architecture of the CDFIG

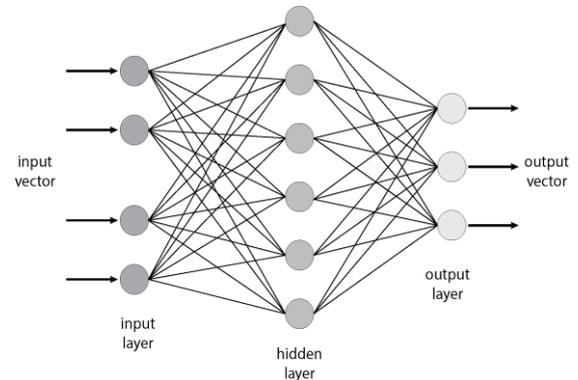


Figure 4. The structure of a single hidden layer neural network

The relation between the input and the output of a SHLNN is given by Rahmani and Belkheiri (11):

$$y_{ni} = \sum_{j=1}^{N_2} [w_{ij} \varphi_j (\sum_{k=1}^{N_1} v_{jk} x_k + \vartheta_{vj}) + \vartheta_{wi}] \quad (48)$$

$$i = 1, \dots, N_3 \quad \text{and} \quad x \in \mathbb{R}^{N_1}$$

The previous NN equation can be represented in compact form as follows (11):

$$y_{nn} = W^T \varphi(V^T x) \quad (49)$$

3. 5. Application The learning procedure can be accomplished through a diverse range of techniques, wherein the weights are iteratively adjusted until the generated output precisely matches the desired target output (31).

The Levenberg-Marquardt (LM) algorithm is widely recognized as one of the most prominent techniques for resolving least-squares problems (32). It serves as an approximation to the Newton technique aiming to achieve the speed of the second-order training without the necessity of calculating the Hessian matrix (31). Following a sequence of approximations and optimization steps, the weights are adapted using Equation 50 (33, 34):

$$W_{k+1} = W_k - (J_k^T + J_k + \mu I)^{-1} J_k e_k \quad (50)$$

The LM method is integrated into Matlab software as a function named “trainlm”, Figure 5 shows the training process using the LM algorithm to train our system within the Matlab environment. Notably, we employ a hidden layer consisting of 10 neurons and conduct 1000 training epochs.

After completing the learning phase, we proceed with implementing two SHLNN blocks in Matlab/Simulink. These blocks are designed to estimate the adaptive terms u_{ad} and u_{aq} , which will be utilized to counteract the targeted disturbance terms Δd and Δq . The adaptive terms are calculated by the following equations:

$$u_{ad} = \widehat{W}^T \varphi(V_0^T \mu_d) \quad (51)$$

$$u_{aq} = \widehat{W}^T \varphi(V_0^T \mu_q) \quad (52)$$

$$\mu_d = [V_{ds2} \ Q_{s1}]^T \quad (53)$$

$$\mu_q = [V_{qs2} \ P_{s1}]^T \quad (54)$$

\widehat{W} is updated according to the following law (35):

$$\widehat{W} = -F_{dq} [2\varphi(V_0^T \mu_{dq}) \tilde{i}_{dq} + k_{dq} (\widehat{W} - W_0)] \quad (55)$$

$$\tilde{i}_{dq} = i_{dqr} - i_{dq} \quad (56)$$

4. THE WIND POWER GENERATION SYSTEM

The wind power conversion system's operational range can be categorized into four distinct zones. The initial

zone, (I), during low wind speeds the electricity generation is nearly impossible. In the second zone, (II), the system utilizes the MPPT algorithm to effectively manage its operations. This optimization process considers fluctuations in wind speed, thereby maximizing power production. Moving on to Zone (III), which occurs when wind speeds exceed their nominal values, the system adjusts the pitch angle to regulate the produced electric power, maintaining it at or near the nominal level. Finally, Zone (IV) represents a state of strong winds that pose a potential threat to the wind turbine. In such emergency situations, the wind turbine is promptly halted to prevent any damage or harm (2).

4. 1. Maximum Power Point Tracking Strategy (MPPT)

The objective of MPPT algorithms is to observe and optimize the utilization of power generated by the WPCS (36). Numerous MPPT algorithms exist for in the realm of wind energy, each of these algorithms have their own merits and demerits. The optimal torque control (OTC) is very simple in usage, fast speed convergence and high effectiveness (37). This algorithm focuses on measuring the generator rotor's speed and computing the required torque or power instead of measuring the wind speed, the following control method is employed to ensure that the generator torque remains at its designated reference value (27):

$$T_{em_ref} = K_{opt} \Omega_{mec}^2 \quad (57)$$

$$K_{opt} = \frac{1}{2} \frac{C_{p_max}}{\lambda_{opt}^3} \rho \pi R_t^3 \frac{1}{G^3} \quad (58)$$

The power coefficient C_p depending on the velocity ratio λ and the pitch angle β reaches its maximum value ($C_{p_max} = 0.44$) when the pitch angle is zero, thus the velocity ratio becomes at its optimal value ($\lambda_{opt} = 7$).

Figure 6 depicts the various components of the wind turbine using a block diagram, with a specific highlight on the MPPT strategy:

4. 2. The WPGS Based on The CDFIG The process of using the CDFIG as a wind turbine generator involves several types of connections, one of them is by linking

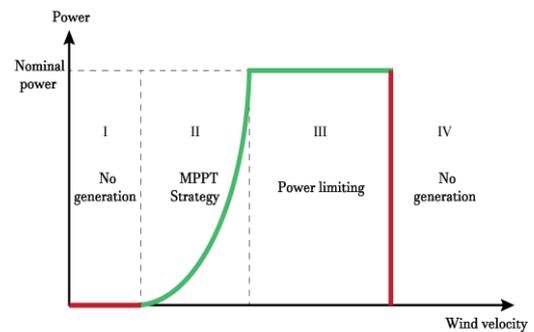


Figure 5. The operational range of the WPCS

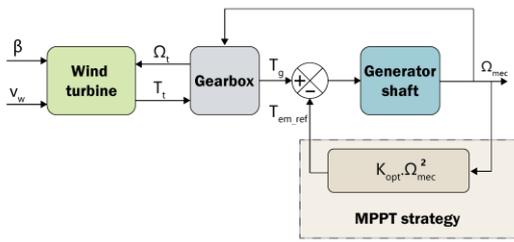


Figure 6. MPPT strategy usage in the wind turbine

the initial stator directly to the grid, while the second stator is linked to the grid via an intermediate AC/AC converter which allows the energy to be controlled and transferred in both directions (6).

The WPCS based on the CDFIG is illustrated in Figure 7.

5. RESULTS AND DISCUSSION

The parameters of the generators and the wind turbine used for the simulation have been selected from literature (25).

The simulation was done using MATLAB/Simulink environment. In the first case, the PI controller implemented alone to control the active power and reactive power. In the second case, the neural network will be added alongside to the PI controller. The results show the difference between both cases. Figure 8 shows the wind velocity profile. Rotor speed in RPM is illustrated in Figure 9.

The wind speed varies between approximately 6 m/s and 13 m/s. The pattern of changes in wind speed is irregular, with several peaks and dips, indicating variability in wind conditions.

The active power of the first generator; first case and second case controlled by the PI controller are shown in

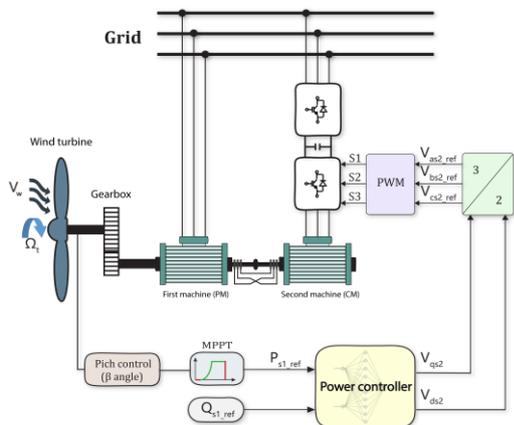


Figure 7. The WPCS based on the CDFIG

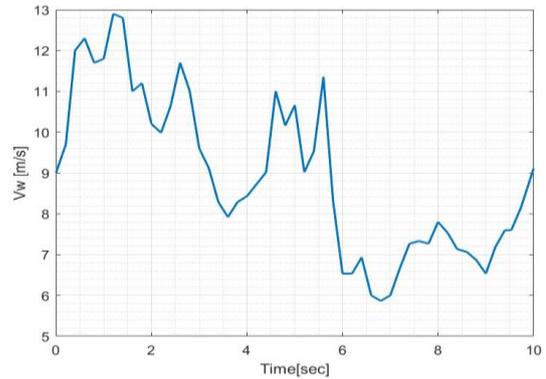


Figure 8. Wind velocity profile

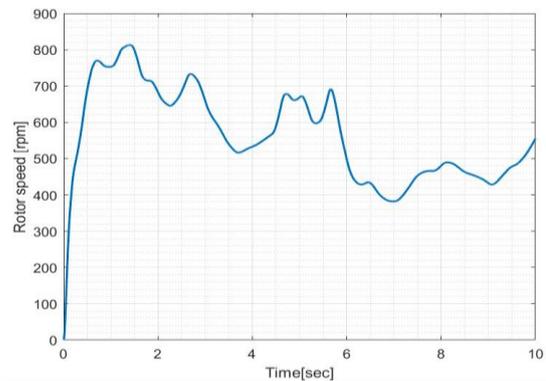


Figure 9. Rotor speed

Figures 10 and 11, respectively. The reactive power of the first generator; first case and second case are shown in Figures 12 and 13, respectively. It is evident that the active and reactive power cannot closely track their reference values during periods of rapid strong wind velocity variation as we can see in Figure 10 the power error amount to 100 kW. However, when neural network is introduced in conjunction with the PI controller, a significant enhancement in results becomes apparent. In Figure 11, the active power plot demonstrates a near-elimination of errors. This improvement is further highlighted by examining the transient response, which is markedly better in the second case. The transient response, a critical aspect of wind generation systems, demonstrates an enhanced ability to quickly and efficiently adjust to changes in wind speed and load demands. This rapid adjustment capability is vital for maintaining system stability and ensuring consistent power output. In Figure 13, the reactive power plot shows a remarkable reduction in errors by 60 KVA.

These results provide compelling justification for the accurate identification of inversion errors (Δ_q) and (Δ_d) by the adaptive terms (u_{aq}) and (u_{ad}), respectively, as depicted in Figures 14 and 15.

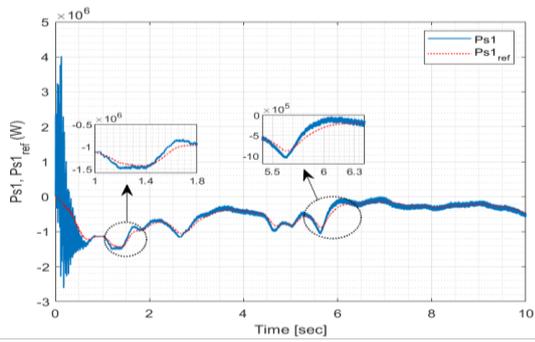


Figure 10. The active power of the first generator; first case

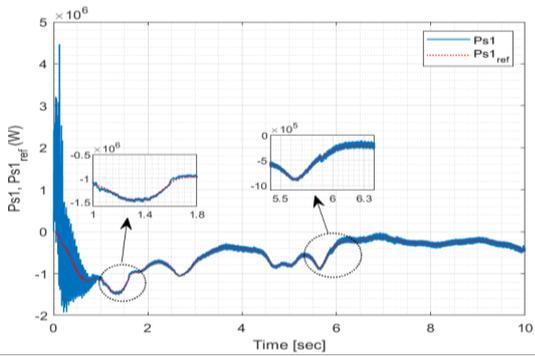


Figure 1. The active power of the first generator; second case

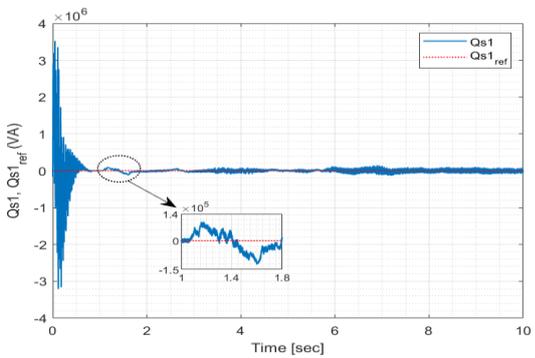


Figure 2. The reactive power of the first generator; first case

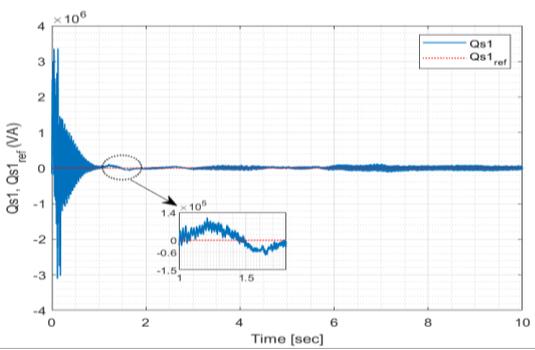


Figure 3. The reactive power of the first generator; second case

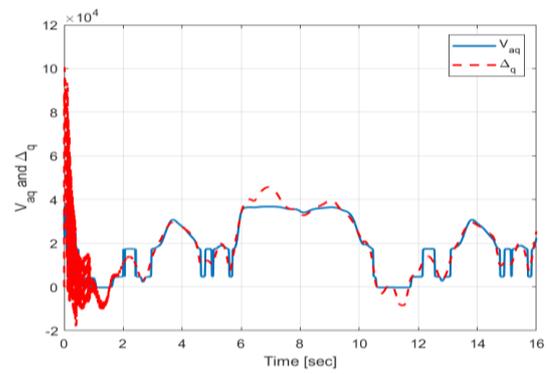


Figure 4. Identification of Δ_q by NN

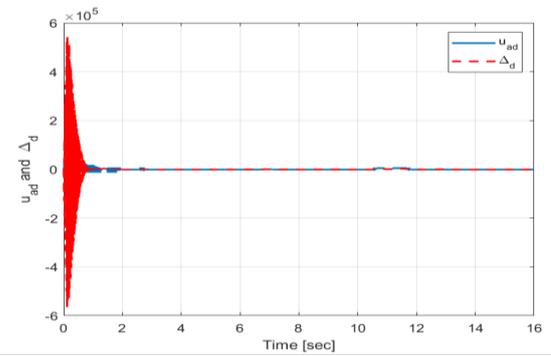


Figure 5. Identification of Δ_d by NN

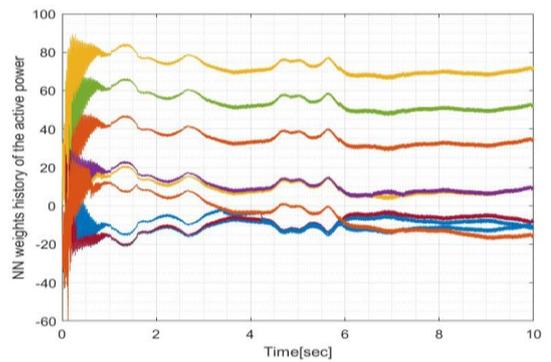


Figure 6. History of NN weights of the active power

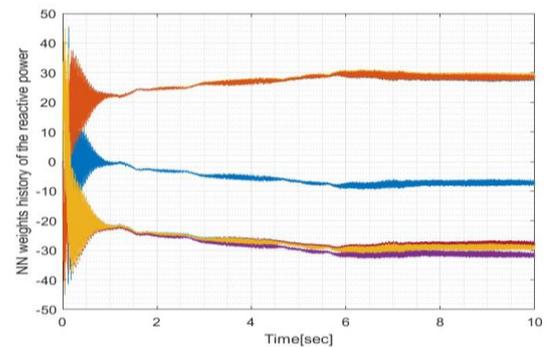


Figure 7. History of NN weights of the reactive power

The neural network weights history of the active power and the reactive power are shown in Figures 16 and 17, respectively. As observed the weights are chosen carefully and accurately to adapt with the system and to provide a better control command the values of this last are in between -40 to 40 for the NN of the reactive power and -40 to 90 and for the NN of the active power the initial weights were selected using the developer expertise and are optimized during the training session for best outcomes.

This improved performance can be attributed to the neural network's ability to predict and compensate for the system's behavior more effectively than the conventional PI controller, which is based on fixed gains and does not adapt to changing conditions dynamically.

6. CONCLUSION

The integration of neural network with a PI controller in wind generator control systems has proven to be a highly effective approach for enhancing performance, particularly in the presence of rapidly changing wind conditions. The results clearly demonstrate a remarkable reduction in errors and a significant improvement in the accuracy of power control.

This study serves as a strong foundation for the application of neural networks in renewable energy systems, offering the potential for improved efficiency and grid stability. The carefully chosen and optimized neural network weights play a critical role in adapting to system dynamics, ensuring a robust control strategy.

As the renewable energy industry continues to grow, innovative solutions like the one explored in this study will become increasingly important in harnessing the full potential of wind energy and addressing the challenges posed by variable wind conditions. It is clear that neural networks hold great promise for the future of the wind energy generation, offering a path toward more reliable and efficient power production.

7. REFERENCES

- Strielkowski W, Civin L, Tarkhanova E, Tvaronavičienė M, Petrenko Y. Renewable Energy in the Sustainable Development of Electrical Power Sector: A Review. *Energies*. 2021;14(24). 10.3390/en14248240
- Chojaa H, Derouich A, Chehaidia SE, Zamzoum O, Taoussi M, Elouatouat H. Integral sliding mode control for DFIG based WECS with MPPT based on artificial neural network under a real wind profile. *Energy Reports*. 2021;7:4809-24. 10.1016/j.egy.2021.07.066
- El Achkar M, Mbayed R, Salloum G, Monmasson E. New voltage compensation of a standalone CDFIG supplying unbalanced loads. *International Journal of Electrical Power & Energy Systems*. 2021;124. 10.1016/j.ijepes.2020.106396
- Olubamiwa OI, Gule N. A Review of the Advancements in the Design of Brushless Doubly Fed Machines. *Energies*. 2022;15(3). 10.3390/en15030725
- Achkar ME, Mbayed R, Salloum G, Patin N, Monmasson E. Voltage Control of a Stand-Alone Cascaded Doubly Fed Induction Generator. *IEEE Transactions on Industrial Electronics*. 2019;66(1):762-71. 10.1109/tie.2018.2856186
- Nguyen D-D, Than N-H, Hoang D-T. The cascade methods of doubly-fed induction machine for generator system. *International Journal of Power Electronics and Drive Systems (IJPEDS)*. 2021;12(1). 10.11591/ijpeds.v12.i1.pp112-120
- Maafa A, Mellah H, Ghedamsi K, Aouzellag D. Improvement of Sliding Mode Control Strategy Founded on Cascaded Doubly Fed Induction Generator Powered by a Matrix Converter. *arXiv preprint arXiv:220811140*. 2022.
- Taghinezhad J, Sheidaei S. Prediction of operating parameters and output power of ducted wind turbine using artificial neural networks. *Energy Reports*. 2022;8:3085-95. 10.1016/j.egy.2022.02.065
- Narimene K, Kheira M, Mohamed F. Robust Neural Control of Wind Turbine Based Doubly Fed Induction Generator and NPC Three Level Inverter. *Periodica Polytechnica Electrical Engineering and Computer Science*. 2022;66(2):191-204. 10.3311/PPee.19921
- Vajdian M, Zahrai S, Mirhosseini S, Zeighami E. Predicting Shear Capacity of Panel Zone Using Neural Network and Genetic Algorithm. *International Journal of Engineering*. 2020;33(8):1512-21.
- Rahmani B, Belkheiri M. Adaptive neural network output feedback control for flexible multi-link robotic manipulators. *International Journal of Control*. 2018;92(10):2324-38. 10.1080/00207179.2018.1436774
- Kashkooli MA, Madani S, Kiyoumars A. Soft Grid Connection of Cascaded Doubly Fed Induction Generator Using Direct Torque & Flux Control for Wind Turbine Application. *Majlesi Journal of Energy Management*. 2016;5.
- Zahedi H, Arab Markadeh G, Taghipour S. Real-time implementation of sliding mode control for cascaded doubly fed induction generator in both islanded and grid connected modes. *Journal of Electrical and Computer Engineering Innovations (JECEI)*. 2020;8(2):285-96.
- Svečko R, Gleich D, Sarjaš A. The Effective Chattering Suppression Technique with Adaptive Super-Twisted Sliding Mode Controller Based on the Quasi-Barrier Function; An Experimentation Setup. *Applied Sciences*. 2020;10(2). 10.3390/app10020595
- Abdolhadi HZ, Markadeh GA, Boroujeni ST. Sliding mode and terminal sliding mode control of cascaded doubly fed induction generator. *Iranian Journal of Electrical and Electronic Engineering*. 2021;17(3):1955-.
- Yan X, Cheng M. A Robustness-Improved Control Method Based on ST-SMC for Cascaded Brushless Doubly Fed Induction Generator. *IEEE Transactions on Industrial Electronics*. 2021;68(8):7061-71. 10.1109/tie.2020.3007087
- Yu X, Feng Y, Man Z. Terminal Sliding Mode Control – An Overview. *IEEE Open Journal of the Industrial Electronics Society*. 2021;2:36-52. 10.1109/ojies.2020.3040412
- Bayhan S, Kakosimos P, Rivera M, editors. Predictive torque control of brushless doubly fed induction generator fed by a matrix converter. 2018 IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG 2018); 2018: IEEE.
- Abdollahimi H, Arab Khaburi D. A novel model predictive voltage control of brushless cascade doubly-fed induction

- generator in stand-alone power generation system. *International Journal of Engineering*. 2021;34(5):1239-49.
20. Dauksha G, Iwanski G. Indirect Torque Control of a Cascaded Brushless Doubly-Fed Induction Generator Operating With Unbalanced Power Grid. *IEEE Transactions on Energy Conversion*. 2020;35(2):1065-77. 10.1109/tec.2020.2972500
 21. Morales LA, Fabara P, Pozo DF. An Intelligent Controller Based on LAMDA for Speed Control of a Three-Phase Inductor Motor. *Emerging Science Journal*. 2023;7(3):676-90. 10.28991/esj-2023-07-03-01
 22. Ismaiel A. Wind Turbine Blade Dynamics Simulation under the Effect of Atmospheric Turbulence. *Emerging Science Journal*. 2022;7(1):162-76. 10.28991/esj-2023-07-01-012
 23. Mahmoudabadi A, Kanaani M, Pourhossein Ghazimahalleh F. Modifying Hidden Layer in Neural Network Models to Improve Prediction Accuracy: A Combined Model for Estimating Stock Price. *HighTech and Innovation Journal*. 2022;3(1):45-55. 10.28991/hij-2022-03-01-05
 24. Thaghipour Boroujeni S. Complex vector modeling of a doubly fed cascaded cage rotor induction machine. *Electrical Engineering*. 2020;102(3):1831-42. 10.1007/s00202-020-00996-7
 25. Maafa A, Aouzellag D, Ghedamsi K, Abdessemed R. Cascaded doubly fed induction generator with variable pitch control system. *Rev Roum Sci Techn-Électrotechn et Énerg*. 2016;61(4):361-6.
 26. Yan X, Cheng M, editors. A robust grid synchronization method for cascaded brushless doubly fed induction generator. 2019 22nd International Conference on Electrical Machines and Systems (ICEMS); 2019: IEEE.
 27. Dahbi A, Reama A, Hamouda M, Nait-Said N, Nait-Said M-S. Control and study of a real wind turbine. *Computers & Electrical Engineering*. 2019;80. 10.1016/j.compeleceng.2019.106492
 28. Kelkoul B, Boumediene A. Stability analysis and study between classical sliding mode control (SMC) and super twisting algorithm (STA) for doubly fed induction generator (DFIG) under wind turbine. *Energy*. 2021;214. 10.1016/j.energy.2020.118871
 29. Boubzizi S, Abid H, El Hajjaji A, Chaabane M. Comparative study of three types of controllers for DFIG in wind energy conversion system. *Protection and Control of Modern Power Systems*. 2018;3:1-12.
 30. El Hamidi K, Mjahed M, El Kari A, Ayad H. Adaptive Control Using Neural Networks and Approximate Models for Nonlinear Dynamic Systems. *Modelling and Simulation in Engineering*. 2020;2020:1-13. 10.1155/2020/8642915
 31. Nazaré G, Castro R, Gabriel Filho LRA. Wind power forecast using neural networks: Tuning with optimization techniques and error analysis. *Wind Energy*. 2019;23(3):810-24. 10.1002/we.2460
 32. Wang M, Xu X, Yan Z, Wang H. An online optimization method for extracting parameters of multi-parameter PV module model based on adaptive Levenberg-Marquardt algorithm. *Energy Conversion and Management*. 2021;245. 10.1016/j.enconman.2021.114611
 33. Ye Z, Kim MK. Predicting electricity consumption in a building using an optimized back-propagation and Levenberg–Marquardt back-propagation neural network: Case study of a shopping mall in China. *Sustainable Cities and Society*. 2018;42:176-83. 10.1016/j.scs.2018.05.050
 34. Peiris AT, Jayasinghe J, Rathnayake U, Vallée F. Forecasting Wind Power Generation Using Artificial Neural Network: “Pawan Danawi”—A Case Study from Sri Lanka. *Journal of Electrical and Computer Engineering*. 2021;2021:1-10. 10.1155/2021/5577547
 35. Abbas HA, Belkheiri M, Zegnini B. Feedback linearisation control of an induction machine augmented by single-hidden layer neural networks. *International Journal of Control*. 2016;89(1):140-55.
 36. Le XC, Duong MQ, Le KH. Review of the Modern Maximum Power Tracking Algorithms for Permanent Magnet Synchronous Generator of Wind Power Conversion Systems. *Energies*. 2022;16(1). 10.3390/en16010402
 37. Apata O, Oyedokun DTO. An overview of control techniques for wind turbine systems. *Scientific African*. 2020;10. 10.1016/j.sciaf.2020.e00566

COPYRIGHTS

©2024 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 مشارکت با هدف بهبود محدودیت‌های یک کنترل‌کننده متعارف تناسبی-انتگرالی-مشتقی (PID) با یکپارچه سازی یک شبکه عصبی تک لایه پنهان (SHLNN) در سیستم کنترل انجام شده است. به‌ویژه، در سیستم تبدیل انرژی بادی (WPGS) بر اساس مولد القایی دوبل تغذیه شده با کاسکاد (CDFIG) هدف از این یکپارچه‌سازی کاهش ضررهای انرژی ناشی از ناپایداری و نادقیقی‌های سیستم‌های سنتی است که به واسطه محدودیت‌های باد ایجاد می‌شوند. علاوه بر این، این روش در جستجوی بهره‌برداری از مزایای مولد القایی دوبل تغذیه شده با کاسکاد (CDFIG) و شبکه عصبی مصنوعی در کنترل WPCS است. در حالی که کنترل‌کننده PI همچنان مکانیزم کنترل اصلی است، محدودیت‌های آن به دلیل غیرخطی بودن و عدم قطعیت‌های WPCS توسط قابلیت‌های تطبیق‌پذیر و پیچیدگی محاسباتی پایین SHLNN مورد توجه قرار گرفته است. این رویکرد از توانایی یادگیری رابطه‌های غیرخطی شبکه‌های عصبی استفاده می‌کند، که با استفاده از داده‌های تاریخی سرعت باد و تولید انرژی آموزش داده شده‌اند، تا به طور مؤثر با تغییرات ذاتی الگوهای باد روبرو شوند. شبیه‌سازی‌ها با استفاده از Matlab/Simulink انجام می‌شوند تا عملکرد کنترل‌کننده PI با یک پیاده‌سازی شبکه عصبی تک‌لایه مقایسه شود. در نهایت، نتایج شبیه‌سازی ارائه و به طور کامل تحلیل می‌شوند و بیش‌هایی در مورد اثربخشی رویکرد پیشنهادی ارائه می‌دهند.



Critical Success Factor Implementation of Land Value Capture on a Toll Road Assignment Scheme: SEM-PLS Approach

R. Fitriadi Kurnia^{*a}, A. Anwar^a, Y. Latief^a, L. B. Sihombing^b

^a Civil Engineering Department, Faculty of Engineering, Universitas Indonesia, Depok, Indonesia

^b Civil Engineering Department, Faculty of Technology and Design, Universitas Pembangunan Jaya, South Tangerang, Indonesia

PAPER INFO

Paper history:

Received 02 October 2023

Received in revised form 28 November 2023

Accepted 21 December 2023

Keywords:

Critical Success Factors

Land Value Capture

Toll Road Assignment

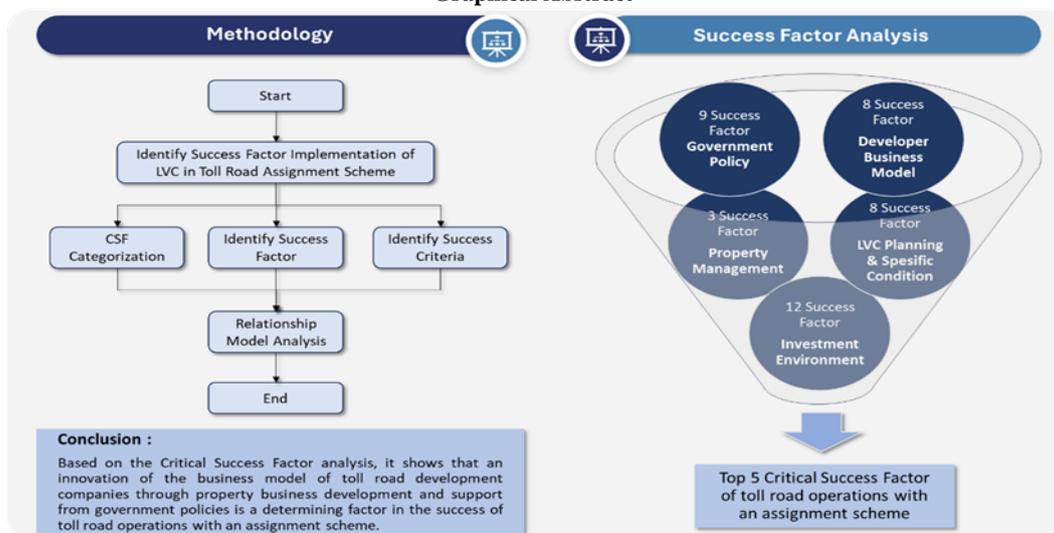
Trans Sumatra Toll Road

ABSTRACT

The Trans-Sumatra Toll Road (TSTR) located in Indonesia operation which is under the management of State-Owned Enterprises (SOE) faces several challenges and requires alternative sources of financing and income, one of which is Land Value Capture (LVC)-based area development. This research aims to identify and analyze the Critical Success Factors needed to implement land value capture in the TSTR project. The results of this research obtained five success factors with the highest ranking and a model of the relationship between variables in implementing land value capture in SOE assigned toll road operations based on the consensus of the expert. Through literature studies, 40 success factors were grouped into five categories and 14 criteria to successfully implement land value capture on toll road-based infrastructure that experts validated. The validated success factors were processed through a series of expert assessments using the Delphi-Method questionnaire, resulting in five success factors with the highest ranking in each category. The relationships between variables were further obtained from PLS-SEM modeling and were analyzed. The analysis of the relationship model produced relationships between variables including Government Policy, Toll Road Developer Business Model, Asset/Property Management, Investment Supporting Environment, LVC Planning and Specific Project Conditions. The present study result may determine factors that can ensure the successful completion of toll road projects with SOE assignment scheme.

doi: 10.5829/ije.2024.37.07a.11

Graphical Abstract



*Corresponding Author Email: remifitriadi@gmail.com (R. Fitriadi Kurnia)

Please cite this article as: R. Fitriadi Kurnia, A. Anwar, Y. Latief, L. B. Sihombing. Critical Success Factor Implementation of Land Value Capture on a Toll Road Assignment Scheme: SEM-PLS Approach. International Journal of Engineering, Transactions A: Basics. 2024;37(07):1317-30.

NOMENCLATURE

<i>TSTR</i>	Trans Sumatra Toll Road	<i>PPP</i>	Public Private Partnership
<i>SOE</i>	State-Owned Enterprises	<i>RII</i>	Relative importance index
<i>LVC</i>	Land Value Capture	<i>SD</i>	Standard Deviation
<i>MTR</i>	Mass Transit Railway	<i>MS</i>	Mean Score
<i>CSF</i>	Critical Success Factors	<i>TOD</i>	Transit Oriented Development
<i>PLS-SEM</i>	Partial Least Square – Structural Equation Modelling	<i>HK</i>	Hutama Karya

1. INTRODUCTION

Infrastructure development is one of the Indonesia's main development targets in 2020-2024 National Medium-Term Development Plan. One of the targets is the construction of 2,500 km of toll roads throughout Indonesia, 1,600 km of which are part of the Trans-Sumatra Toll Road (TSTR). TSTR is a mega project of toll road network that connects Sumatra Island from Lampung to Aceh, with a total length of 2,813 km, and a projected total investment cost of million US\$ 38,297.

The construction of this toll road faces several challenges, where the TSTR is not financially feasible but economically feasible, making the project unattractive for toll road developers to invest in long-term operations. To overcome these challenges, the Indonesian government chose to implement the project with the SOE Assignment scheme. The assignment of SOE to TSTR is regulated through Presidential Decree Number 100 of 2014, Presidential Decree Number 117 of 2015 (1st amendment), and Presidential Decree Number 131 of 2022 (2nd amendment) concerning the Acceleration of Toll Road Development on Sumatra island. SOE assignments are designed to be implemented on toll roads that are projects economically feasible but not financially feasible due to government budget limitations.

In its implementation, the assigned TSTR Concession development company faced several challenges, including low financial viability, low average daily traffic, constraints on government budget allocation, and increasing concession company debt, resulting in declining company performance. To face these challenges, the TSTR Concession Company needed alternative sources of financing and income, one of which is area development based on Land Value Capture (LVC).

LVC is based on the principle of a virtuous value cycle, where value in the form of additional economic improvements created from infrastructure investments is then captured (either in part or in full) (1) through value capture mechanisms to recover the capital costs of the investment or reinvest in the land or area.

Figure 1 shows the Virtuous Value Cycle which provides a policy-based framework to create value that can increase economic growth, using a value capture mechanism and providing economic growth results that will be devoted to replacing the initial investment capital

costs used in infrastructure projects (2). The implementation of the virtuous value cycle as an alternative source of income can reduce the dependence of toll road infrastructure on income originating from toll payments.

Through utilizing land value capture, TSTR concession companies can have additional sources of income apart from toll fees. In its application to the urban transit infrastructure of the Hong Kong MTR and Hyderabad Metro, the concession company is implementing LVC-based property development, integrated with transit infrastructure based on urban area development master plans that enable the company to generate alternative revenue streams, with the Hyderabad Metro and Hong Kong MTR respectively accounted for 45% and 38% of its total revenue (3).

The use of LVC is a great opportunity to become an alternative source of TSTR financing due to the economic improvements generated by infrastructure investment (2) and the principle of beneficiary pays, where infrastructure is financed by the beneficiaries (4).

As a large-scale infrastructure project, TSTR is expected to significantly increase connectivity, connect economic activity nodes, reduce transportation and logistics costs, and reduce travel time and distance (2). Currently, the implementation of LVC in Indonesia is not yet optimal (5), so it cannot be used as an alternative source of income and financing for toll road infrastructure.

A study of the Critical Success Factors (CSF) can identify things that need to be prioritized or require

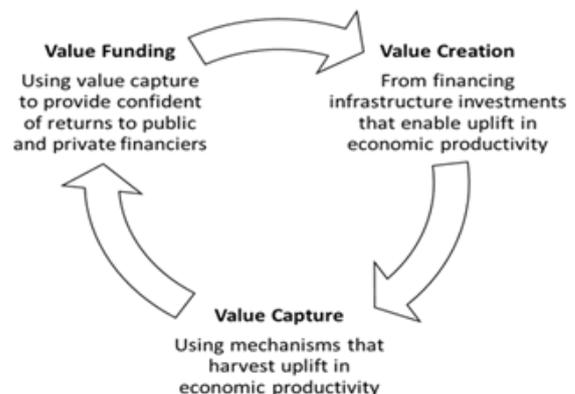


Figure 1. Virtuous Value Cycle Source: processed from ADB, 2021 [5]

special attention when implementing LVC on the TSTR. Modeling between the determining factors for success and the success of implementing land value capture can map the relationship between critical success factors and the success of implementing land value capture on the TSTR. Partial Least Square (PLS) can analyze cause and effect relationships and structural relationships based on latent variables and manifest variables in the application of land value capture on the Trans-Sumatra Toll Road with SOE Assignment. The advantage of using the PLS-SEM method is that it uses a small sample (6).

Based on literature studies regarding land value capture, it is known that there are several gaps i.e. no studies identify the critical success factors and success criteria in implementing land value capture especially on toll roads with the Assignment scheme and there is no research to develop a model of the relationship between critical success factors and the success criteria to implementation of land value capture mechanisms on SOE Assignment toll roads. To complement previous research gaps, in this research a Structural Equation Model regarding factors that influence the successful implementation of the land value capture mechanism in the Trans-Sumatra Toll Road Assignment was developed.

Through CSF identification and CSF land value capture modeling which have been successfully applied to the Trans-Sumatra Toll Road Assignment, it is hoped that it can encourage the successful implementation of land value capture as an alternative source of income on the Trans-Sumatra Toll Road. This paper aims to identify and model the influence of critical success factors on the success of implementing land value capture as an alternative source of financing for the Trans-Sumatra Toll Road. It is also hoped that the results of this research can be a reference in implementing LVC, especially on other SOE Assignment toll roads in Indonesia.

2. LITERATURE REVIEW

2.1. Trans-Sumatera Toll Road as Assignment SOE Toll Road Scheme

The construction of the Trans-Sumatera Toll Road is one of the largest mega construction projects in Indonesia and even the world with total financing of IDR 572 trillion. Its 2,813 km road length consists of 24 sections stretching from Bakauheni to Banda Aceh, which is divided into four construction stages. TSTR is regulated in Presidential Decree Number 100 of 2014 regarding the Acceleration of Toll Road Development in Sumatra, which was subsequently updated in Presidential Decree No. 117 of 2015 and Presidential Decree No. 131 of 2022.

The Assignment SOE scheme was chosen because the TSTR does not yet have financial viability, even though it has economic viability, which makes this project

unattractive for private investment (7). According to the Head of the General Section of BPJT Mahbullah Nurdin (7), the assignment of the Trans-Sumatra Toll Road Concession to PT Hutama Karya (HK) as a Toll Road Business Entity was carried out to a SOE that was 100% owned by the government and had the greatest technical and strong financial capacity. Through the Assignment of SOE, HK as the Toll Road Developer was assigned to carry out toll road business concession which includes funding, technical planning, construction execution work and also operation and maintenance toll road.

2.2. Land Value Capture (LVC) on Infrastructure Projects

Recent studies in land value capture are mostly on transit infrastructure context in several highly populated cities such as Hyderabad, Hong Kong, Wuhan, and Guangzhou (3, 8). Based on the LVC best practices, researchers emphasize the pivotal role of government bodies in multi-level authorities in establishing policies in land value capture implementation (3, 5, 8, 9). In the case of urban transit infrastructure, successful LVC implementation highly depends on the business model of the transit agencies and its capability in asset and property management (3, 4, 10). Although LVC research has mostly been carried out on rail-based infrastructure, currently the related research of LVC on toll road infrastructure has begun to be carried out through the development of the Road Plus Property Developer business model framework (11).

Several studies also expressed the need for a favorable investment environment in the form of a supportive legal framework and regulation (3-5, 8, 12), zonation and spatial planning (3, 13), economic policies (13-15), good governance (13), political support (3, 4, 9, 12), and attractive financing package (8) in infrastructure financing involving private investment.

Specific project conditions and solid land value capture planning in terms of good value creation based on profitable and contextual business model (3, 8, 12), integrated property development with the main infrastructure (3), alignment with medium-to-long term government strategic objectives (3), high-quality operational services (8), beneficiary stakeholder support (4, 5, 8, 9) and potential application of a combination of LVC instruments (3, 8) are also several success criteria mentioned in the recent studies.

Based on the LVC best practices on transit infrastructures, these success factors need further studies in the context of toll roads in Indonesia, mainly on the SOE Assignment scheme that would likely utilize the financing scheme.

2.3. Critical Success Factor (CSF) Implementation Land Value Capture on Assignment Toll Road

Critical Success Factor (CSF) is an action, decision, condition, or situation that must go well to achieve the

desired goals of a project (15). In general, research related to CSF in the construction industry, including toll road projects shows that factors related to human resources are important factors from the owner and contractor perspective, while project characteristics are the most important factors from the consultant's perspective.

Based on previous research (16), the characteristics of critical success factors, including: the form of events that cannot be predicted but have significant risks and, involving the performance of key individuals, it can have a good or bad effect on the organization's ability to achieve its goals and involves critical operations, systems, or facilities that need to be monitored or subject to contingency plans. CSF have different CFSs in the context of toll road business assignment schemes for each country and project and depend on dynamic existing environmental conditions along with changes in policy and the industrial environment.

This is based on the benchmarking carried out through literature studies regarding best practices for implementing land value capture in transit infrastructure. The success factors are then grouped into five categories, including: Government Policy, Toll Road Developer Business Model, Asset/Property Management, Profitable Investment Environment, and Land Value Planning and Special Project Conditions.

2. 4. Relationship between Variables

2. 4. 1. The Relationship between Government Policy and the Successful Implementation of LVC

Implementing LVC involving various government agencies at various levels across institutions requires a collaborative framework (8) and good coordination. Furthermore, the importance of government support in implementing LVC, whether in the form of financial support, zoning, or land acquisition, was emphasized (3). In addition, the need for implementing authorities, as well as the importance of preparing infrastructure investment packages in supporting LVC implementation was also emphasized (8).

2. 4. 2. The Relationship between Government Policy and the Successful Implementation of LVC

Through the previous study (8), the role of the concession company's business model in implementing LVC, including competency/capability in the property market, long-term oriented contract preparation, economic growth-oriented development that aligns with transportation infrastructure, as well as cooperation in land acquisition with the government were identified. Other studies (3, 4) also identified the importance of toll road developer competency/capability in business model planning and asset management as a success factor in implementing LVC.

2. 4. 3. The Relationship between Government Policy and the Successful Implementation of LVC

Studies that have been carried out before (9, 10) identified transparency in the implementation of LVC as a factor that supports the implementation of LVC. Meanwhile, other studies (3, 8) suggested asset management collaboration between concession companies and property developers as a success factor in implementing LVC. In connection with several studies (3, 8) also stated the importance of property developer competency or capability in LVC implementation areas as a factor that supports the success of LVC implementation.

2. 4. 4. The Relationship between a Supporting Investment Environment and the Successful Implementation of LVC

Two studies (3, 8) have identified spatial planning, transportation, and zoning; as well as support for good macroeconomic conditions and economic growth as conditions that support the successful implementation of the LVC mechanism. One of the research (8) further revealed a financing package that includes land management around transportation infrastructure as a supporting factor for the successful implementation of LVC. Based on the other studies (5, 9), regulatory support in implementing LVC was expressed.

2. 4. 5. The Relationship between LVC Planning and Specific Project Conditions with the Successful Implementation of LVC

The application of LVC to transportation infrastructure is closely related to the quality of the transportation infrastructure. This is supported by literature (4, 5, 8) that good value creation is needed in implementing the LVC mechanism as well as Li et al. (8) which reveals that quality infrastructure operational services can support the implementation of LVC.

3. RESEARCH METHODOLOGY

This research applied qualitative methods to develop success factors for implementing LVC on the Assignment SOE toll road. Currently, land value capture is widely used as an alternative financing for urban transport infrastructure. Based on the literature review, the author identified 40 success factors and 14 success criteria for implementing LVC in urban transport infrastructure. These success factors were then categorized into five categories consisting of X1-Government Policy (9 SF), X2-Toll Road Company Business Model (8 SF), X3-Asset/Property Management (3 SF), X4-Profitable Investment Environment (12 SF), and X5-Land Value Planning and Special Project Conditions (8 SF). Figure 2 presents the research model.

The obtained success factors, success criteria, and success factors categories were then validated through multiple rounds of Delphi Method questionnaires. This decision was due to the limited study of the LVC concept in Indonesia and the limited number of LVC experts on toll road infrastructure in Indonesia. Through the Delphi Method, the research aims to obtain a consensus among the experts in providing an assessment based on an anonymous iterative process with a small number of respondents. A recent study shows that the Delphi Method with sample sizes greater than or equal to five combined with 1-7 linear Likert Scale (17).

Figure 2 shows the six latent variables. Arrows between variables illustrate the cause-and-effect relationship between variables, for example: X1 (government policy) influences X2 (toll road company business model), X3 (asset/property management), X4 (favorable investment environment), X5 (project specific planning and conditions), and Y1 (successful implementation of land value capture).

The identified success factors, success factor categories, and success criteria were then validated using an expert judgment questionnaire in the context of SOE Assignment toll roads. The expert assessment questionnaire was given to five expert practitioners with at least 10 years of experience in infrastructure projects, infrastructure financing and investment, or public-private partnerships and occupied the position of manager or higher in their respective institutions, as shown in Table 1.

After obtaining valid data from each expert, the data were further tabulated and analyzed. Success factor items or invalid wording were eliminated when the majority of experts (three or more) did not agree with the success factor/success criteria. The success factors that have been validated were then ranked based on the expert's assessments according to their level of importance through a Delphi Method questionnaire round, to obtain expert consensus regarding the determining factors for success in LVC implementation. Critical success factors (CSF) are a number of actions, decisions, conditions, or

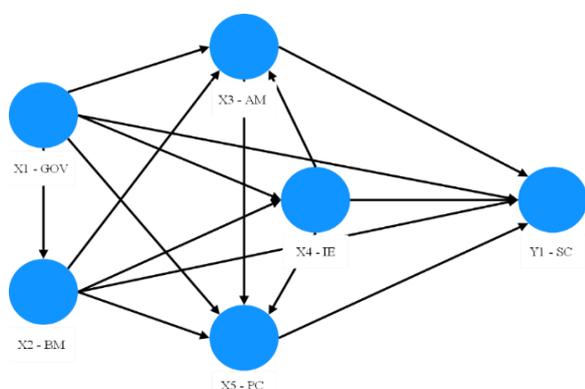


Figure 2. Research model

TABLE 1. Expert Data

Expert	Institution	Job Function	Experience
Expert 1	State Owned Enterprise	Director of Toll Road Company	30 Years
Expert 2	Consultancy Firm/ Private University	Senior Project Manager/ Lecturer	30 Years
Expert 3	Ministry of Public Works and Housing	Secretary of the Directorate General	22 Years
Expert 4	State Owned Enterprise	Director of Toll Road Company/ Head of Toll Road Procurement Unit	32 Years
Expert 5	State Owned Enterprise	Vice President of Strategic Planning	17 Years

circumstances that must be met to achieve success, both in the context of an organization and a project (15, 16). The CSF approach was used to provide an understanding of the business environment and the actions that need to be taken (16).

Before the second round of assessment, the experts were presented with the first assessment data, then asked to provide a re-assessment based on this data as a reference for the second round of assessment and so on. Consensus is considered achieved if all experts provide the same assessment as the previous assessment. The assessment uses a 7-point Likert scale, because the sensitivity is better to get a more accurate evaluation (18). The assessment of success factors was analyzed based on the relative importance index (RII), standard deviation (SD), and mean score (MS) (16). RII is calculated using Equation 1.

$$RII = \frac{\text{Total point score}}{A \times N} \quad (0 < RII \leq 1) \quad (1)$$

RII = Relative importance index

Total point score = summation of all ratings for a given factor

A = maximum rating possible, in this case 7

N = number of respondents for the factor

If there are success factors with similar RII values, ranking is done using the lower SD value. If the success factors also have similar SD values, then the ranking is carried out using the larger MS value. The results are the top five success factors from the five categories mentioned above as determining factors for the success of implementing LVC on toll road assignment of SOE in Indonesia based on expert consensus.

The CSF and success criteria obtained were then sent as a pilot survey questionnaire to 10 respondents to ensure clarity of survey indicators. A final questionnaire was then sent to respondents involved in toll road projects, transportation infrastructure, infrastructure financing and investment, and also Public-Private Partnership (PPP) projects to assess the importance of 23

critical success factors and 14 success criteria. The minimum sample requirements for this research are based on the PLS model which has five structural paths that lead to the implementation of LVC as an endogenous variable, thus requiring a minimum sample of 10 times the number of structural paths (50 samples) (19). Furthermore, data analysis and modeling were carried out using SmartPLS 4. The categories of success factors and application of LVC were used as latent variables, while the determining factors for success and success criteria were used as indicators in the PLS model.

The criteria adopted to assess the measurement model in Figure 2 are as follows:

- a. **Indicator reliability**
The acceptance criteria are outer loading values above 0.7. This is an iterative process, where the indicators with outer loading value below 0.7 will be eliminated (19).
- b. **Internal consistency**
The acceptance criteria of the constructs were based on the composite reliability value of above 0.7 and Cronbach's alpha of above 0.6. A high Cronbach's alpha value indicates the indicators have similar range and meaning (19)
- c. **Convergent validity**
Convergent validity are subtypes of validity that is tested to assess the construct validity (20). In order to achieve convergent validity, each construct should account for at least 50% of the average variance explained (AVE ≥ 0.5) (19).
- d. **Discriminant validity**
Discriminant validity is another sub type of validity used to assess the construct validity (20). By establishing discriminant validity, it is implied that a construct is unique and captures phenomena not represented by other constructs in the model. The assessment is carried out by examining the cross loadings of the indicators. An outer loading value on an associated construct should be greater than all of its loadings on other constructs (19).

The criteria adopted to assess the structural model in Figure 2 are as follows:

- a. **R-squared**
R-squared or coefficient of determination evaluates the variation of the dependent variables (endogeneous) that were explained by the independent variables (exogeneous) and one of the most common method used to measure the inner

model (18, 19). Based on theoretical concepts, R² value of <0.3 indicates weak, $\geq 0.3-0.6$ as moderate, and >0.6 as substantial relationship between variables.

- b. **Path coefficient**
Path coefficient value describes the magnitude of the direct influence among latent variables in the SEM model. The path coefficient value ranges from -1 to +1. A value close to -1 indicates a negative direct relationship between latent variables, while a value close to +1 indicates a positive relationship (18). P value of <0.05 shows significance of the relationship between latent variables.

4. RESULTS AND DISCUSSION

This study identified 40 success factors and 14 criteria for successful implementation of land value capture. Based on the 40 identified success factors, they were grouped into 5 success factor categories consisting of X1 - Government Policy (9 success factors), X2 - Toll Road Company Business Model (8 success factors), X3 - Asset/Property Management (3 success factors), X4 - Favorable Investment Environment (12 success factors), and X5 - Land Value Planning and Specific Project Conditions (8 success factors).

The identified success factors, success factor categories, and success criteria were then validated using an expert judgment questionnaire in the context of SOE Assignment toll roads. The identified success factors were then ranked based on expert assessments through a series of Delphi Method questionnaires to obtain the determinants of success based on expert consensus. This consensus was reached after four rounds of the Delphi Method questionnaire, due to several changes in each round of assessment by the experts involved. The top five success factors in each category are the factors that most significantly contribute to the implementation of land value capture on the Trans-Sumatra Toll Road as a SOE Assignment project (20). The success factors validated and ranked in five categories are presented in Table 2, while the success criteria are presented in Table 3.

Based on the results of the analysis of the Critical Success Factor rankings that determine the success of toll road operations using the assignment scheme in accordance with Table 2 above, the top 5 (five) rankings are as follows:

TABLE 2. Ranked validated success factor and categories on LVC implementation

Success Factor Categories	Success Factor	Relative Importance Index (RII)	Rank	Reference
X1 - Government Policy	X1.8 Concession agreements that are clear and properly describe the obligations of the government and private parties	0.94	1	(3)

	X1.4 Good division of authorities between government agencies	0.89	2	(5)
	X1.1 Collaboration and coordination framework between stakeholders	0.89	3	(3, 16)
	X1.2 Government support (financial, zoning, and land acquisition)	0.86	4	(3, 16)
	X1.5 Preparation of transportation infrastructure investment packages	0.86	5	(16)
	X1.3 LVC implementing agency authority	0.83	6	(16)
	X1.9 Long-term development vision	0.83	7	(3)
	X1.6 Technical and managerial competency/capability of government institutions	0.83	8	(5)
	X1.7 Risk mitigation plan	0.83	9	(3, 5, 18)
	X2.5 Toll Road Company's competency/ capability in planning business models	1.00	1	(3, 4)
	X2.1 Toll Road Company's competency/ capability in the property market (real estate) and its development process	0.91	2	(16)
	X2.7 Toll Road Company's competency/ capability in monetizing accessibility as an asset	0.91	3	(19)
X2 - Toll Road Company's Business Model	X2.3 Toll Road Company's that are oriented towards economic growth in harmony with transportation infrastructure and land development, through collaboration with property developers	0.89	4	(16)
	X2.6 Toll Road Company's competency/ capability in asset management	0.89	5	(3)
	X2.2 Long-term income-oriented LVC contract arrangement	0.83	6	(16)
	X2.8 Toll Road Company's competency/ capability in aligning stakeholder interests in various project phases	0.83	7	(3)
	X2.4 Land acquisition cooperation and land development revenue sharing with the government	0.77	8	(16)
X3 - Asset/ Property Management	X3.3 Property developer competency/ capability	0.94	1	(3, 16)
	X3.1 LVC implementation transparency	0.89	2	(18, 19)
	X3.2 Asset management collaboration with property developers	0.80	3	(3, 16)
	X4.4 Availability of supporting regulations	0.94	1	(5, 18)
	X4.7 Supporting legal framework	0.91	2	(4, 5, 20)
	X4.8 Clear and fair regulations and agreements in sharing costs, benefits, and risks (cost, benefit, risk sharing) between stakeholders	0.91	3	(3, 20)
	X4.10 Healthy and profitable economic policies	0.91	4	(21, 22)
	X4.11 Good governance	0.89	5	(21)
X4 - Favorable Investment Environment	X4.5 LVC scheme transparency	0.83	6	(5, 16)
	X4.12 Public/ private sector commitment and responsibility	0.83	7	(21)
	X4.1 Supporting spatial, transportation, and zoning planning	0.83	8	(3, 16)
	X4.6 Strong political support	0.83	9	(3, 4, 18)
	X4.2 Supportive macroeconomic conditions and good economic growth	0.83	10	(3, 16)
	X4.9 Stable political system	0.80	11	(21, 22)
	X4.3 Financing package that includes land management around the transportation infrastructure	0.77	12	(16)
X5 - Land Value Capture Planning and Specific Project Conditions	X5.5 Profitable business model	0.91	1	(3, 18)
	X5.7 The application of LVC is supported by strategic objectives and regional and national long-term development plans	0.91	2	(3)
	X5.3 Quality transportation infrastructure operational services	0.91	3	(16)
	X5.2 Beneficiary stakeholder support	0.89	4	(4, 5, 16, 18)

X5.8 LVC-based property development has integration with the main transportation infrastructure	0.89	5	(3)
X5.4 Contextual LVC that fits the regional context and meets needs (including property, industrial estate, and tourism)	0.86	6	(18)
X5.1 Good value creation	0.83	7	(4, 5, 16)
X5.6 Potential application of a combination of LVC instruments	0.83	8	(3, 16)

TABLE 3. Five Highest Ranked Critical Success Factors in LVC implementation

ID	Critical Success Factor	RII	Rank	Description
X2.5	Toll Road Company's competency/ capability in planning business models	1,00	1	Business Model Improvement
X1.8	Concession agreements that are clear and properly describe the obligations of the government and private parties	0,94	2	Government Policy
X3.3	Property developer competency/ capability	0,94	3	Business Model Improvement
X4.4	Availability of supporting regulations	0,94	4	Government Policy
X5.5	Profitable business model	0,91	5	Business Model Improvement

TABLE 4. Success criteria of LVC implementation

Successful LVC Implementation	Success Criteria	Reference
Y1 – Successful LVC Implementation	Y1.1 Improving project feasibility without increasing toll (maintaining the affordability of PPP projects, subsidizing the lack of toll-based revenue)	(18)
	Y1.2 Increasing property values around transportation infrastructure (creating value, encourages the value cycles)	(3)
	Y1.3 Become an additional source of income during the construction and operational-maintenance period	(3-5)
	Y1.4 LVC instruments as an adequate and stable source of income	(16)
	Y1.5 Improving market efficiency by better linking costs (charged) and benefits of developing transportation infrastructure	(3)
	Y1.6 Increasing daily traffic volume	(3)
	Y1.7 Reducing dependence on government (and SOE) financing	(3)
	Y1.8 Enabling completion of projects that were not previously possible	(3)
	Y1.9 Access resources (financial, expertise, innovation) from private sector partners (property developers)	(3)
	Y1.10 LVC as a risk sharing mechanism	(3)
	Y1.11 Job creation	(3)
	Y1.12 Encouraging business growth and economic development	(3)
	Y1.13 Applying the beneficiary principle (benefit-received principle), in which value capture is applied to the incremental value beneficiaries	(3, 5)
	Y1.14 Contributing to development of the LVC implementation area	(5, 18)

After the critical success factors and the success criteria were finalized, a pilot survey was conducted to ensure the clarity of the survey indicators, resulting in no adjustment required to the questionnaire. This study received 56 questionnaires; 6 responses did not meet the respondent criteria, resulting in 50 appropriate responses to be used in the model. The PLS model is presented in Figure 3.

Structural equation modeling (SEM) was used to examine the hypothesized model based on Figure 3. The

measurement model in this study consists of 37 indicators/ measurement in six construct/ latent variables, as shown in Figure 3.

Table 5 shows the result of the measurement model assessment. The shown indicators on Table 5 were reduced based on the indicator reliability assessment. The iterative assessment was carried out three times and eliminated 12 unreliable indicators, resulting in 25 reliable indicators. Table 6 shows the result of R-squared value among variables. Table 7 shows the path

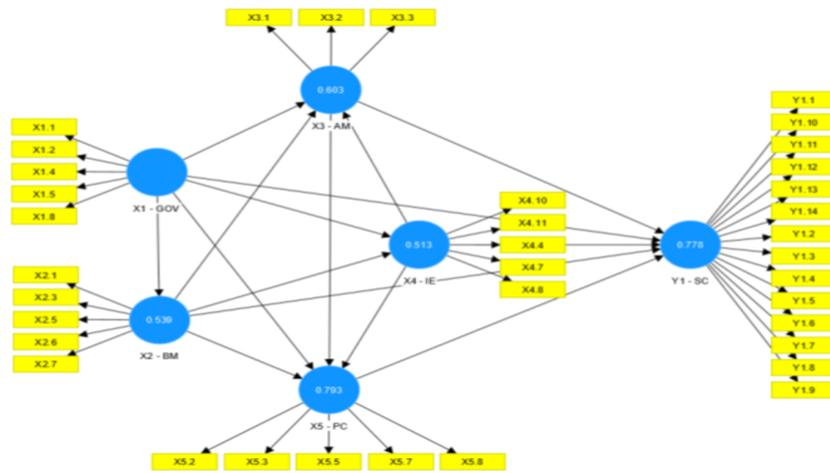


Figure 3. Partial Least Square Model

TABLE 5. Measurement Model Assessment Result

Latent Variables	Indicators	Outer Loading	Composite Reliability	Cronbach's Alpha	Average Variance Extracted (AVE)
Government Policy (GOV)	X1.2	0.779	0.872	0.807	0.630
	X1.4	0.822			
	X1.5	0.778			
	X1.8	0.795			
Toll Road Company's Business Model (BM)	X2.5	0.785	0.902	0.835	0.754
	X2.6	0.914			
	X2.7	0.901			
Asset/ Property Management (AM)	X3.1	0.754	0.836	0.709	0.631
	X3.2	0.716			
	X3.3	0.902			
Favorable Investment Environment (IE)	X4.4	0.905	0.919	0.868	0.791
	X4.7	0.855			
	X4.10	0.908			
Land Value Capture Planning and Specific Project Conditions (PC)	X5.2	0.728	0.914	0.882	0.681
	X5.3	0.842			
	X5.5	0.800			
	X5.7	0.881			
	X5.8	0.866			
Successful Land Value Capture Implementation (SC)	Y1.4	0.720	0.906	0.879	0.580
	Y1.5	0.739			
	Y1.9	0.745			
	Y1.10	0.764			
	Y1.11	0.738			
	Y1.12	0.745			
	Y1.14	0.868			

TABLE 6. R-Squared

Latent Variable	R ²	Result
X1 - GOV		
X2 - BM	0.374	Moderate
X3 - AM	0.597	Moderate
X4 - IE	0.399	Moderate
X5 - PC	0.769	Substantial
Y1 - SC	0.721	Substantial

coefficient value of the model, resulting in three negative results on the hypothesis testing on H1d (X1→X5), H3 (X3→Y1), and H4a (X4→X3). These 3 paths were then eliminated to determine the relationship between influential variables in the model to represent the relationship between the critical success factors on implementing land-value-capture-based SOE assignment toll road project in Indonesia. The result of the second path coefficient evaluation is shown in Table 7.

This study results in 18 critical success factors in five categories and seven success criteria in successful LVC implementation on SOE-assignment Trans Sumatera Toll Road. Figure 4 is the final model with R2 value in each construct, path coefficient, and p value. On each indicator among the six constructs are the indicator loading and p value.

Government policy has a pivotal role in implementing land value capture as an alternative financing scheme on the SOE assignment Trans Sumatera toll road. The existence of government support in terms of financial, zoning, and land acquisition support (X1.2) has a key role in the successful implementation of land value capture in Hong Kong, Guangzhou, and Hyderabad (3, 8). Zoning support would optimize the value capture process in the LVC catchment area throughout the Trans-Sumatera Toll Road. The LVC catchment area needs to be studied further because the development of toll road infrastructure has a more significant impact on increasing accessibility compared to transit infrastructure through its function as logistics distribution routes. On the other hand, land acquisition is one of the issues in infrastructure provision, with a contribution of 29% to the cost of procuring/providing national strategic projects in Indonesia.

Regulatory aspects and division of authority are some of the challenges in infrastructure financing in Indonesia [5] due to overlapping regulations at various levels of government. This emphasizes the importance of good division of authority between government agencies (X1.4) in implementing land value capture. An example of the division of authority in regional development in Indonesia is in the Jakarta MRT transit-oriented-development area (TOD), where Mass Rapid Transit Jakarta (a regionally-owned-enterprise) as the main

TABLE 7. Path Coefficients results for hypothesis (1)

Hypothesis	Paths	Path Coefficients	P Value	Result
H1	X1 - GOV -> Y1 - SC	0.271	0.099	Positive, insignificant
H1a	X1 - GOV -> X2 - BM	0.612	0.000	Positive, significant
H1b	X1 - GOV -> X3 - AM	0.413	0.008	Positive, significant
H1c	X1 - GOV -> X4 - IE	0.106	0.545	Positive, insignificant
H1d	X1 - GOV -> X5 - PC	-0.042	0.789	Negative, insignificant
H2	X2 - BM -> Y1 - SC	0.118	0.670	Positive, insignificant
H2a	X2 - BM -> X3 - AM	0.490	0.001	Positive, significant
H2b	X2 - BM -> X4 - IE	0.562	0.003	Positive, significant
H2c	X2 - BM -> X5 - PC	0.356	0.025	Positive, significant
H3	X3 - AM -> Y1 - SC	-0.028	0.834	Negative, insignificant
H3a	X3 - AM -> X5 - PC	0.204	0.127	Positive, insignificant
H4	X4 - IE -> Y1 - SC	0.169	0.284	Positive, insignificant
H4a	X4 - IE -> X3 - AM	-0.067	0.699	Negative, insignificant
H4b	X4 - IE -> X5 - PC	0.491	0.000	Positive, significant
H5	X5 - PC -> Y1 - SC	0.449	0.029	Positive, significant

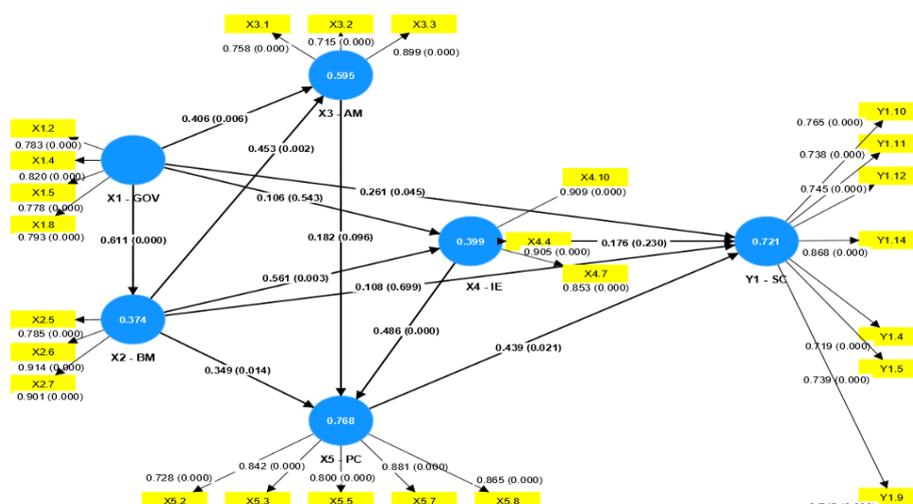


Figure 4. Final model with path coefficient, loadings, R², and p value

operator is given the authority to manage the TOD area. Determining authority in regional development requires an understanding between the parties involved so that the government plays a role in preparing and enforcing a memorandum of understanding between relevant institutions in developing land value capture-based areas on SOE-assignment toll roads. A clear division of authority can prevent conflicts between toll road companies and government institutions in managing the value capture process in the catchment areas.

Another form of government support is a sound investment package (X1.5) on infrastructure investment that allows LVC implementation as an alternative financing scheme. The clarity of the concession agreement (X1.8) is also a critical success factor in implementing land value capture on the SOE-assignment Trans-Sumatera toll road. The agreement clarity on LVC-based infrastructure is a must, due to the various parties involved such as government institutions, state-owned enterprises, and property developers.

In implementing land value capture on SOE-assignment toll roads, the toll road company's business model as the concession company needs to include the function of sustainable area development as a new source of income and to increase average daily traffic. The Regional development may have a lengthy time framework, so the toll road company requires a competency/ capability in asset management (X2.6) to maximize recurring income from property development in the LVC catchment area as a long-term business portfolio. The business model applied by Trans-Sumatera toll road company for the Bakauheni – Terbanggi Besar section must be based upon the monetization of increased accessibility from toll road investment (X2.7), as well as facilitating regional economic growth based on the existing potential (X2.5). Currently, the region has

various potentials, such as industrial areas in South Lampung Regency and Central Lampung Regency, agriculture-related product industry in South Lampung Regency and East Lampung Regency, and tourism-and-maritime-based industry in South Lampung Regency.

Another construct required to successfully implement LVC is asset/ property management. LVC-based area development is based on the joint creation and sharing of land value increments principle (3) and hence requires transparency in its implementation (X3.1), especially regarding the distribution of roles and benefits received. In general, granting land concession rights from the government to the private sector as property developers has a negative public prejudice (3).

Case studies of LVC implementation in Hyderabad and Guangzhou (3, 8) revealed the importance of property developer competence in area development (X3.3), where concession companies in both cities already have competency in area/property development. Another case study in Hong Kong reveals the importance of asset management collaboration between concession companies and property developers (X3.2). As a concession company, Hong Kong MTR cooperates with property developers as a risk mitigation measure in developing the area by accessing the competence and resources of property developer partners in the fields of finance, expertise, and innovation (3, 12). Such asset management collaboration has been developed on the Bakauheni–Terbanggi Besar section to develop Bakauheni Harbor City as an integrated tourism area of 214 ha by involving a joint venture between the Lampung provincial government and the state-owned companies in various sectors, such as sea transportation, construction, and tourism.

The supportive investment environment construct consists of the availability of regulations and legal

frameworks (X4.4, X4.7), as well as healthy and profitable economic policies (X4.10). In general, LVC-based regional development involves private-sector investment. A supportive investment environment attracts private investment in the development, as well as providing a guarantee for private investment. Currently, several regulations are related to the authority and incentives for regional development in Indonesia, mainly on TOD.

Though there are extended development rights for Trans-Sumatera toll road company to further develop the region near the toll road right-of-way based on Presidential Decree No 131 of 2022, the decree has not included land value capture as an alternative source of infrastructure financing (2). Private sector investment in LVC-based regional development requires sound and profitable economic policies to attract private sector participation. In Indonesia, the Coordinating Ministry for Economic Affairs has an integral role in economic policy-making, as well as coordinating various government institutions in drafting regulations related to LVC-based regional development to meet the development target in the National Medium-Term Development Plan (2).

LVC planning and specific project conditions are other constructs influencing the successful land value capture implementation on the TSTR as an SOE assignment project with well planning and controlling. A construction project requires tools that can be used for project planning and control in order to provide high performance (23).

On its implementation, LVC required support from beneficiaries (X5.2), as the value-capturing process is imposed on beneficiaries of the value creation based on toll road investment and LVC-based regional development. Opposition from beneficiaries in terms of land parcels acquisition, zoning adjustments, and the imposition of new types of taxation, may hinder the value capture process. LVC implementation on transportation infrastructure must be based on the quality of operational services (X5.3), as the toll road infrastructure and the fulfillment of minimum service standards influence the average daily traffic and value uplift on the LVC catchment area. A profitable business model (X5.5) proposed by the toll road company is another critical success factor required to attract investors to the LVC-based regional development.

The proposed business model for TSTR needs to consider the regulatory support from the government, the toll road company's regional development capability, regional resource potential, and investment environment. Another critical success factor is the support and alignment of the LVC-based regional development with regional and national long-term development plans (X5.7).

Lesson learnt on Hong Kong MTR reveals the role of the Hong Kong government's strategic decision to further increase the integration of mixed-use development areas in the LVC catchment area (3). Steps that can be taken by toll road companies as the concessionaire are to align the regional development with the government's strategic objectives stated in the National Medium Term Development Plan, obtain zoning support, and accommodate regional economic growth based on existing potential in the region. LVC-based property development integrated with the toll road (X5.8) encourages a simultaneous increase in value between the toll road as the main infrastructure, the developed region, and the catchment area. By integrating LVC-based regional development and toll road, the toll road company provides a clearer increase in value for the beneficiaries.

Based on the structural equation model, there are seven success criteria for successful LVC implementation on the SOE-assignment-based Trans-Sumatera toll road. One of the success criteria is an adequate and stable source of income for concession companies (Y1.4) to serve as a source of income and return on investment, both for government investment and toll road companies as the concessionaire. To become a stable source of income, the toll road company needs to collaborate with the government to implement various LVC instruments based on area development by the toll road company and imposing taxes to the beneficiaries in the catchment area by the government authorities. Another success criterion for implementing LVC is risk mitigation for the toll road company (Y1.10), as the toll road company faces low financial viability, depends on government financing, and has low average daily traffic. Through collaboration with property developers, toll road companies obtain access to financial resources, expertise, and property developer innovation (Y1.9), thereby reducing the financial burden and risks in property development.

LVC-based regional development is based on the increased accessibility through the availability of toll road services and existing regional potential that further drive the business growth, and economic and industry development in the region (Y1.12).

5. CONCLUSION

This research has identified Critical Success Factors that influence the implementation of land value capture in toll road with the SOE assignment scheme. By conducting an assessment and gap analysis, it can be seen that it is necessary to identify the top five of Critical Success Factors that must be paid attention to by toll road developer companies. After that, relationship modeling

can also be carried out regarding the factors that influence the success of implementing the land value capture mechanism and use it to identify potential improvements and as a lesson for toll road operations with other assignment schemes in the future.

The results of the Critical Success Factor analysis, it shows that an improvement/innovation of the business model of toll road development companies through property business development and support from government policies is a determining factor in the success of toll road operations with an assignment scheme. The following are the top 5 (five) Critical Success Factors that determine the success of toll road operations using an assignment scheme:

1. Business Entity's capability in planning a business model.
2. A clear concession agreement that clearly describes the obligations of the government and the private sector.
3. Property development competencies and capabilities.
4. Availability of supporting regulations.
5. Profitable business model.

The structural equation model produces 18 determinants of success into five construct categories. Based on the path coefficient, government policy has the most significant influence on the success of implementing land value capture for the TSTR SOE assignment. This study also shows that government policies significantly impact other sectors, especially toll road companies' business models.

This research is limited to case studies of toll road development companies with SOE assignment schemes in Indonesia. Therefore, further research needs to be developed for larger data samples. The limitation of this research is to analyze the gap between critical success factors in achieving successful toll road operations but does not discuss the inhibiting factors in achieving success. Further research is needed to determine the best LVC and regional development instruments to be applied on toll roads with the SOE assignment scheme in Indonesia.

6. AUTHOR'S NOTE

The authors declare that this article is free from plagiarism and has no conflicts of interest regarding its publication.

7. REFERENCES

1. Roukouni A, Macharis C, Basbas S, Stephanis B, Mintsis G. Financing urban transportation infrastructure in a multi-actors environment: the role of value capture. *European Transport Research Review*. 2018;10:1-19. <https://doi.org/10.1007/s12544-017-0281-5>
2. Advisory PPI. *Innovative Infrastructure Financing through Value Capture in Indonesia*. 2021. <https://dx.doi.org/10.22617/SPR200093-2>
3. Suzuki H, Murakami J, Hong Y-H, Tamayose B. *Financing transit-oriented development with land values*: World Bank Publications Washington, DC, USA; 2015.
4. Connolly C, Wall A. Value capture: A valid means of funding PPPs? *Financial Accountability & Management*. 2016;32(2):157-78. <https://doi.org/10.1111/faam.12083>
5. Iskandar DA. Land Value Capture as Financial Resource for Infrastructure Development in Palembang City. *The Journal of Indonesia Sustainable Development Planning*. 2021;2(1):74-83. <https://doi.org/10.46456/jisdep.v2i1.96>
6. Sarwono J, Narimawati U. *Membuat skripsi: Tesis dan Disertasi dengan Partial Least Square SEM (PLS-SEM)*. Yogyakarta ...; 2015.
7. Fakhri MR, Riantini LS, Latief Y, editors. *Risk-based institutional evaluation of trans Sumatera Toll road infrastructure development to improve time performance*. AIP Conference Proceedings; 2023: AIP Publishing.
8. Li X, Love PE, Luo H, Fang W. A systemic model for implementing land value capture to support urban rail transit infrastructure projects. *Transportation Research Part A: Policy and Practice*. 2022;156:90-112. <https://doi.org/10.1016/j.tra.2021.12.010>
9. Asgari M, Kheyroddin A, Naderpour H. Evaluation of project critical success factors for key construction players and objectives. *International Journal of Engineering, Transactions B: Applications*. 2018;31(2):228-40. <http://doi.org/10.5829/ije.2018.31.02b.06>
<http://doi.org/10.5829/ije.2018.31.02b.06>
10. Medda F. Land value capture finance for transport accessibility: a review. *Journal of Transport Geography*. 2012;25:154-61. <https://doi.org/10.1016/j.jtrangeo.2012.07.013>
11. Kurnia RF, Latief Y, Soehodo S, Rarasati AD. *Conceptual Framework of Road Plus Property Developer Business Model on Toll Roads Assignment Scheme to Improve Investment Performance*. EasyChair; 2023. Report No.: 2516-2314.
12. Zhao ZJ, Das KV, Larson K. Joint development as a value capture strategy for public transit finance. *Journal of Transport and Land Use*. 2012;5(1):5-17. <https://www.jstor.org/stable/26201681>
13. Hai DT, Toan NQ, Van Tam N. Critical success factors for implementing PPP infrastructure projects in developing countries: the case of Vietnam. *Innovative Infrastructure Solutions*. 2022;7(1):89. <https://doi.org/10.1007/s41062-021-00688-6>
14. Ahmad U, Waqas H, Akram K. Relationship between project success and the success factors in public-private partnership projects: A structural equation model. *Cogent Business & Management*. 2021;8(1):1927468. <https://doi.org/10.1080/23311975.2021.1927468>
15. Molwus JJ, Erdogan B, Ogunlana S. Using structural equation modelling (SEM) to understand the relationships among critical success factors (CSFs) for stakeholder management in construction. *Engineering, construction and architectural management*. 2017;24(3):426-50. <https://doi.org/10.1108/ecam-10-2015-016>
16. Sihombing L, Latief Y, Rarasati AD, Wibowo A. Developing a toll road financing model with a hybrid of deep discount project bonds and land leases in Indonesia. *Int J of Civil Eng Tech*. 2018;9(6):1310-23. <https://doi.org/10.13140/RG.2.2.32051.89125>

17. Franc JM, Hung KKC, Pirisi A, Weinstein ES. Analysis of Delphi study 7-point linear scale data by parametric methods: use of the mean and standard deviation. *Methodological Innovations*. 2023;16(2):226-33. <https://doi.org/10.1177/20597991231179393>
18. Finstad K. Response interpolation and scale sensitivity: Evidence against 5-point scales. *Journal of usability studies*. 2010;5(3):104-10. <https://www.researchgate.net/publication/265929744>
19. Sarstedt M, Ringle CM, Hair JF. Partial least squares structural equation modeling. *Handbook of market research*: Springer; 2021. p. 587-632.
20. Samani SA, Rasid SZA, Sofian S. The influence of personal control and environmental distraction in open-plan offices on creative outcome. *Performance Improvement Quarterly*. 2017;30(1):5-28. <https://doi.org/10.1002/piq.21239>
21. Muhwezi L, Acai J, Otim G. An assessment of the factors causing delays on building construction projects in Uganda. *International journal of construction engineering and management*. 2014;3(1):13-23. <https://doi.org/10.5923/j.ijcem.20140301.02>
22. Alvanchi A, Shiri N, Alikhani H. In-depth investigation of project planning and control software package application in the construction industry of Iran. *International Journal of Engineering*. 2020;33(10):1817-25. <https://doi.org/10.5829/ije.2020.33.10a.01>
23. Ahmatabadi AA, Heravi G. The effect of critical success factors on project success in Public-Private Partnership projects: A case study of highway projects in Iran. *Transport policy*. 2019;73:152-61. <https://doi.org/10.1016/j.tranpol.2018.07.004>

COPYRIGHTS

©2024 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

جاده عوارضی ترانس سوماترا (TSTR) واقع در عملیات اندونزی که تحت مدیریت شرکت‌های دولتی (SOE) قرار دارد، با چالش‌های متعددی مواجه است و به منابع مالی و درآمد جایگزین نیاز دارد که یکی از آنها مبتنی بر جذب ارزش زمین (LVC) است. توسعه منطقه هدف این تحقیق شناسایی و تجزیه و تحلیل عوامل موفقیت حیاتی مورد نیاز برای اجرای ثبت ارزش زمین در پروژه TSTR است. نتایج این تحقیق پنج عامل موفقیت با بالاترین رتبه را به دست آورد و مدلی از رابطه بین متغیرها در اجرای تصرف ارزش زمین در عملیات تخصیص عوارض راهداری شرکت‌های دولتی بر اساس اجماع متخصصین به دست آمد. از طریق مطالعات ادبیات، ۴۰ عامل موفقیت در پنج دسته و ۱۴ معیار برای اجرای موفقیت‌آمیز ثبت ارزش زمین در زیرساخت‌های مبتنی بر جاده‌ای که کارشناسان اعتبارسنجی کردند، گروه‌بندی شدند. فاکتورهای موفقیت تایید شده از طریق یک سری ارزیابی کارشناسان با استفاده از پرسشنامه دلفی-متد پردازش شدند که منجر به پنج عامل موفقیت با بالاترین رتبه در هر دسته شد. روابط بین متغیرها بیشتر از مدل سازی PLS-SEM به دست آمد و مورد تجزیه و تحلیل قرار گرفت. تجزیه و تحلیل مدل رابطه روابط بین متغیرهایی از جمله سیاست دولت، مدل کسب و کار توسعه دهنده جاده عوارضی، مدیریت دارایی/املاک، محیط حمایت از سرمایه گذاری، برنامه ریزی LVC و شرایط پروژه خاص را ایجاد کرد. نتایج مطالعه حاضر ممکن است عواملی را تعیین کند که می‌توانند از تکمیل موفقیت‌آمیز پروژه‌های جاده‌ای با عوارض با طرح تخصیص SOE اطمینان حاصل کنند.



Accurate Analytical Modeling of Drain Current of Heterojunction Tunneling Field Effect Transistor

F. Peyravi, S. E. Hosseini*

Department of Electrical Engineering, Ferdowsi University of Mashhad, Iran

PAPER INFO

Paper history:

Received 10 October 2023

Received in revised form 13 December 2023

Accepted 31 December 2023

Keywords:

Analytical Model

Heterojunction Tunneling Field Effect Transistor

Band-to-band Tunneling

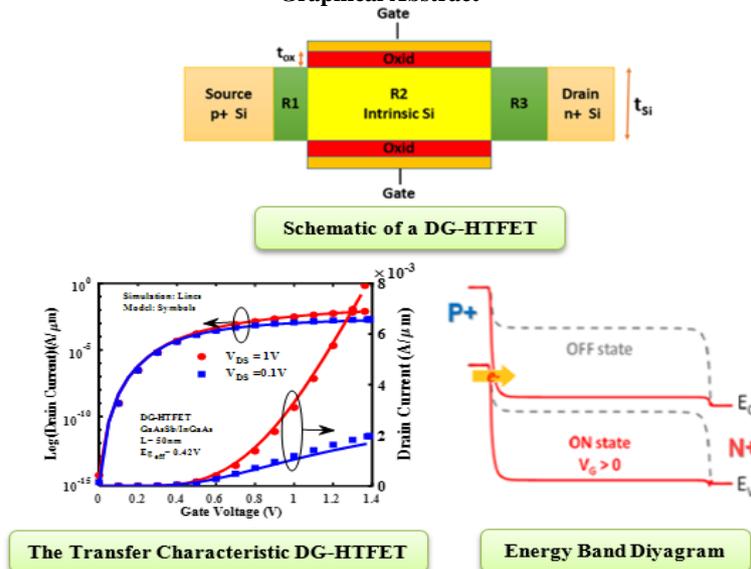
Tangent Line Approximation

ABSTRACT

An accurate analytical model is presented for drain current of the heterojunction tunneling field effect transistor, taking into account the source depletion region, mobile charges and the effect of the drain voltage. This model accurately predicts the potential distribution not only on the surface but also within the semiconductor depth by utilizing newly formulated mathematical relationships. Using the tangent line approximation method and considering the channel region as well as the source depletion region' We analytically calculate the band-to-band tunneling current from the source to the channel by integrating the tunneling generation rate. Compared to simulation results, the proposed model demonstrates significant accuracy in predicting drain current.

doi: 10.5829/ije.2024.37.07a.12

Graphical Abstract



1. INTRODUCTION

In recent years, the Tunneling Field-Effect Transistor (TFET) has garnered significant attention due to its

ability to achieve a sub-threshold swing of less than 60 mV/dec, which is a fundamental limitation in Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs). This characteristic positions TFET as a

*Corresponding Author Email: ehosseini@um.ac.ir (S. E. Hosseini)

promising alternative to MOSFETs in low-power applications and future electronic devices technology. Additionally, TFET demonstrates low leakage current and increased immunity against short-channel effects (SCEs) (1-4).

To facilitate circuit design utilizing TFETs, the development of an accurate analytical model for the drain current is of paramount importance (5). A robust analytical model should rely on physical phenomena, enabling the expression of device characteristics through mathematical equations. Numerous numerical and analytical models have been proposed for TFETs in the existing literature (6, 7). While numerical and semi-analytical models offer reasonably accurate results; their efficiency in circuit-level simulations is limited. Conversely, analytical models are compatible with a wide range of circuit-level simulation tools, making them more practical for this purpose. Analytical models for TFETs generally rely on surface potential computation. However, some models have previously utilized the average electric field across the tunneling junction to calculate the tunneling generation rate. This approach, unfortunately, has limited the accuracy of the model (8, 9). To address this limitation, it is recommended to consider the complete electric field profile along the tunneling junction. By incorporating the full electric field profile, the model accuracy can be significantly improved.

Tangent line method is adopted for drain current calculation to consider total generation profile (10-12). While these models take into account the electric field profile across the tunneling junction to enhance the accuracy of the model, they primarily focus on the electric field along the channel depletion region and overlook the depletion region in the source region.

Furthermore, in order to calculate the generation rate, the electric field at the semiconductor surface is determined, assuming that the electric field throughout the depth of the semiconductor layer remains constant and equal to the surface electric field. However, in reality, this assumption is rarely valid in actual devices (13). Artificial intelligence can be useful in optimizing analytical models and the control of various systems as a powerful tool at the global level (14).

In this paper, we proposed an accurate two-dimensional (2D) solution to Poisson equation in the channel of the TFET.

In addition to the surface potential, we considered the potential distribution in the semiconductor depth, and therefore, we were able to calculate the potential more accurately than reported literature, by utilizing newly formulated mathematical relationships.

This solution provides an accurate prediction of the electric field not only at the surface but also throughout the depth of the semiconductor layer. Finally, we developed an analytical model for drain current using the tangent line approximation method, which considers the

source depletion region. Comparing with simulation results are shown that the proposed model has impressive accuracy.

This paper is organized as follows: In section 2, the device structure under consideration is introduced. In section 3, an analytical surface potential model is proposed in detail. In section 4, an analytical model for drain current by applying the tangent line approximation method is expressed. The results of analytical model and comparison with simulation results are presented in section 5.

2. DEVICE STRUCTURE

The schematic diagram of the H-TFET is depicted in Figure 1. In this device, two depletion regions are formed at the source-channel junction and the channel-drain junction. At a certain value of the gate voltage V_{GS} , the valence band in the source is aligned with the conduction band in the channel, and electrons can tunnel from the source valence band into the channel conduction band through the potential barrier formed by the bandgap. Device specifications are summarized in Table 1 (15-17).

3. ANALYTICAL MODEL DESCRIPTION

Analytical modeling of TFET typically involves three steps. The first step is calculating the surface potential of the transistor. The second step is determining the tunneling generation rate, and in the third step, integrating the tunneling rate to calculate drain current.

3. 1. Surface Potential and Electric Field Model

A pseudo-2D method together with the parabolic

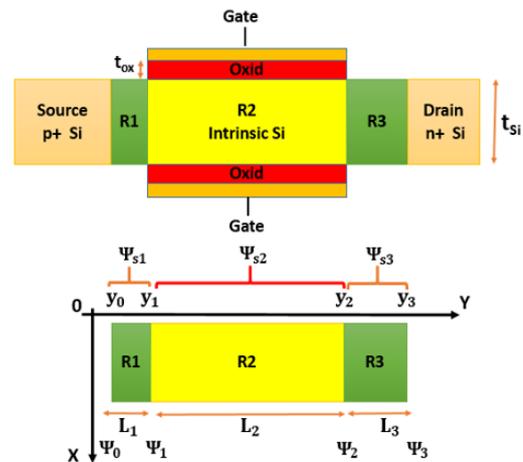


Figure 1. schematic diagram of the proposed double gate heterojunction tunnel field effect transistor (DG-HTFET) considering the areas included (source depletion region R1, channel region R2, drain depletion region R3)

TABLE 1. Device parameters used in simulations and analytical model

Physical parameters	Value
Gate length (L_G)	50 nm
Source length (L_S)	30 nm
Drain length (L_D)	30 nm
Source doping (p-type)	10^{20} cm^{-3}
Drain doping (n-type)	$5 \times 10^{18} \text{ cm}^{-3}$
Channel doping	10^{17} cm^{-3}
Channel thickness (t_{si})	10 nm
Oxide thickness (t_{ox})	2.3 nm
Permittivity of Si (ϵ_{Si})	$11.68 \epsilon_0$
Permittivity of SiGe (ϵ_{SiGe})	$13.9 \epsilon_0$
Permittivity of GaAsSb (ϵ_{GaAsSb})	$14.3 \epsilon_0$
Permittivity of InGaAs (ϵ_{InGaAs})	$13.9 \epsilon_0$
Permittivity of SiO ₂ (ϵ_{ox})	$3.9 \epsilon_0$
Work function of gate material (ϕ_m)	4.5, 4.8 eV
Kane's parameter (A_{BTBT})	$3.5 \times 10^{21} \text{ eV}^{1/2}/\text{cm.s.V}^2$
Kane's parameter (B_{BTBT})	$22.5 \times 10^6 \text{ V/cm.eV}^{3/2}$

potential approximation for solving Poisson equation is used to obtain a 2D potential profile. The Poisson equation given below:

$$\frac{\partial^2 \Psi(x,y)}{\partial x^2} + \frac{\partial^2 \Psi(x,y)}{\partial y^2} = \frac{-qN_s}{\epsilon_s} \quad (1)$$

In this equation, $\Psi(x,y)$ is the electrostatic potential in the device, which is measured with respect to substrate Fermi level, ϵ_s is the semiconductor permittivity, q is electron charge, and N_s is effective doping density. N_s is equal to $-N_1$ in p-type source, and is equal to N_2 and N_3 in n-type channel and drain regions, respectively. In Equation 1 mobile charges are ignored and its effect is considered in section B.

Pseudo 2-D method is adopted to convert the 2-D Poisson equation into an effective 1-D equation which is further used to obtain the potential distribution along the thickness of the device. Poisson equation is solved in three regions, i.e., the source depletion region (R1), the low doped channel (R2) and the drain depletion region (R3) using proper boundary conditions. Assuming parabolic solution for the Poisson equation, the 2-D potential distribution is written as (15, 18).

$$\Psi(x,y) = a_0(y) + a_1(y)x + a_2(y)x^2 \quad (2)$$

Coefficients $a_0(y)$, $a_1(y)$ and $a_2(y)$ in Equation 2 are computed using boundary conditions, i.e., continuity of the potential and the electric displacement at the insulator-semiconductor interfaces. The potential distribution at the front-oxide-channel interface ($x=0$), and back-oxide-semiconductor interface ($x = t_s$) are denoted as $\Psi_s(y)$ and $\Psi_b(y)$, respectively.

$$\Psi(0,y) = \Psi_s(y) \quad (3a)$$

$$\Psi(t_s,y) = \Psi_b(y) \quad (3b)$$

The electric field displacement at front and back surfaces are continuous and are expressed as follows:

$$E_x(0,y) = \frac{C_{ox}}{\epsilon_s} [\Psi_s(y) - \Psi_G] \quad (3c)$$

$$E_x(t_s,y) = \frac{C_{ox}}{\epsilon_s} [\Psi_G - \Psi_b(y)] \quad (3d)$$

where C_{ox} demonstrates the oxide capacitances at each region, which is $C_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$ in the R2 region. In this analytical model, the fringing fields effect is approximated by the conformal mapping techniques in R1 and R3 regions and is considered as $C_{ox} = \frac{2}{\pi} \times \frac{\epsilon_{ox}}{t_{ox}}$ (19). The gate potential is demonstrated as Ψ_G , where $\Psi_G = V_{GS} - \phi_m + \chi + E_g/2$, χ is the electron affinity of silicon, E_g is the energy bandgap of silicon, and ϕ_m is work function of the gate metal.

By applying boundary conditions, coefficients $a_0(y)$, $a_1(y)$ and $a_2(y)$ are calculated as functions of the surface potential $\Psi_s(y)$:

$$a_0(y) = \Psi_s(y) \quad (4a)$$

$$a_1(y) = \frac{C_{ox}}{\epsilon_s} (\Psi_s(y) - \Psi_G) \quad (4b)$$

$$a_2(y) = \frac{C_{ox}}{\epsilon_s t_s} (\Psi_G - \Psi_s(y)) \quad (4c)$$

By substituting coefficients 4 in Equation 2, the two-dimensional Poisson equation becomes a second-order one-dimensional linear differential equation written as follows:

$$\Psi_s''(y) - 2 \frac{C_{ox}}{\epsilon_s t_s} \Psi_s(y) = -2 \frac{C_{ox}}{\epsilon_s t_s} \Psi_G - \frac{qN_s}{\epsilon_s} \quad (5)$$

This equation can be written for three regions as:

$$\Psi_{s,i}''(y) - k_i^2 \Psi_{s,i}(y) = -k_i^2 \Psi_{d,i} \quad (6)$$

$$k_i = (2 \frac{C_{ox}}{\epsilon_s t_s})^{0.5} \quad (7)$$

where $i = 1, 2$ and 3 denote R1, R2 and R3 regions and k_i and $\Psi_{d,i}$ are coefficient values in i -th region. The parameter $\Psi_{d,i}$ is equal to the solution of one-dimensional Poisson equation (long-channel approximation), and is expressed as Equation 8 for $i=2, 3$:

$$\Psi_{d,i} = \Psi_G + \frac{qN_s t_s}{2C_{ox}} \quad (8)$$

and for $i=1$ as expression 9.

$$\Psi_{d,i} = \Psi_G - \frac{qN_s t_s}{2C_{ox}} \quad (9)$$

k_i indicates reverse decay lengths or the characteristic lengths of the surface potential in each region. According

to Equation 7, k_i is not a function of V_{GS} . Finally, the general solution of 2D potential in Equation 6 for the surface potential is obtained. The surface potential in region R_i ($y \in [y_i, y_{i-1}]$) is obtained as follows:

$$\Psi_{si}(y) = b_i e^{k_i(y-y_{i-1})} + c_i e^{-k_i(y-y_{i-1})} + \Psi_{di} \quad (10)$$

To obtain the coefficients b_i and c_i in Equation 10, suitable boundary conditions are considered at the R_1/R_2 and R_2/R_3 boundaries, which indicate continuity of potential and electric displacement.

$$b_i = \frac{1}{2 \sinh k_i L_i} (-\Psi_{di}(1 - e^{-k_i L_i}) - \Psi_{i-1} e^{-k_i L_i} + \Psi_i) \quad (11)$$

$$c_i = \frac{1}{2 \sinh k_i L_i} (+\Psi_{di}(1 - e^{k_i L_i}) + \Psi_{i-1} e^{k_i L_i} - \Psi_i) \quad (12)$$

In Equations 11 and 12 $L_i = y_i - y_{i-1}$ for $i = 1, 2, 3$ and the terms Ψ_i and Ψ_{i-1} express potential values at the boundary of each region. According to these equations $\Psi_{si}(y) = \Psi_i$, using the following boundary conditions, the coefficients b_i and c_i are obtained (8):

$$\Psi_0 = b_1 + c_1 + \Psi_{d,1} \quad (13a)$$

$$\Psi_1 = b_1 e^{k_1 L_1} + c_1 e^{-k_1 L_1} + \Psi_{d,1} = b_2 + c_2 + \Psi_{d,2} \quad (13b)$$

$$\Psi_2 = b_2 e^{k_2 L_2} + c_2 e^{-k_2 L_2} + \Psi_{d,2} = b_3 + c_3 + \Psi_{d,3} \quad (13c)$$

$$\Psi_3 = b_3 e^{k_3 L_3} + c_3 e^{-k_3 L_3} + \Psi_{d,3} \quad (13d)$$

The surface potential values at the source end and at the drain end are Ψ_0 and Ψ_3 , respectively.

$$\Psi_0 = -v_t \ln \frac{N_1}{n_i} \quad (14)$$

$$\Psi_3 = V_D + v_t \ln \frac{N_3}{n_i} \quad (15)$$

where V_D demonstrates the drain voltage, v_t is the thermal voltage $v_t = kT/q$ and the n_i represent the intrinsic carrier concentration.

3. 2. The Effect of Channel Mobile Charges and the Drain Voltage

In this section, the mobile charges are considered to obtain an accurate analytical model for the potential. To consider the effect of channel mobile charges, the new mid-surface potential and the new inverse decay length are obtained as follows:

$$\Psi_{d,2new} = \frac{1}{2} (\Psi_{d,2} + \Psi_\tau - \sqrt{(\Psi_{d,2} - \Psi_\tau)^2 + \delta^2}) \quad (16)$$

where $\Psi_{d,2}$ is the mid-surface potential that does not include the effect of mobile charges. Ψ_τ is given below where α and β are the fitting parameters. δ and N_{tran} are considered as: $\delta = 0.04$ and $N_{tran} = 10^{14} \text{ cm}^{-3}$ (16).

$$\Psi_\tau = V_D + \varphi + \alpha(\Psi_{d,2} - V_D - \varphi) + \beta(\Psi_{d,2} - V_D - \varphi)^2 \quad (17)$$

$$\varphi = v_t \ln \frac{N_2 N_{tran}}{n_i^2} \quad (18)$$

The effect of mobile charges on the reverse decay length in the channel (k_2) is modeled as follows:

$$k_2 = \sqrt{2 \frac{C_{ox}}{\epsilon_{si} t_{si}} - \frac{\sigma N_{inv}}{\epsilon_{si} t_{si} (\Psi_0 - \Psi_{d,2})}} \quad (19)$$

$N_{inv} = 2C_{ox}(\Psi_g - \Psi_{d,2})$ represents the inversion charge in the channel and σ is a fitting parameter, $1 < \sigma < 2$.

To accurately model the potential in short channel devices, it is important to consider the impact of high drain voltages. These voltages lead to the formation of an inversion layer in the channel, which affects the potential distribution. Therefore, to enhance the accuracy of the model, the new gate potential $\Psi_{G,eff}$ is used instead of Ψ_G (7), where $\alpha = 0.6$ and $\theta_t = 0.1$ are fitting parameters.

$$\Psi_{G,eff} = \Psi_G - \alpha \theta_t \ln[1 + \exp(\frac{\Psi_G - \Psi_3}{\theta_t})] + \alpha \theta_t \ln[1 + \exp(\frac{\Psi_0 - \Psi_G}{\theta_t})] \quad (20)$$

Figure 2, shows the front surface potential along the lateral distance in y-axis at different gate voltages. Obviously, as the gate voltage increases, the potential in the channel region increases. With increasing the gate voltage, the mobile charges accumulate at the channel surface below the gate and gradually an inversion layer is formed; consequently, the channel potential connects to the drain Fermi level, and becomes almost independent of the gate voltage. In addition, by increasing inversion charges, drain depletion region would gradually become smaller; thus, the source depletion region would increase.

3. 3. Modified Surface Potential In the parabolic approximation method used in literature to solve the 2-D Poisson equation, surface potential is usually obtained and tunneling current is computed based on surface potential (15, 16, 18). Although these analytical models

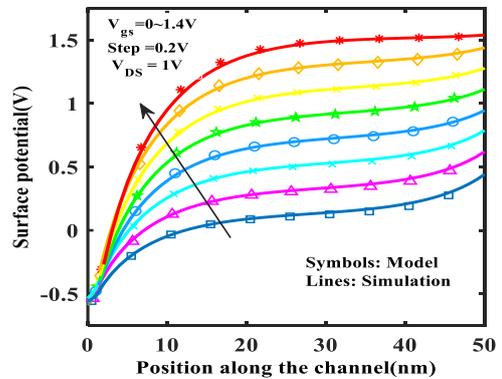


Figure 2. The front surface potential distribution obtained from simulation (line) and analytical model (symbol) at gate voltages from $V_g = 0V$ to $V_g = 1.4V$ in steps of $0.2V$

can predict the surface potential accurately, the model cannot predict the potential profile in the channel depth.

The validation of the surface potential distribution behavior of the model against the TCAD simulation results on the surface and at the different depths of the channel for $V_g = 0, 0.4, 0.8, 1.2 V$ and $V_d = 1V$ are depicted in Figure 3.

The potential distribution is shown inside the channel at different depths inclusive the surface of the channel, 3 nm below the surface, 5 nm below the surface in the center of the channel without considering the modified surface potential.

Figure 3a, shows the surface potential, Figure 3b, shows the potential distribution at 3 nm below the surface in ($y = 3nm$) and Figure 3c, depicts the potential distribution, 5 nm beneath the surface (center of the channel, $y = 5nm$). Note that the semiconductor thickness is 10 nm.

The analytical model well captures the effect of the mobile charges and effect of the drain voltage. Accuracy of the simulation results and analytical model can be simply seen in Figure 3a, while the model results are not consistent with the TCAD simulation data at the depth of 3 nm and 5 nm in Figures 3b and 3c.

In the parabolic approximation method, in fact, the potential is resolved along the path where the current dominates (along the oxide-semiconductor junction); hence, the accuracy of this method is just well along this path while close to the center of the channel, the accuracy decreases.

To achieve 2D potential distribution correctly, Equation 2 is used to convert the 2-D Poisson equation into a second-order one-dimensional linear differential equation in terms of the surface potential. Unlike the models in literature (15, 16, 18) which use $\frac{\partial^2 \psi(x,y)}{\partial x^2}$ and $\frac{\partial^2 \psi(x,y)}{\partial y^2}$ at $x=0$ (surface) an equation for arbitrary x is derived which gives the potential distribution from the surface down to the channel depth. Equation 5 turns into the following equation to include x dependency.

$$-2 \frac{\eta}{t_s^2} (\Psi_s(y) - \Psi_G) + \Psi_s''(y) \left[1 + \frac{\eta}{t_s} x - \frac{\eta}{t_s^2} x^2 \right] = \frac{-qN_s}{\epsilon_s} \quad (21)$$

where $\eta = \frac{C_{ox}}{C_{si}}$ and for simplicity, A' is defined as $A' = 1 + \frac{\eta}{t_s} x - \frac{\eta}{t_s^2} x^2$; therefore,

$$\Psi_s''(y) - \frac{1}{A'} \frac{2\eta}{t_s^2} \Psi_s(y) = -\frac{qN_s}{\epsilon_s} \frac{1}{A'} - \frac{1}{A'} \frac{2\eta}{t_s^2} \Psi_G \quad (22)$$

By defining $k_{i,new}^2 = \frac{1}{A'} \frac{2\eta}{t_s^2}$ and mid-potential $\Psi_{di,new} = \frac{t_s^2}{2\eta} \left(\frac{qN_s}{\epsilon_s} + \frac{2\eta}{t_s^2} \Psi_G \right)$, the differential equation for all regions can be written in a closed form analytical expression for the surface potential as follows:

$$\Psi_{si}''(y) - k_{i,new}^2 \Psi_{si}(y) = -k_{i,new}^2 \Psi_{di,new} \quad (23)$$

$$k_{i,new} = \sqrt{\frac{2\epsilon_{ox}}{\epsilon_{si}t_{ox}t_{si} + \epsilon_{ox}t_{si}x - \epsilon_{ox}x^2}} \quad (24)$$

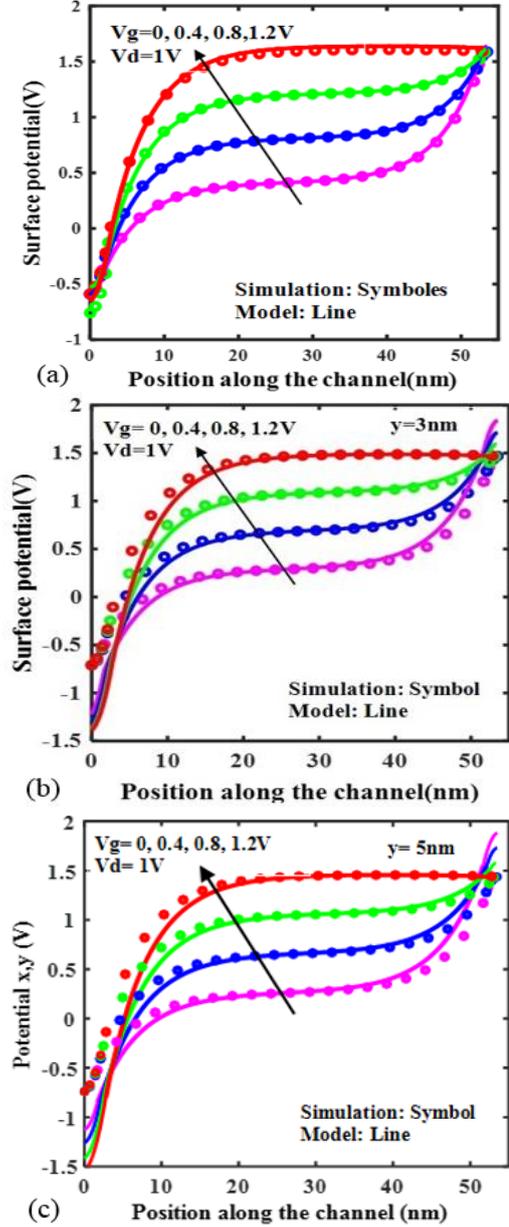


Figure 3. The potential distribution parabolic method is compared with the simulation results. In (a) model shows excellent matching with simulation data in surface potential, here the effect of the mobile charges and effect of the drain voltage is included. (b) displays the potential distribution profile along the channel for 3 nm below the device surface for $V_{ds} = 1V$ and $V_{gs} = 0V$ to $1.2V$ in steps of $0.4V$ and, (c) shows potential distribution in the center of the channel ($y = 5nm$). In the DG-TFET, by moving towards the center of the channel and away from the surface of the semiconductor, the match between the model and the simulation becomes less and the analytical model cannot predict the potential profile well

Equation 6 using new reverse decay length ($k_{i,new}$) in Equation 24 has been modified to represent the potential distribution at any depth of the device.

The electric field is computed by differentiating electrostatic potential with respect to x and y coordinates in each region.

$$E_{yi} = -\frac{\partial \psi(x,y)}{\partial y} = b_i k_i e^{k_i(y-y_{i-1})} - k_i c_i e^{-k_i(y-y_{i-1})} \quad (25)$$

$$E_{xi} = -\frac{\partial \psi(x,y)}{\partial x} = a_1(y) + 2a_2(y)x \quad (26)$$

The electric field along the surface of the 2-D TFET in the x and y directions have been shown in Figure 4. As can be seen, the lateral electric field in the y direction is larger than the vertical electric field in the x direction at the source-channel interface. Hence, for simplicity in calculating the tunneling generation rate, the vertical field can be ignored. The length of the depletion regions on the source and drain sides is indicated by L_1 and L_3 , respectively, and are calculated as follows:

$$L_1 = \sqrt{\frac{2\epsilon_s(\psi_1 + v_t \ln \frac{N_1}{n_i})}{qN_1}} \quad (27)$$

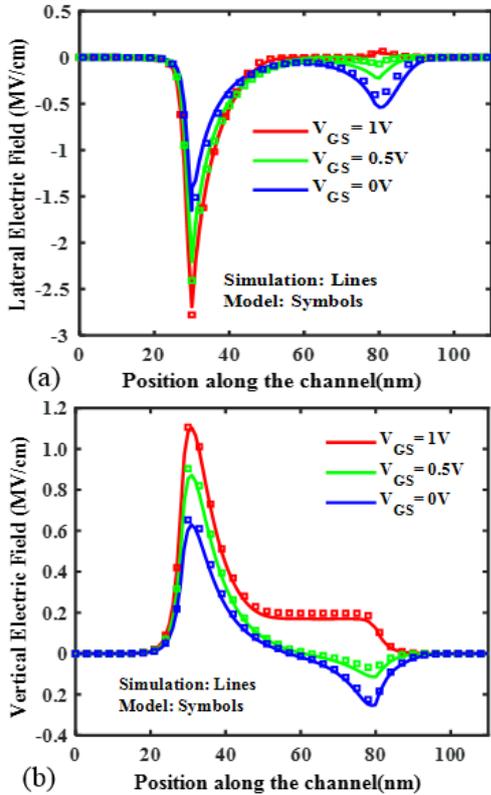


Figure 4. (a) The lateral electric field, (b) the vertical electric field was obtained from the simulator (line) and the analytical model (symbol) at different gate voltages. The dielectric constant SiO_2 is $\epsilon_{ox} = 3.9$, and the drain voltage is $V_D = 1V$

$$L_3 = \sqrt{\frac{2\epsilon_s(V_D + v_t \ln \frac{N_3}{n_i} - \psi_2)}{qN_3}} \quad (28)$$

L_2 is the channel length. Using the continuity of potential at the boundary of regions R_1/R_2 and R_2/R_3 , the depletion lengths can be calculated (16, 19).

4. ANALYTICAL MODEL FOR THE DRAIN CURRENT

The predominant current mechanism in TFETs is band-to-band tunneling (BTBT) current from the valance band of the source to the conduction band of the channel. In order to obtain the drain current, the amount of BTBT tunneling generation in Equation 30 must be integrated over the effective tunneling volume. The drain current is expressed as:

$$I_D = q \int_0^{L_{si}} \int_0^L \int_0^W G_{btbt} dx dy dz \quad (29)$$

The band-to-band generation rate (G_{BTBT}) is taken from the Kane's model (20).

$$G_{BTBT} = A \frac{|E|^P}{\sqrt{E_g}} \exp \left[-B \frac{E_g^{3/2}}{|E|} \right] \quad (30)$$

In Equation 30, A and B are Kane's model parameters and are dependent on the material, the semiconductor energy gap and the effective mass of the carriers (10, 15, 16, 18). $|E|$ represents electric field amplitude $|E| = \sqrt{E_x^2 + E_y^2}$. In a heterojunction, $E_{g,eff}$ is used as the effective bandgap, and is usually determined from the bandgaps of the bulk and electron affinities materials. $E_{g,eff} = E_{g2} - \Delta E_V$ and $\Delta E_V = (\chi_2 - \chi_1) + (E_{g2} - E_{g1})$.

ΔE_V is the off-set between the valance bands of the two materials (18). BTBT generation rate is dependent on the electric field at each point. Integrating G_{BTBT} in Equation 29 is the main challenge in TFET analytical modeling. Equation 29 cannot be computed analytically; therefore, several methods have been proposed in the literature for approximating this integral, in which accuracy is a challenge. The integral solution method in these papers (21, 22) is based on the exponential dependence of the G_{BTBT} on the electric field and considering the average electric field (E_{avg}) in the minimum tunneling length ($l_{tum,min}$). In this paper, to calculate this integral, the tangent line approximation method is used (9-11).

4. 1. Tunneling Generation Rate in the Channel

In the tangent line approximation (TLA) method, G_{BTBT} curve is considered along the channel length, and the tangent lines on the curve are drawn to intersect the horizontal axis and form triangles. The sum of these triangular areas gives the beneath area of the G_{BTBT} curve, which is the integral of the tunneling value.

At the source-channel interface, G_{BTBT} has the highest value, and along the channel it reduces steeply to a negligible amount. In Figure 5, l_1 , l_2 and l_3 are tangent lines and the area under these lines can be expressed as G_1 , G_2 and G_3 , respectively. To simplify the method, at the source-channel junction y is assumed $y = 0$.

$$L_1 = \frac{G_{btbt}(0)}{G'_{btbt}(0)} \quad (31)$$

$$G_1 = \frac{1}{2} L_1 G_{btbt}(0) = \frac{1}{2} L_1^2 G'_{btbt}(0) \quad (32)$$

The slope of the tangent line l_1 at $y = 0$ is denoted as $G'_{btbt}(0)$. L_1 is the length between $y = 0$ and the y-intercept of l_1 . Then, at $y = L_1$, the tangent line l_2 is determined with the slope of $G_{btbt}(y)$ to the point where it intersects the previous line l_1 . The length between y-intercepts of l_1 and l_2 is called L_{2d} .

$$L_2 = \frac{G_{btbt}(L_1)}{G'_{btbt}(L_1)} \quad (33)$$

$$G_2 = \frac{1}{2} L_{2d}^2 G'_{btbt}(L_1) \quad (34)$$

$$L_{1d} = \frac{G'_{btbt}(0) L_2}{G'_{btbt}(0) - G'_{btbt}(L_1)} \quad (35)$$

The L_{1d} expresses the distance between the intersection of line l_2 on the y-axis and the intercept of l_1 and l_2 lines. This process is repeated and tangent line l_3 is drawn from its y-intercept ($L_1 + L_2$) to the point it intersects the tangent line l_2 .

$$L_3 = \frac{G_{btbt}(L_1+L_2)}{G'_{btbt}(L_1+L_2)} \quad (36)$$

$$G_3 = \frac{1}{2} L_{2d}^2 G'_{btbt}(L_1 + L_2) \quad (37)$$

$$L_{2d} = \frac{G'_{btbt}(L_1) L_3}{G'_{btbt}(L_1) - G'_{btbt}(L_1+L_2)} \quad (38)$$

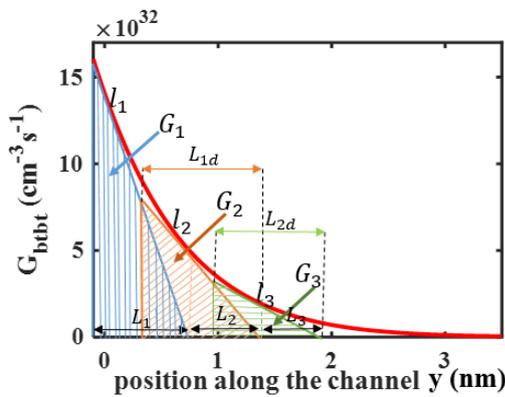


Figure 5. The band to band generation rate as a function of y along the channel length at $V_{GS} = 1.5V$ and $V_{DS} = 1V$. The tangent lines l_1, l_2, l_3 on the G_{btbt} curve and creating the triangular areas with the area G_1, G_2, G_3 , respectively; which can be regarded as the total area of the shaded regions

The G_{1d} represents the common area under the tangent lines l_1 and l_2 . Furthermore, G_{2d} is the common areas between tangent lines l_2 and l_3 as shown in Figure 6. The general expression are given in literature (9).

$$G_{1d} = \frac{1}{2} (L_{1d} - L_2)^2 G'_{btbt}(0) \quad (39)$$

$$G_{2d} = \frac{1}{2} (L_{2d} - L_3)^2 G'_{btbt}(L_1) \quad (40)$$

Finally, a closed-form equation for total G_{BTBT} is obtained using TLA method that gives $\int G_{BTBT} dv = G_{total}$. Final analytical expression is given in Equation 41.

$$G_{btbt \text{ channel}} = G_1 + G_2 + G_3 + \dots + G_n - G_{1d} - G_{2d} - \dots - G_{n-1d} \quad (41)$$

4. 2. Tunneling Generation Rate in the Source Depletion

To calculate the drain current accurately, the tunneling rate in the source depletion region should be calculated in addition to tunneling rate in the channel region. Figure 7 illustrates the tunneling rate in the source depletion region calculated by using the TLA method, which increases the accuracy of the model at high gate voltages.

$$G_{btbt \text{ source depletion}} = G_{1,sdep} + G_{2,sdep} + \dots + G_{n,sdep} - G_{1d,sdep} - G_{2d,sdep} - \dots - G_{n-1d,sdep} \quad (42)$$

Considering both $G_{btbt \text{ channel}}$ and $G_{btbt \text{ source depletion}}$, the total G_{BTBT} can be expressed as follows:

$$G_{BTBT \text{ total}} = G_{btbt \text{ channel}} + G_{btbt \text{ source depletion}} \quad (43)$$

Thus, the drain current per unit width is calculated as follows:

$$I_D = q t_{si} G_{BTBT \text{ total}} f_{fermi} \quad (44)$$

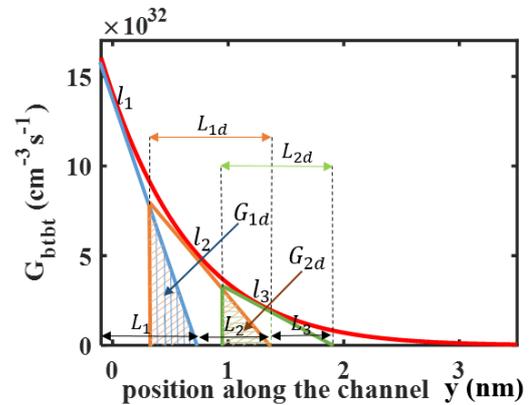


Figure 6. The tangent line approximation method for solving the integral of the tunneling value. The common area of the tangent lines l_1 and l_2 on the G_{btbt} curve and forming a triangular area with the area G_{1d} . Also G_{2d} is the common area under the tangent lines l_2 and l_3 on the G_{btbt} curve

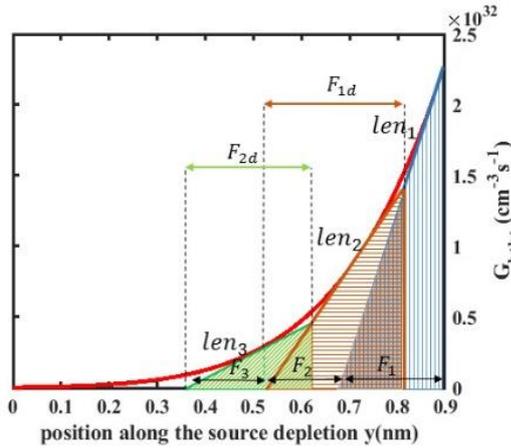


Figure 7. The tunneling generation rate as a function of y along the source depletion region at $V_{GS} = 1.5V$ and $V_{DS} = 1V$. The tangent lines len_1, len_2, len_3 on the G_{btbt} curve and creating the triangular areas with the area $G_{1,sep}, G_{2,sep}, G_{3,sep}$ respectively. Common areas of the tangent lines len_1 and len_2 , len_2 and len_3 on the G_{btbt} curve in the source depletion region forming a triangular area with the area $G_{1d,sep}$ and $G_{2d,sep}$ respectively

where f_{fermi} is a correction factor to insure $I_{DS} = 0$ at $V_{DS} = 0$ and η is an experimental coefficient (9, 19).

$$f_{fermi} = 1 - \frac{2}{(1 + e^{\frac{V_{DS}}{\eta V_t}})} \quad (45)$$

5. MODEL VERIFICATION AND RESULTS

To verify the proposed model, in Figure 1, the investigated H-TFET is simulated with the ATLAS device simulator and is compared with the proposed analytical model. The device parameters used for simulation and analytical model are mentioned in Tables 1 and 3.

The BTBT generation rate in the source channel junction at the beginning of the channel is more than the amount of tunneling in the source depletion region, and in the calculations, the amount of tunneling in the channel region is dominant. With considering the amount of

tunneling in the source depletion region along with the channel region, the accuracy of drain current increases as are mentioned in Table 2. The accuracy can be improved by increasing the number of tangent lines. Accuracy of TLA method is reported in Table 2 for number of repetition steps. The accuracy 97.2% can be achieved by using eight tangent lines.

In this section, comparisons are made for a number of different biases, dielectric constants for oxide devices and different materials and mole fractions of material. For TFET and heterojunction TFET based on different material devices, parameters like bandgap, charge carriers tunneling masses, constants of Kane's model are adjusted.

In Table 3, parameters like charge carriers tunneling masses, intrinsic carrier concentration, permittivity and constants of Kane's model for TFET and H-TFETs based on different material devices are adjusted.

Here, a number of results from the proposed analytical model for the n-type DG-TFET structure as well as the double gate H-TFET structure are reviewed and evaluated. For this purpose, a different set of the results of proposed model are compared with simulation results with the calibrated numerical simulator.

In Figure 8, the transfer characteristic ($I_D - V_{GS}$) for a double-gate TFET with the specifications mentioned in Table 1 is drawn for the drain voltage $V_{DS} = 1V$ and $V_{DS} = 0.1V$. In this figure, the transition characteristic is drawn for both logarithmically and linearly. As can be seen from Figure 8, in a wide range of biases, the obtained results from our analytical model are consistent with the TCAD simulation results.

TABLE 2. Accuracy of TLA method

Number of repetition steps	Accuracy considering channel	Accuracy considering channel & source depletion
1	53.95%	60.95%
2	83.04%	85.48%
3	92.52%	93.41%
4	95.66%	96.02%

TABLE 3. BTBT model parameters

Material	Si	SiGe	GaAs _{0.5} Sb _{0.5}	In _{0.53} Ga _{0.47} As
<i>me.tunnel</i> (Kg)	0.322 m_0	0.328 m_0	0.053 m_0	0.0332 m_0
<i>mh.tunnel</i> (Kg)	0.549 m_0	0.549 m_0	0.421 m_0	0.471 m_0
Intrinsic carrier concentration (<i>ni</i>)	1.45×10^{10}	2.86×10^{12}	1.37×10^{12}	7.62×10^{11}
Permittivity (ϵ_r)	3.9 ϵ_0	13.9 ϵ_0	14.3 ϵ_0	13.9 ϵ_0
Material TFET	A_{BTBT} (eV ^{1/2} /cm.s.V ²)		B_{BTBT} (V/cm.eV ^{3/2})	
Si	3.5×10^{21}		22.5×10^6	
SiGe/Si	4.1×10^{22}		2.45×10^7	
GaAsSb/InGaAs	1.3×10^{20}		6.3×10^6	

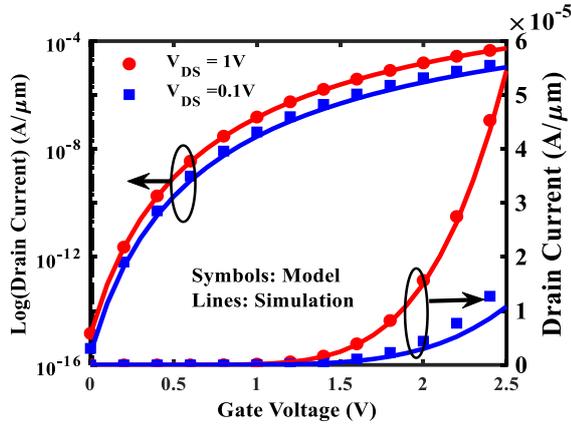


Figure 8. Comparison of the results obtained from (Lines) simulations and (symbols) our model for the transfer characteristic DG-TFET based on Si for different drain voltages as Logarithmically and linearly

In the TFET based on the hetero dielectric and the hetero structure, different materials from groups III-V with group IV are used to increase the performance of H-TFET (23-25). In the analytical models, the effect of using materials with hetero structure is considered in addition to the tunneling phenomenon to the appropriate placement of the valance and conduction bands in the TFET, and an effective bandgap $E_{g,eff}$ is defined for the model (10, 18).

In the heterojunction (n-type) TFET, the use of low bandgap materials at the source side along with a lower tunneling mass causes more overlap between the energy bands at the channel-source junction and smaller the tunneling distance. As a result, the on-state current increases (26).

The comparison between switching performance of homojunction TFET and double gate H-TFET using $Si_{0.2}Ge_{0.8}$ in the source and Si in the channel and drain regions is shown in Figure 9.

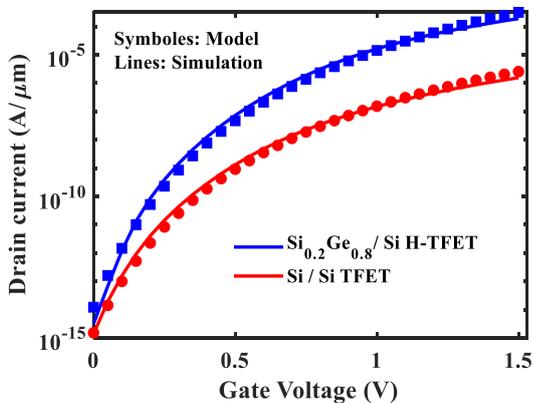


Figure 9. Transfer characteristics of homojunction TFET and H-TFETs predicted by the TCAD simulators (line) and analytical models (symbol)

As expected, for the same gate voltage range, H-TFET structure provide more I_{ON} current than homojunction TFET structure, and our proposed analytical model can follow the simulation curve well.

In Figure 10, the transfer characteristics of three double gate H-TFETs based on different molar ratio of $Si_{(1-x)}Ge_x$ are shown. As the results of simulation and analytical model are shown, $Si_{0.2}Ge_{0.8}$ has more current than the other two structures. The reason is the smaller energy gap in the source region and the more energy band bending at the tunneling region of $Si_{0.2}Ge_{0.8}$ compared to $Si_{0.5}Ge_{0.5}$ and $Si_{0.7}Ge_{0.3}$ structures.

Mostefai (27) studied on energy gap (E_g) as a function of different composition x and y in $Ga_xIn_{1-x}As_ySb_{1-y}$ at $T = 300 K$ and intrinsic carrier concentration (ni) as a function of temperature and effective density of states N_c and N_v in the conduction band and valence band, respectively. So far several articles worked on $GaAs_{1-y}Sb_y$ and $In_{1-x}Ga_xAs$ (10, 28, 29). The effective energy bandgap ($E_{g,eff}$) for heterojunction TFET is an important design parameter. Although there is no definite value for $E_{g,eff}$ in the $GaAs_{0.5}Sb_{0.5}/In_{0.53}Ga_{0.47}As$. Various ranges of effective bandgap can be seen in the literature (30, 31). With comparing measured $I_D - V_{GS}$ curve for the $GaAs_{0.5}Sb_{0.5}/In_{0.53}Ga_{0.47}As$ H-TFET, $E_{g,eff} = 0.42eV$ is more suitable (31).

For best matching between the results of the TCAD simulator and the results of the analytical model. Figure 11 shows transfer characteristics $I_D - V_{GS}$ for different $E_{g,eff}$ in $V_{Ds} = 1V$ and dielectric constant Al_2O_3 ($\epsilon_{ox} = 8.5$). Parameters mentioned in Table 2 are accounted for $GaAsSb/InGaAs$ and its results are compared with simulation results.

In Figure 12, the effect of drain-source voltages on transfer characteristic of a DG-HTFET $GaAs_{0.5}Sb_{0.5}/$

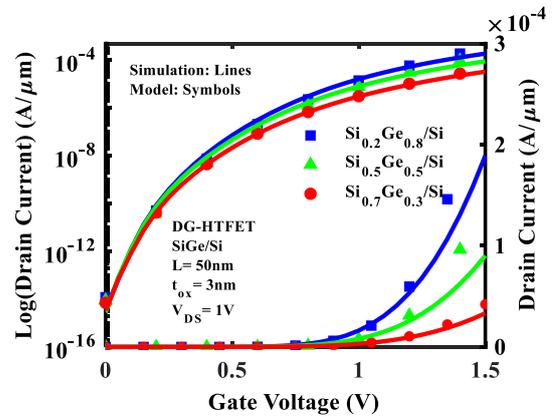


Figure 10. $I_D - V_{GS}$ curve obtained for $Si_{(1-x)}Ge_x/Si/Si$ H-TFET for molar ratios $X= 0.3, 0.5, 0.8$, where the simulations are shown as lines and the analytical model is shown as symbols

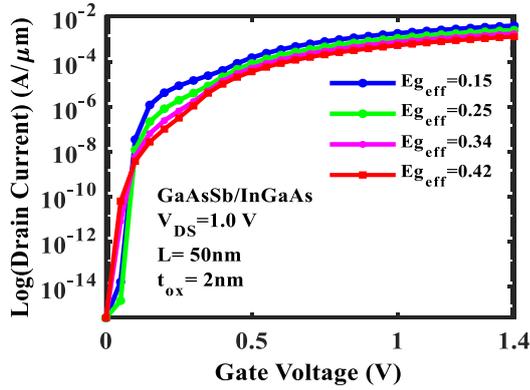


Figure 11. The transfer characteristics of $GaAs_{0.5}Sb_{0.5}/In_{0.53}Ga_{0.47}As$ for several effective energy bandgap ($E_{g,eff}$). our model predicted by TCAD simulations (lines) and by the analytical model (symbols)

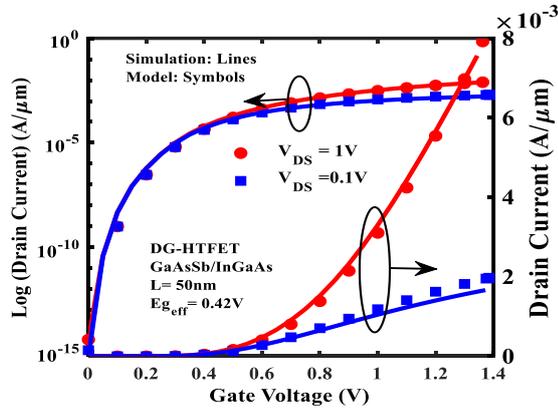


Figure 12. Comparison of the results obtained from simulations (Lines) and our model (symbols) for the transfer characteristic DG-HTFET based on $GaAs_{0.5}Sb_{0.5}/In_{0.53}Ga_{0.47}As$ for various drain voltage ($V_{DS} = 1V$) and ($V_{DS} = 0.1V$) as Logarithmically and linearly

$In_{0.53}Ga_{0.47}As$ for $V_{DS} = 1V$ and $V_{DS} = 0.1V$ is shown. In short channel devices, the applied drain-source voltage affects the distance of tunneling path. As can be seen in Figure 12, it modulates the I_{ON} reasonably while the subthreshold swing remains almost unchanged.

The results obtained from our analytical model are consistent with the TCAD simulator and shows that the model predicts accurately.

To reduce power density and better switching performance, TFETs are considered in low power applications (32-34). Figure 13, shows the $\log I_D - V_{GS}$ curves of a $GaAs_{0.5}Sb_{0.5}/In_{0.53}Ga_{0.47}As$ DG-HTFET for different dielectric constants.

The channel length and other device parameters are the same. Thin oxides with high permittivity lead to a greater oxide capacitance, and provides higher charge being formed in the channel at the same gate voltage, which results in a higher control of the gate on the electrostatics of the channel. This makes higher I_{ON} and lower bias voltages. As expected from Figure 13, HfO_2 provides more current at the same range of gate voltages and our analytical model are consistent with the TCAD results.

Table 4, shows comparative analysis of the proposed H-TFET structural parameters with existing literature. Transfer characteristics of DG-HTFET for two different heterojunctions, $SiGe/Si$ and $GaAsSb/InGaAs$, are shown in Figure 14, by considering TCAD simulations (Lines) and analytical model (symbols). The parameters

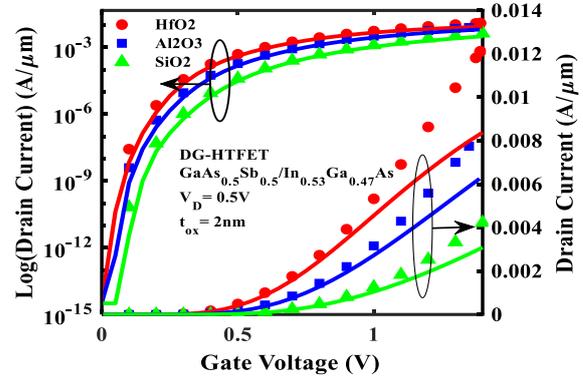


Figure 13. $I_D - V_{GS}$ curves obtained of the DG-HTFET $GaAs_{0.5}Sb_{0.5}/In_{0.53}Ga_{0.47}As$ from simulation (lines) and our model (symbols) for different dielectric constant material, HfO_2 ($\epsilon_{OX} = 22$), Al_2O_3 ($\epsilon_{OX} = 8.5$) and SiO_2 ($\epsilon_{OX} = 3.9$)

TABLE 4. The comparative H-TFET with existing literature

	Source	Channel	Eot (nm)	Oxide	I_{ON} (A/ μm)	I_{OFF} (A/ μm)	SS (mV/dec)
(9)	GaAsSb	InGaAs	2	SiO_2	10^{-6}	10^{-13}	45
(30)	SiGe	Si	2	HfO_2	10^{-3}	10^{-13}	60
(30)	InGaAs	InP	2	SiO_2	10^{-8}	10^{-17}	80
(21)	InGaAs	InP	2	SiO_2	10^{-5}	10^{-13}	80
(26)	Si	Si	2	SiO_2	10^{-4}	10^{-15}	55
This paper	SiGe	Si	3	SiO_2	10^{-4}	10^{-15}	50
This paper	GaAsSb	InGaAs	2	SiO_2	10^{-2}	10^{-16}	30

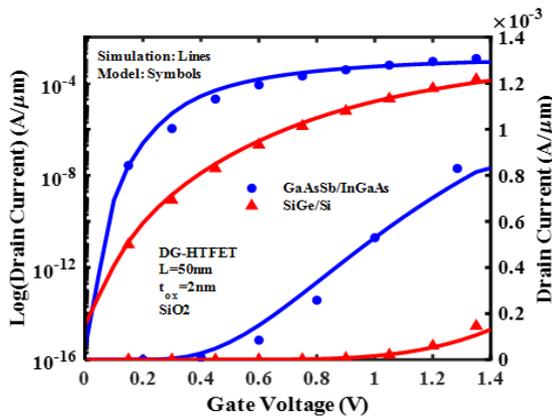


Figure 14. Transfer characteristics of the H-TFET $\text{Si}_{0.2}\text{Ge}_{0.8}/\text{Si}$ and $\text{GaAs}_{0.5}\text{Sb}_{0.5}/\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ predicted by the TCAD simulators (line) and analytical models (symbol)

in Table 2 are considered for both numerical simulation and analytical model. The $\text{GaAs}_{0.5}\text{Sb}_{0.5}/\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ has higher current than the $\text{Si}_{0.2}\text{Ge}_{0.8}$. According to Figure 14, matching the results of proposed model and the numerical simulations confirm the validity of the model.

6. CONCLUSIONS

In this paper, a universal physics-based analytical model for DG H-TFETs is proposed which accurately predicts the potential and the drain current. The fringing fields and effects of mobile charges on the reverse decay length in the channel are considered. The effect of high drain voltages on the potential, using new gate potential $\Psi_{G,eff}$, is introduced, as well. The model consists of a complete 2D expression for the potential profile to predict the potential distribution in the channel depth in addition to the surface of the device. The mid-potential $\Psi_{di,new}$ and new reverse decay length $k_{i,new}$ are modified to represent the potential distribution at any depth of the device. For drain current modeling, the effect of the extended depletion region in the source side is considered in TLA method, and G_{BTBT} is calculated in the channel and the source depletion region, which improves the accuracy of the model at high gate voltages by 97.2%. The proposed model is compared with TCAD results for different material and device parameters, which confirms the accuracy of the proposed analytical model.

7. REFERENCES

1. Maity N, Maity R, Baishya S. An analytical model for the surface potential and threshold voltage of a double-gate heterojunction

- tunnel FinFET. Journal of computational electronics. 2019;18:65-75. 10.1007/s10825-018-1279-5
2. Pan A, Chui CO. A quasi-analytical model for double-gate tunneling field-effect transistors. IEEE electron device letters. 2012;33(10):1468-70. 10.1109/LED.2012.2208933
3. Abdulwahid AH. Artificial intelligence-based control techniques for hvdc systems. Emerging Science Journal. 2023;7(2):643-53. 10.28991/ESJ-2023-07-02-024
4. Divayana DGH, Suyasa PWA, Santiyadnya N, Andayani MSL, Sundayana IM, Astawa IND, et al. Utilization of the Weighted Product-Based CIPP Evaluation Model in Determining the Best Online Platform. HighTech and Innovation Journal. 2023;4(1):233-48. 10.28991/HIJ-2023-04-01-0155
5. Zhang L, Chan M. SPICE modeling of double-gate tunnel-FETs including channel transports. IEEE Transactions on Electron Devices. 2014;61(2):300-7. 10.1109/TED.2013.2
6. Graef M, Holtij T, Hain F, Kloes A, Iniguez B. A 2D closed form model for the electrostatics in hetero-junction double-gate tunnel-FETs for calculation of band-to-band tunneling current. Microelectronics Journal. 2014;45(9):1144-53. 10.1016/j.mejo.2014.04.033
7. Horst F, Graef M, Hosenfeld F, Farokhnejad A, Hain F, Luong GV, et al., editors. Implementation of a DC compact model for double-gate Tunnel-FET based on 2D calculations and application in circuit simulation. 2016 46th European Solid-State Device Research Conference (ESSDERC); 2016: IEEE. 10.1109/ESSDERC.2016.7599684
8. Dutta R, Sarkar SK. Analytical modeling and simulation-based optimization of broken gate TFET structure for low power applications. IEEE transactions on electron devices. 2019;66(8):3513-20. 10.1109/TED.2019.2925109
9. Devi WV, Bhowmick B, Pukhrambam PD. N+ pocket-doped vertical TFET for enhanced sensitivity in biosensing applications: modeling and simulation. IEEE Transactions on Electron Devices. 2020;67(5):2133-9. 10.1109/TED.2020.2981303
10. Vishnoi R, Kumar MJ. An accurate compact analytical model for the drain current of a TFET from subthreshold to strong inversion. IEEE Transactions on Electron Devices. 2015;62(2):478-84. 10.1109/TED.2014.2381560
11. Guan Y, Li Z, Zhang W, Zhang Y, Liang F. An analytical model of gate-all-around heterojunction tunneling FET. IEEE Transactions on Electron Devices. 2018;65(2):776-82. 10.1109/TED.2017.2783911
12. Kumar S, Singh K, Baral K, Singh PK, Jit S. 2-D analytical model for electrical characteristics of dual metal heterogeneous gate dielectric double-gate TFETs with localized interface charges. Silicon. 2021;13:2519-27. 10.1007/s12633-020-00564-5
13. Gholizadeh M, Hosseini SE. A 2-D analytical model for double-gate tunnel FETs. IEEE Transactions on Electron Devices. 2014;61(5):1494-500. 10.1109/TED.2014.2313037
14. Ponnian J, Pari S, Ramadass U, Ooi CP. A Unified Power-Delay Model for GDI Library Cell Created Using New Mux Based Signal Connectivity Algorithm. Emerging Science Journal. 2023;7(4):1364-94. 10.28991/ESJ-2023-07-04022
15. Bardon MG, Neves HP, Puers R, Van Hoof C. Pseudo-two-dimensional model for double-gate tunnel FETs considering the junctions depletion regions. IEEE Transactions on Electron Devices. 2010;57(4):827-34. 10.1109/TED.2010.2040661
16. Jain P, Yadav C, Agarwal A, Chauhan YS. Surface potential based modeling of charge, current, and capacitances in DGTFTFET including mobile channel charge and ambipolar behaviour. Solid-State Electronics. 2017;134:74-81. 10.1016/j.sse.2017.05.012
17. Keighobadi D, Mohammadi S. Physical and analytical modeling of drain current of double-gate heterostructure tunnel FETs.

- Semiconductor Science and Technology. 2018;34(1):015009. 10.1088/1361-6641/aaeeeb
18. Kaur S, Raman A, Sarin RK. Analytical modeling of surface potential, capacitance and drain current of heterojunction TFET. Applied Physics A. 2020;126:1-11. 10.1007/s00339-020-03945-0
 19. Lin S, Kuo J. Modeling the fringing electric field effect on the threshold voltage of FD SOI nMOS devices with the LDD/sidewall oxide spacer structure. IEEE Transactions on Electron Devices. 2003;50(12):2559-64. 10.1109/TED.2003.816910
 20. Kane E. Zener tunneling in semiconductors. Journal of Physics and Chemistry of Solids. 1960;12(2):181-8. 10.1016/0022-3697(60)90035-4
 21. Mohammadi S, Keighobadi D. A universal analytical potential model for double-gate Heterostructure tunnel FETs. IEEE Transactions on Electron Devices. 2019;66(3):1605-12. 10.1109/TED.2019.2895277
 22. Ranjith R, Suja K, Komaragiri RS. An analytical model for a TFET with an n-doped channel operating in accumulation and inversion modes. Journal of Computational Electronics. 2021;20:1125-36. 10.1007/s10825-021-01683-x
 23. Howldar S, Balaji B, Srinivasa Rao K. Design and Analysis of Hetero Dielectric Dual Material Gate Underlap Spacer Tunnel Field Effect Transistor. International Journal of Engineering, Transactions C: Aspects. 2023;36(12):2137-44. 10.5829/ije.2023.36.12c.01
 24. Howldar S, Balaji B, Srinivasa Rao K. Design and qualitative analysis of hetero dielectric tunnel field effect transistor device. International Journal of Engineering, Transactions C: Aspects. 2023;36(6):1129-35. 10.5829/ije.2023.36.06c.11
 25. Keighobadi D, Mohammadi S, Fathipour M. An analytical drain current model for the cylindrical channel gate-all-around heterojunction tunnel FETs. IEEE Transactions on Electron Devices. 2019;66(8):3646-51. 10.1109/TED.2019.2922232
 26. Dong Y, Zhang L, Li X, Lin X, Chan M. A compact model for double-gate heterojunction tunnel FETs. IEEE Transactions on Electron Devices. 2016;63(11):4506-13. 10.1109/TED.2016.2604001
 27. Mostefai A. Basic Characteristics of Gallium Indium Arsenide Antimonide ($GaxIn_{1-x}As_ySb_{1-y}$) Semiconductors Using MATLAB. Journal of Nano-and Electronic Physics. 2022;14(4). 10.21272/jnep.14(4).04027
 28. Convertino C, Zota CB, Schmid H, Caimi D, Czomomaz L, Ionescu AM, et al. A hybrid III-V tunnel FET and MOSFET technology platform integrated on silicon. nature electronics. 2021;4(2):162-70. 10.1038/s41928-020-00531-3
 29. Arabhavi AM, Ciabattini F, Hamzeloui S, Flückiger R, Saranovac T, Han D, et al. InP/GaAsSb double heterojunction bipolar transistor emitter-fin technology with $f_{MAX} = 1.2$ THz. IEEE Transactions on Electron Devices. 2022;69(4):2122-9. 10.1109/TED.2021.3138379
 30. Smets Q, Verhulst AS, El Kazzi S, Gundlach D, Richter CA, Mocuta A, et al. Calibration of the effective tunneling bandgap in GaAsSb/InGaAs for improved TFET performance prediction. IEEE transactions on electron devices. 2016;63(11):4248-54. 10.1109/TED.2016.2604860
 31. Detz H, Silvano de Sousa J, Leonhardt H, Klang P, Zederbauer T, Andrews AM, et al. InGaAs/GaAsSb based two-dimensional electron gases. Journal of Vacuum Science & Technology B. 2014;32(2). 10.1116/1.4863299
 32. Zare M, Peyravi F, Hosseini SE. Impact of hetero-dielectric ferroelectric gate stack on analog/RF performance of tunnel FET. Journal of Electronic Materials. 2020;49:5638-46. 10.1007/s11664-020-08315-3
 33. Ghosh S, Chattopadhyay A, Tewari S. Optimization of hetero-gate-dielectric tunnel FET for label-free detection and identification of biomolecules. IEEE transactions on Electron Devices. 2020;67(5):2157-64. 10.1109/TED.2020.2978499
 34. Arunkumar N, Senathipathi N, Dhanasekar S, Malin Bruntha P, Priya C. An ultra-low-power static random-access memory cell using tunneling field effect transistor. International Journal of Engineering, Transactions B: Applications. 2020;33(11):2215-21. 10.5829/ije.2020.33.11b.13

COPYRIGHTS

©2024 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 مدل تحلیلی دقیق برای جریان درین ترانزیستورهای تونلی ناهمگون، با در نظر گرفتن ناحیه تخلیه سورس، بارهای متحرک، و اثر ولتاژ درین ارائه شده است. این مدل با استفاده از روابط ریاضی جدید فرموله شده، توزیع پتانسیل را نه تنها در سطح، بلکه در عمق نیمه هادی به طور دقیق پیش بینی می کند. با استفاده از روش تقریب خط مماس با در نظر گرفتن ناحیه کانال و همچنین ناحیه تخلیه سورس، جریان تونل زنی باند به باند از ناحیه سورس به کانال به صورت تحلیلی و با استفاده از انتگرال مقدار تولید تونل زنی مدل کین محاسبه می شود. در مقایسه با نتایج شبیه سازی، مدل پیشنهادی دقت قابل توجهی در پیش بینی جریان درین دارد.



Analysis of Nucleation Time of Gas Hydrates in Presence of Paraffin During Mechanized Oil Production

G. Y. Korobov, A. A. Vorontsov*, G. V. Buslaev, V. T. Nguyen

Petroleum Engineering Department, Empress Catherine II Saint Petersburg Mining University, Saint Petersburg, Russia

PAPER INFO

Paper history:

Received 19 December 2023

Received in revised form 21 February 2024

Accepted 07 March 2024

Keywords:

Gas Hydrates

Asphalt-resin-paraffin Deposits

Waxes

Gas Hydrate Formation Time

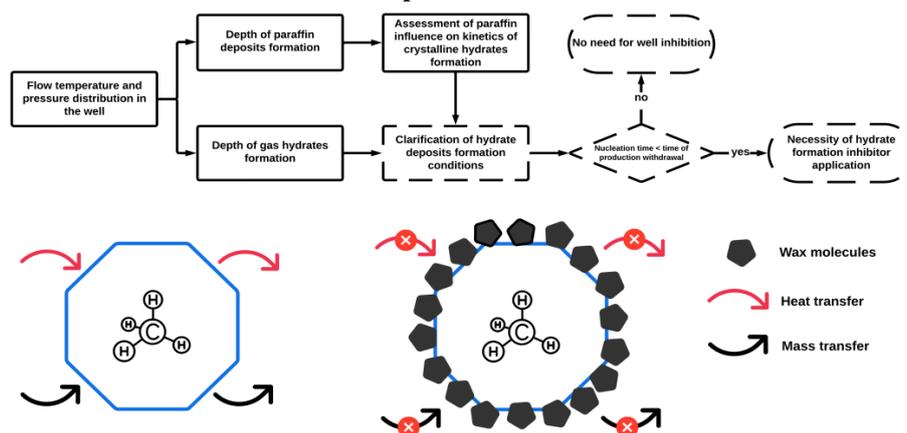
Kinetic Inhibitor of Hydrate Formation

ABSTRACT

The objective of this study is to investigate the nucleation timing of gas hydrate molecules in oil flows. This research focuses on examining how paraffin particles impact the formation timing of hydrate deposits during the mechanical production of oil. A thorough comprehension and control over the formation of organic deposits within the wellbore can substantially mitigate equipment maintenance expenses, enhance the safety and consistency of production, and bolster the economic viability of extracting hydrocarbons. The initial segment of the paper outlines a methodology for identifying the formation depths of gas hydrates and asphaltene-resin-paraffin deposits (ARPD) in operational oil wells through the resolution of thermobaric differential equation systems. Subsequent laboratory experiments were conducted to assess the nucleation timing of gas hydrates in the presence of paraffin. These tests were performed in a specialized high-pressure autoclave that enables the establishment of requisite thermobaric conditions. An internal agitator in the autoclave facilitates the needed dispersion within the system to emulate well flow conditions. Experimental findings revealed that paraffin particles impede the formation of gas hydrate deposits and decelerate their nucleation process. Notably, a 3% increase in paraffin concentration within the mixture was observed to prolong the nucleation timing of gas hydrates by a factor of nine. Based on the review of available literature, it is deduced that further comprehensive investigations are essential for the advancement of a temporal model governing the operational dynamics of production wells under the influence of gas hydrate and ARPD formation.

doi: 10.5829/ije.2024.37.07a.13

Graphical Abstract



Disturbance of mass and heat exchange between hydrates due to wax particles

*Corresponding Author Email: andrey-vorontsov98@yandex.ru (A.A. Vorontsov)

1. INTRODUCTION

Gas hydrates, also known as clathrates, are crystalline structures where water molecules form a lattice that encases gas molecules under specific temperature and pressure conditions. The thermodynamic prerequisites for gas hydrate formation vary according to gas type, temperature, pressure, gas saturation, and the salinity of water (1) as well as the compositional makeup of other phases present. Methane (CH_4), ethane (C_2H_6), propane (C_3H_8), butane (C_4H_{10}), carbon dioxide (CO_2), hydrogen sulfide (H_2S), nitrogen (N_2), and other gases in hydrocarbon production can form hydrates under suitable conditions (2). Methane-based natural gas hydrates are particularly notable for their energy potential, with one cubic meter of hydrate yielding approximately 165 standard cubic meters of CH_4 .

The formation of gas hydrates can be classified into anthropogenic and natural occurrences. Anthropogenic hydrates may develop in oil and natural gas production systems characterized by a high gas-to-oil ratio, including areas such as the bottom hole zone, wellbores, and field pipelines (3). In contrast, natural hydrates may aggregate into clusters or exist in a dispersed state (4). The genesis of anthropogenic natural gas hydrates poses considerable challenges to the operational efficiency of oil and gas production facilities by complicating technological processes, escalating energy demands, shortening equipment maintenance intervals, and amplifying risks in the exploration of new Arctic oil and gas ventures (5, 6).

The formation of asphaltene-resin-paraffin deposits (ARPD) in oilfield equipment is a more prevalent issue than gas hydrate deposition. The development of numerous fields in the Volga-Ural and West Siberian oil and gas provinces is hindered by ARPD formation (7).

The severe climatic conditions of the Far North, exacerbated by permafrost cooling of flows, coupled with the rigorous climate, amplify the challenges associated with gas hydrate and ARPD formation during oil and gas extraction (8, 9). The coalescence of gas hydrates and ARPD into a unified deposit further complicates the operation of production wells. Therefore, refining the theoretical understanding of gas hydrate and ARPD formation is crucial for the enhanced exploitation of Arctic territories within the Russian Federation (10, 11).

For the most part, it is imperative to acknowledge that the thermobaric conditions prevailing within a well exert a profound influence on the process of hydrocarbon production process. The pressure and temperature parameters of the flow within the well are critical determinants in the formation of organic deposits (12). In this context, a methodology for assessing the flow temperature inside the borehole was introduced by researchers (13). This approach is grounded in differential equations of heat conduction, incorporating

the heat transfer coefficient between the surrounding rock formations and the well's flow. The determination of this coefficient leverages field-derived data (14). The phenomenon of convective heat transfer in vertical gas-liquid two-phase hydrocarbon flows has been examined (15, 16), with laboratory experiments revealing the relationship between the heat transfer and the velocity and structure of the mixture. The theoretical endeavors have culminated in the formulation of an elaborate mechanistic model for heat transfer. This model accounts for the flow's pattern, distinguishing among bubble, slug, and annular structures, and is adept at predicting the flow structure before calculating the flow's hydrodynamics and heat transfer characteristics based on the anticipated structure. The results of comparison between the model and real data showed that presented model calculates the heat transfer coefficient with an error margin of 20%, 30%, and 25% for bubble, annular, and slug flows, respectively. Furthermore, the integration of permafrost attributes such as the latent heat of fusion (17), alongside considerations of permafrost thawing, the migration of water from proximal zones around the well to more remote areas, and the temperature gradient, has facilitated the introduction of a novel thermal model for a producing well by the authors (18). Field tests of the model showed high convergence of the calculation results with real data.

Additionally, the impact of operating an electrical submersible pump (ESP) on the heat exchange dynamics between the pumped fluid and its surroundings warrants attention. Investigations (14, 19) have delved into the thermal interactions between an active ESP unit and its immediate environment, proposing a model that accurately predicts the temperature variations of both the unit itself and the fluid in transit through and around the pump. Moreover, the processes of well killing and the injection of specialized fluids can induce cooling in the bottom hole zone and wellbore, potentially leading to the augmented formation of ARPD and gas hydrates within the well space. This phenomenon poses challenges to the subsequent reactivation and operation of the well (20, 21).

Research documented in studies (22, 23) has highlighted the considerable impact of flow dynamics on hydrate formation, specifically through mechanisms such as destruction, coalescence, and deformation of gas bubbles, as well as their interactions with mixture flow vortices and collisions among themselves. Further investigation (24) elucidated the nuances of hydrate formation across varying levels of water content within the flow, yielding the following insights:

1. High water content facilitates hydrate nucleation at the gas-water phase boundary, significantly accelerating the agglomeration process of gas hydrates.
2. At medium water content, hydrate nucleation predominantly occurs at the oil-water interface, resulting

in fewer hydrate deposits and a reduced rate of hydrate formation.

3. Systems with low water content exhibit decreased density and thickness in gas hydrate deposits.

These findings underscore the critical importance of understanding hydrate formation dynamics under various well conditions.

Hydrate deposit management commonly employs the use of hydrate formation inhibitors (25), which are categorized by most researchers based on literature (26):

- Their mechanism of action, dividing them into thermodynamic, kinetic, and anti-agglomerant types.
- Composition, distinguishing between single-component and multi-component inhibitors.
- The number and nature of functions they perform.
- Physical and chemical properties, such as density, volatility, and freezing point.

The accumulation of asphalt-resin-paraffin deposits (ARPD) on field equipment and pipeline surfaces poses significant challenges to oil production, transportation, and refining processes (27). Key factors contributing to ARPD formation in field equipment have been identified (28) as:

- The presence of heavy oil components that agglomerate into large ARPD particles.
- The release of oil-dissolved gases when flow pressure falls below the gas saturation pressure of oil.
- Reduction of flow temperature beneath the paraffin saturation point of oil.
- Situations where disruptive flow forces, dependent on velocity, regime, and structure, are outweighed by the cohesive forces within the deposits.

The most effective technological approach for ARPD management involves the use of chemical inhibitors with a crucial focus on minimizing inhibitor loss (29) due to adsorption by the surrounding rock formations (30). A comprehensive methodology for mitigating ARPD formation in the wellbore area has been proposed (31), incorporating well washing with specialized solvents to dissolve existing deposits. Additionally, other strategies for combating ARPD formation in production wells have been explored (32). Investigations (33-35) have examined how the material composition of production pipes influences ARPD formation conditions. Another study (36) introduced a heating cable system as a method for controlling paraffin deposition in production wells, while research (37) presented a novel approach involving magnetic reagent treatment of oilfield equipment.

As previously discussed, the concurrent formation of gas hydrates and asphalt-resin-paraffin deposits (ARPD) significantly complicates hydrocarbon productions via mechanized methods (38, 39). Research presented by Wang et al. (40) introduced an advanced model for hydrate nucleation in the presence of ARPD and surfactants, with asphaltenes potentially acting as such

surfactants, as illustrated in Figure 1. The study yielded several key findings:

1. The nucleation process of hydrate molecules is decelerated by the presence of ARPD or surfactants.
2. ARPD molecules obstruct the mass and heat transfer process between hydrate molecules.
3. The volume of hydrate deposits escalates in the absence of surfactants but in the presence of ARPD. This increase is attributed to the integration of ARPD molecules into the hydrate agglomerates, further complicating their disintegration.

Therefore, the primary aim of this research is to delve into the dynamics of gas hydrate deposit formation within a producing well, particularly in scenarios where oil flow contains wax particles. Laboratory observations could facilitate the development of a temporal model for a producing well. This model enables the prediction of the time required for product removal and the formation of organic deposits. Under certain conditions, this model suggests that gas hydrates may not accumulate in bottom hole equipment to a significant extent. Consequently, this allows for a strategic adjustment in the delivery of hydrate formation inhibitors to surface equipment rather than directly to the bottom hole. Such a shift in the inhibitor delivery strategy is anticipated to streamline equipment maintenance and substantially reduce the financial burden associated with the injection of reagents into the wellbore.

2. METHODOLOGY

2. 1. Methodology for Determining Depth of Gas Hydrate and ARPD Formation in a Well

The methodology for identifying the initial depths of organic deposits formation is structured around a comprehensive algorithm that encompasses several critical steps:

1. Initial Setup: Define the initial well configuration and production conditions based on field data to establish a baseline for further analysis.
2. ESP Operating Parameters: Determine the operational parameters of the Electrical Submersible Pump (ESP) unit, incorporating adjustments for emulsions and free gas content to ensure accurate modeling of the pump's performance under field conditions.

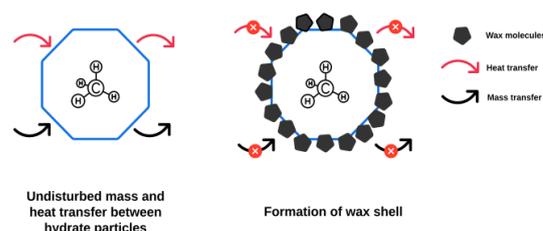


Figure 1. Disturbance of mass and heat exchange between hydrates due to wax particles

3. Temperature Distribution Evaluation: Assess the temperature distribution along the well's depth, taking into account the heat generated by the ESP unit (38). This step is crucial for understanding the thermal dynamics within the wellbore.

4. Pressure Distribution Calculation: Utilize the updated Poetman-Carpenter method to calculate the pressure distribution at various depths within the well for both the production string and tubing. This method provides a refined approach to estimating pressure profiles.

5. Linking Pressure to Temperature: Correlate the depth-specific pressure values with corresponding flow temperature readings. This correlation specifies the P(T) distribution within the well, offering a detailed view of the thermobaric conditions.

6. Well Operation Modeling: Estimate the true operational parameters of the well under the specified conditions using a well operation model. This step involves specifying the flow rate, bottom hole pressure, and receiving pressure, thereby refining the operational parameters.

7. Hydrate Formation Equilibrium: Compare the P(T) distribution against the equilibrium conditions for hydrate formation to identify the onset depth of hydrate formation (22). This comparison is vital for pinpointing where hydrates begin to form within the well.

8. Thermogram and Wax Crystallization: Match the well's thermogram against the temperature profile for wax crystallization across the well's depth to determine the formation depth of paraffin.

A block diagram illustrating this algorithm is depicted in Figure 2; providing a visual representation of the process flow.

Upon implementing this algorithm, it is possible to achieve accurate determination of the initial depths for ARPD and crystalline hydrate formation. To further enhance this algorithm, incorporating a time variable into the calculations can account for the kinetics of gas hydrate formation in the presence of ARPD, making the model more dynamic and reflective of real-world conditions (41, 42).

For the initial determination of organic deposits formation depths, the pressure characteristic of the ESP

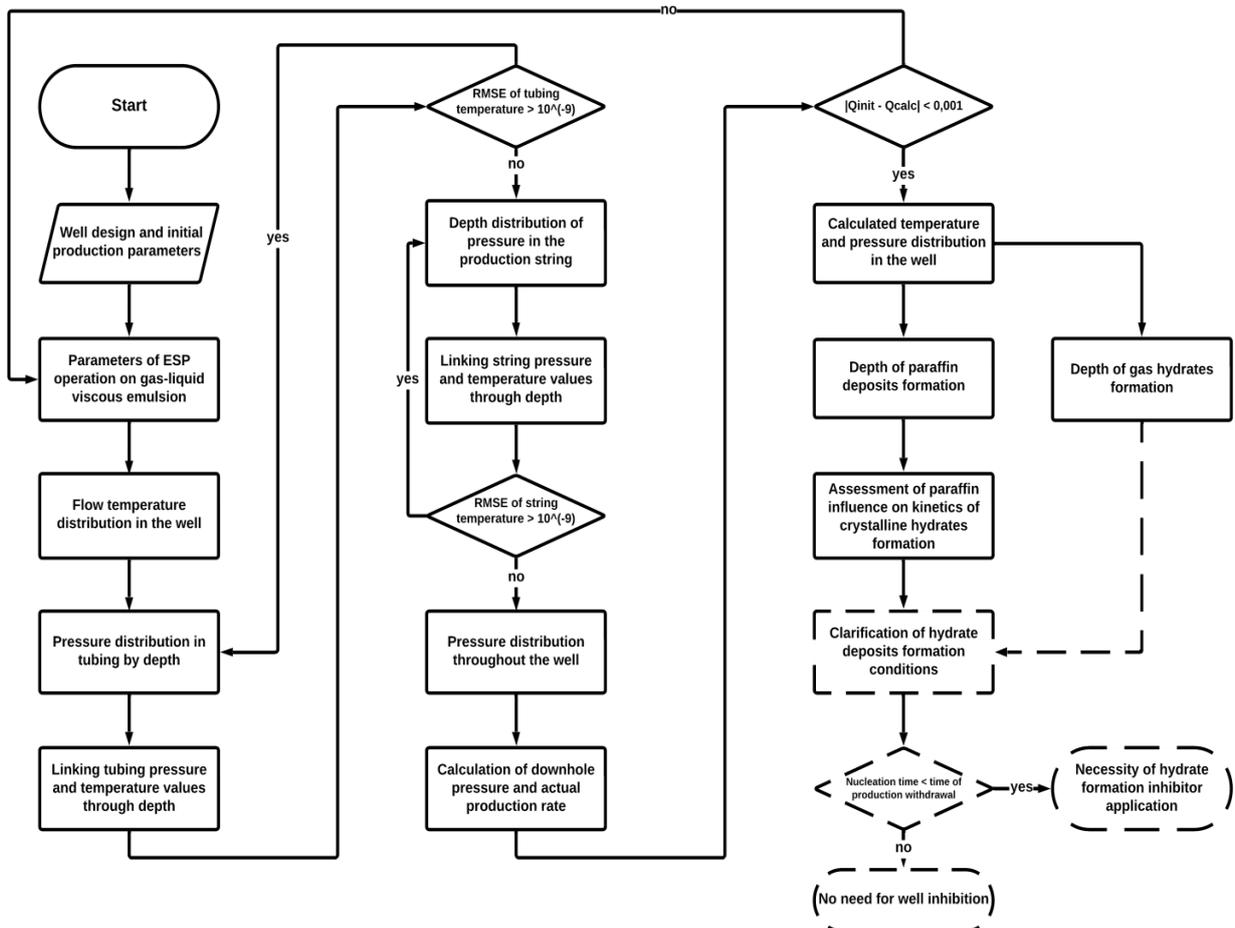


Figure 2. Algorithm to determine the depth of organic sediment formation

in emulsion with gas presence is established. The pump's pressure output is derived from its passport characteristics, which are then recalibrated for a viscous liquid using Lyapkov method (43). This recalibration involves determining the pump stage's coefficient of rapidity, denoted as n_s , as outlined in Equation 1:

$$n_s = 193 \cdot n \cdot \left(\frac{Q_{liq,optim.}}{86400}\right)^{0,5} \cdot \left(g \cdot \frac{H_{liq,optim.}}{z}\right)^{-0,75} \quad (1)$$

where n is the number of pump revolutions, rpm; $Q_{liq,optim.}$, $H_{liq,optim.}$ - optimum flow rate (m³/day) and pump head on water (m); z - number of stages.

Reynolds number of the flow in the channels of ESP, Equation 2:

$$Re_H = \frac{4,3+0,816 \cdot n_s^{0,274}}{n_s^{0,575}} \cdot \frac{Q_{liq}}{\left(\frac{\mu_{mix}}{\rho_{mix}}\right) \cdot 86400} \cdot \sqrt[3]{\frac{n \cdot 86400}{Q_{liq,optim.}}} \quad (2)$$

where μ_{mix} , ρ_{mix} are mixture viscosity (Pa · s) and density (kg/m³); Q_{liq} - well flow rate, m³/day.

Conversion factors for viscous fluid, Equations 3 and 4:

$$K_{HQ} = 1 - (3,585 - 0,821 \cdot \lg(Re_H)) \cdot \left(0,027 + 0,0485 \frac{Q_{liq}}{Q_{liq,optim.}}\right) \quad (3)$$

$$K_\eta = \begin{cases} 0,485 \cdot \lg(Re_H) - 0,63 - 0,26 \frac{Q_{liq}}{Q_{liq,optim.}} & \text{if } Re_H < 2320 \\ 0,274 \cdot \lg(Re_H) - 0,06 - 0,14 \frac{Q_{liq}}{Q_{liq,optim.}} & \text{if } Re_H > 2320 \end{cases} \quad (4)$$

Coefficients for recalculation of the water-air mixture are now calculated using the corresponding nomogram in Figure 3 (44).

The gas content is found by Equation 5:

$$\Gamma = G / \left(\frac{\rho_{oil,degas}}{\rho_{gas}} + G\right) \quad (5)$$

where G - gas factor, m³ /m³; $\rho_{oil,degas}$ - density of degassed oil at standard conditions, kg/m³; ρ_{gas} - density of gas at standard conditions, kg/m³.

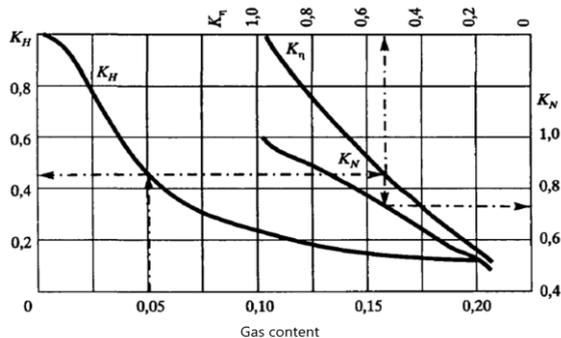


Figure 3. Nomogram for calculating pump performance parameters for water-air mixtures (44)

To accurately calculate the pressure produced by the pump P_{pump} for any given well flow rate, the approach involves obtaining conversion factors for gas. These factors are then multiplied by the coefficients determined for viscous fluids, enabling the characterization of the pump's performance with the actual emulsion present in the well. This step is crucial for understanding the pump's ability to maintain optimal pressure levels under various operational conditions, thereby ensuring the efficient transport of the fluid mixture to the surface.

The flow temperature in any section between the reservoir roof and the pump (temperature in the production string) is determined with Equation 6, the flow temperature in any section of the tubing is determined with Equation 7:

$$T_{str} = T_{res} - (H_{res} - H) \cdot \cos\alpha \cdot \frac{0,0034+0,79 \cdot Grad \cdot \cos\alpha}{10 Q_{liq}/86400 \cdot 20 \cdot d_{str,inner}^{2,67}} \quad (6)$$

$$T_{tub} = T_{res} - (H_{res} - H_p) \cdot \cos\alpha \cdot \frac{0,0034 + 0,79 \cdot Grad \cdot \cos\alpha}{10 Q_{liq}/86400 \cdot 20 \cdot d_{tub,inner}^{2,67}} - (H_p - H) \cdot \cos\alpha \cdot \frac{0,0034 + 0,79 \cdot \Gamma \cdot \cos\alpha}{10 Q_{liq}/86400 \cdot 20 \cdot d_{tub,inner}^{2,67}} + \Delta t_{pump} \quad (7)$$

where α is the average well inclination angle; $Grad$ is the temperature gradient, grad/m; H_p - vertical depth of pump suspension, m; T_{res} - reservoir temperature, °C; H_{res} - vertical depth of the reservoir roof, m; $d_{str,inner}$ - inner diameter of production string, m; $d_{tub,inner}$ - inner diameter of tubing, m. The temperature distribution along the borehole is shown in Figure 4.

Temperature rise at the pump outlet (45) is determined with Equation 8:

$$\Delta t_{pump} = \frac{g H_{pump}}{c} \left(\frac{1}{\eta_{pump} \eta_{motor}} - \frac{0,5}{\eta_{pump}}\right) \quad (8)$$

where H_{pump} - is the head created by the pump, m; η_{pump} - pump efficiency, shares; η_{motor} - motor efficiency equal to 0.9.

A modernized Poetman-Carpenter methodology can be used to determine the pressure distribution in the well. Modernization includes modified equations to account

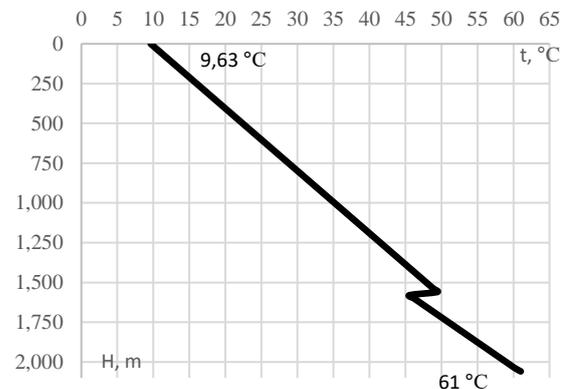


Figure 4. Flow temperature distribution along the borehole

for separated and dissolved gas in the tubing string after the ESP unit. The separation factor of the pump, which is used for modernization, is determined with Equation 9:

$$K_{sep} = \frac{1}{\left(\frac{1+0,75 \cdot \frac{Q_{liq}}{86400 \cdot (0,02 \cdot \left(\frac{d_{str,inner}^2 - d_{pump}^2}{4} \right))}}{\right)} \right)} \quad (9)$$

The Poetman-Carpenter methodology is used to find the pressure gradient in the well, Equation 10:

$$\frac{dP}{dH} = \rho_{mix} \cdot 9,81 \cdot 10^{-6} \cdot \cos \alpha + \frac{[f \cdot Q^2 \cdot (1 - \beta_{water})^2 \cdot M_{mix}^2]}{(2,3 \cdot 10^{15} \cdot \rho_{mix} \cdot d_{tub,inner}^5)} \quad (10)$$

where ρ_{mix} is mixture density, kg/m³; β_{water} - water cut, shares; M_{mix} - specific density of the gas-liquid mixture, kg/m³. These parameters are found according to the Poetman-Carpenter methodology.

Calculations are executed in a "top-down" manner within the tubing to the point of saturation pressure, and similarly, within the production string, extending from the pump inlet pressure to the bottomhole pressure (39). By employing the iterative method (46) key operational parameters of the well can be determined. Through numerical integration, we ascertain the pressure distribution along the tubing. Initially, pressures at designated depths are computed via a mathematical interpolation technique, subsequently linking these pressures with new temperature values. This process involves multiple iterations to accurately determine the pressure distribution throughout the depth of both the tubing and production string, culminating in the precise calculation of the true bottom hole pressure, P_{bot} . The derived pressure distribution along the wellbore is depicted in Figure 5.

Then, through a well productivity coefficient K_{prod} , the true flow rate of the well is found. Vogel correction (47), in Equation 11, is also introduced in the calculations which takes into account the movement of aerated liquid flow.

$$Q = \begin{cases} K_{prod} \cdot (P_{res} - P_{bot}), & \text{if } P_{bot} > P_{sat} \\ K_{prod} \cdot (P_{res} - P_{bot}) + \frac{K_{prod} \cdot P_{sat}}{1,8}, & \\ \cdot \left(1 - 0,2 \cdot \frac{P_{bot}}{P_{sat}} - 0,8 \cdot \left(\frac{P_{bot}}{P_{sat}} \right)^2 \right), & \text{if } P_{bot} \leq P_{sat} \end{cases} \quad (11)$$

The calculated flow rate value is reintegrated at the beginning of the calculation, and the entire algorithm is iterated until the desired precision is achieved. Subsequently, the equilibrium conditions for the formation of crystalline hydrates are established utilizing the calculation methodologies shown in the referenced monographs (48, 49). These methodologies are specifically tailored for application to both natural and associated petroleum gases.

First, using Equation 12, the pressure of gas hydrate formation p_m^0 (in MPa) is calculated (p_m^0 is determined

for temperature $T_0 = 273.15$ K). Equation 12 is applicable for hydrates of cubic structure II, which, in turn, are typical for oil and gas condensate fields, which are considered in this paper.

$$\left[1 + p_m^0 (2,5y_{CH_4} + 1,4y_{CO_2} + 0,67y_{N_2} + 46,1y_{H_2S}) \right]^2 = \frac{1}{p_m^0 \left(\frac{y_{CH_4}}{231} + \frac{y_{C_2H_6}}{2,3} + \frac{y_{C_3H_8}}{0,176} + \frac{y_{i-C_4H_{10}}}{0,113} + \frac{y_{n-C_4H_{10}}}{1,6} + \frac{y_{CO_2}}{26,3} + \frac{y_{N_2}}{2323} + \frac{y_{H_2S}}{10,47} \right)} \quad (12)$$

where y is the mole fraction of each component of the gas mixture. The content of each component in the gas is also determined after field studies or according to field development guidelines.

The desired pressure (p_m) at temperatures higher than 273.15 K is found using sets of reference curves of hydrate formation proposed by Semenov et al. (50) by the following Equation 13.

$$p_m = \frac{p_m^0}{z} \exp \left(A_1 \left(\frac{1}{T_0} - \frac{1}{T} \right) \right) \quad (13)$$

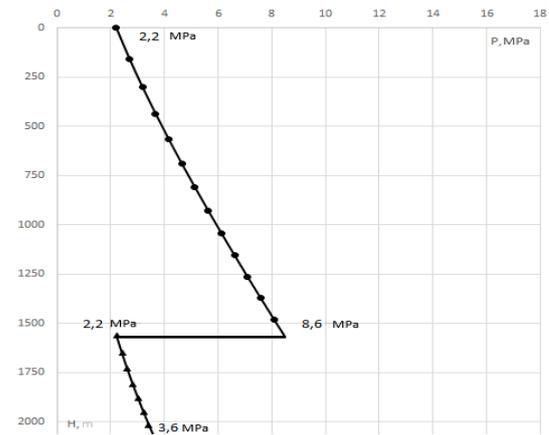


Figure 5. Pressure distribution along the wellbore obtained with modernized Poetman-Carpenter method

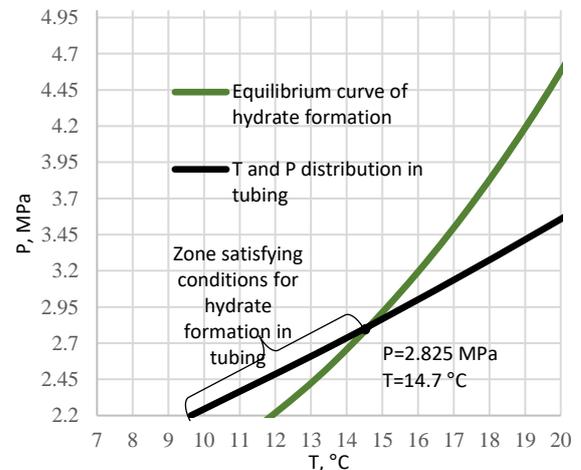


Figure 6. Determination of hydrate formation interval in tubing

z^0 , z are gas compressibility coefficient for p_m^0 and T_0 conditions and calculated conditions, respectively; A_1 - empirical coefficient (50).

Employing this methodology yields an equilibrium condition curve for hydrate formation plotted in pressure-temperature (P-T) coordinates, as illustrated in Figure 6. The point where this curve intersects with another indicates the initial depth at which gas hydrate begins to form within the tubing.

To calculate the depth of ARPD formation in the well, it is necessary to find the wax crystallization temperature distribution (51). The values of pressure and volume gas content are taken from previously calculated Poetman-Carpenter method. The dependence of paraffin crystallization temperature on depth is presented in Equation 14:

$$T_{wax}(P_i) = t_0 + 0,2 P_i - 0,1 V_{gas.sat,i} \quad (14)$$

where t_0 is found by Equation 15 ($^{\circ}\text{C}$); P_1 and $V_{gas.sat,i}$ - borehole section pressure (MPa) and residual gas saturation of oil (m^3/m^3), determined with the Poetman-Carpenter methodology.

$$t_0 = 11,398 + 34,084 \lg C_{wax} \quad (15)$$

where C_{wax} - mass content of paraffin, %.

The outcomes of this research are depicted in Figure 7. The algorithm introduced is slated for enhancement through the incorporation of supplementary differential equations. These equations aim to ascertain the duration required for the formation of organic deposits within the borehole space, as well as the extraction time for the fluid lifted from the borehole to the wellhead.

This advancement seeks to refine the predictive accuracy of the algorithm, providing a more comprehensive understanding of the dynamics involved in the deposition and extraction processes within the well operation.

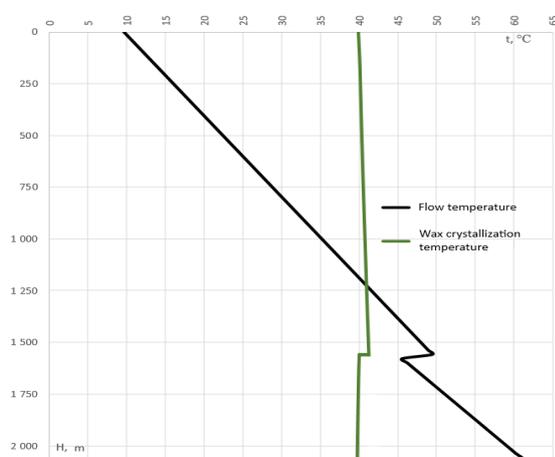


Figure 7. Determining the depth of ARPD formation in the well

2. 2. Methodology for Determining Thermobaric Conditions of Gas Hydrates Formation in Presence of Paraffin

The concurrent formation of crystalline hydrates and waxes poses a substantial challenge to oil production infrastructure. The accumulation of gas hydrate and wax deposits within pipelines and wells can significantly diminish productivity and, in severe cases, entirely obstruct the flow of liquids (27, 41, 52). Therefore, investigating this phenomenon is crucial for devising effective strategies for the prevention of deposit formation and for the efficient removal of such deposits during the operational lifecycle of oil fields.

In this section, we delve into the methodology employed in laboratory experiments designed to evaluate the impact of wax on gas hydrate formation conditions. By varying the mass content of dissolved paraffin, it is feasible to delineate the correlation between the conditions of gas hydrate formation and the quantity of paraffin present in the system. The experiments were conducted using the specialized reactor-autoclave Gas Hydrate Autoclave GHA 350 and a model gas preparation system, enabling the investigation of hydrate formation at temperatures ranging from -10°C to 60°C and pressures up to 35 MPa, as depicted in Figure 8. This research methodology was meticulously developed and validated by the personnel of the Scientific Center “Arctic” at the Empress Catherine II Saint Petersburg Mining University.

The laboratory experiment methodology, inspired by the approach detailed by Semenov et al. (53), is outlined as follows:

1. Preparation of Paraffin Solution in Kerosene:

- Aviation kerosene TS-1 (GOST 10227-86) with paraffin P-2 (GOST 23683-89) dissolved in it serves as the simulated hydrocarbon medium.
- Initially, 125 ml of the kerosene solution containing paraffin and 200 ml of water are utilized to



Figure 8. Laboratory complex for gas hydrate research: A. Autoclave GHA-350; B. Top-driven stirrer; C. Huber Ministat 240 thermostat; D. Gas boosters with maximum pressures of 40.0 MPa (left) and 15.0 MPa (right); E. Model gas preparation system.

create a phase interface within the camera lens. The mass for 125 ml of kerosene is accurately measured, considering the container's mass, yielding 96.71 g of kerosene. Following the measurement of the requisite volume, the container is sealed to prevent the vaporization of kerosene.

- The necessary quantity of paraffin is determined based on its desired mass content in the kerosene, ranging from 2-10% wt. The precise amount of paraffin wax is calculated using the appropriate proportion.

- The paraffin is then integrated into the pre-measured kerosene, followed by thorough mixing. To ensure containment, containers with screw-on lids, specifically designed for reagents, are employed to maintain airtight conditions upon heating.

- To expedite the dissolution of paraffin wax in kerosene, the container is placed within an oven preheated to temperatures of 50-60°C. This heating facilitates the complete dissolution of paraffin into kerosene within 30 minutes.

2. Autoclave Preparation for the Experiment:

- The autoclave's internal cavity is meticulously cleaned of residues from prior experiments using a solvent followed by alcohol, ensuring a contaminant-free environment for the new experiment.

- It is then loaded with 125 ml of kerosene solution and 200 ml of water. This specific ratio is crucial for visibly distinguishing phase separation in the first camera and gas-liquid interaction in the second camera.

- The experiment's requisite pressure is established, initially set to 40 bar for the commencement of the experiment.

- To generate the necessary emulsion, the solution is agitated at room temperature at a speed of 800 rpm for a duration of 30 minutes, rendering the autoclave primed for experimentation.

3. Conducting the Experiment to Identify the Equilibrium Point of Hydrate Formation:

- The system is pressurized with methane to slightly above the required pressure to ensure that, following methane dissolution and system equilibration, the pressure remains at or above the experiment's threshold. In this instance, the system is pressurized to 45 bar before stabilizing at 40 bar.

- The agitator's speed is adjusted to a moderate 100-200 rpm, facilitating the saturation of the liquid with methane.

- The system is then cooled to a temperature below the anticipated hydrate formation threshold. Maintaining the system in a super cooled state for a period allows for equilibrium establishment without the risk of hydrate formation due to oversaturation, as the stirrer operates at low speeds.

- To induce gas hydrate formation, the agitator speed is increased to 500 rpm, establishing a turbulent regime within the autoclave. In these oversaturated and

super cooled conditions, hydrate crystals form abruptly throughout the autoclave volume, marked by a significant pressure drop and a minor temperature increase—indicative of the exothermic nature of hydrate formation. Additionally, a noticeable increase in stirrer torque occurs as the forming hydrate crystals impede the agitation process.

- Identifying the equilibrium point involves gradual heating of the system at a constant rate. The equilibrium conditions for hydrate formation and dissociation align, but the stochastic nature of hydrate formation complicates direct equilibrium point determination during cooling. During dissociation, the equilibrium point is discerned at the pressure curve's inflection point. As the system is slowly heated, hydrate crystals disintegrate, releasing gas and elevating autoclave pressure. Once hydrate decomposition is complete, the pressure curve stabilizes, rising more gradually as pressure increases are solely attributed to heating. The experiment maintains a heating rate of 1°C every two hours, facilitating precise equilibrium point determination. To chart the equilibrium curve for hydrate formation, equilibrium points across various thermobaric conditions are identified.

2. 3. Methodology for Studying the Nucleation Time of Gas Hydrates in Presence of Paraffin

During laboratory studies on hydrate formation in the presence of paraffin, it was discovered that under conditions of low system pressures and low temperatures, the hydrate nucleation time significantly increases. This observation led to the conclusion that paraffin particles play a role in affecting the kinetics of the gas hydrate crystal formation process. Specifically, due to the pronounced super cooling at the lower pressures where hydrate formation occurs, paraffin particles tend to precipitate first within the liquid's volume. These particles, therefore, interfere with mass and heat transfer processes during the nucleation of gas hydrates (54), subsequently delaying the formation of initial gas hydrate agglomerates. This delay is part of an accumulative process concerning the potential energy involved in the agglomeration of crystalline hydrates. Paraffin, by postponing the formation of hydrate crystals, contributes to the sudden formation of large agglomerates when its role as a kinetic inhibitor of hydrate formation becomes overwhelmed (55).

The main factors that promote the formation of crystalline hydrates in the reactor are the super cooling of the system and the super saturation of the system with free gas. The super cooling aspect precisely defines the thermobaric conditions favorable for the nucleation process of gas hydrates. The kinetics of hydrate formation is greatly dependent on the level of super saturation. To accelerate the formation of gas hydrate deposits within the reactor, it is simply necessary to enhance the degree of gas super saturation in the system

by increasing the stirring speed. However, for the purpose of this experiment, the degree of super cooling and super saturation of the system is to be kept constant.

Accordingly, to investigate the effect of paraffin on the nucleation rate and kinetic characteristics of gas hydrate formation, the following methodology was developed. The process of nucleation and formation of crystalline hydrates is recognized as a relatively stochastic process, which introduces additional challenges in the investigation of this phenomenon (56). The variability in the time intervals for the formation of hydrate agglomerates necessitates the identification of dissociation conditions when determining the equilibrium point. Hydrate dissociation consistently occurs at a uniform rate, reaching equilibrium conditions simultaneously. However, this approach does not specifically cater to the study of the nucleation rate of hydrate crystals. Thus, the methodology proposed herein aims to evaluate the impact of varying concentrations of paraffin in the system on the nucleation rate of gas hydrate crystals.

The methodology for this laboratory experiment is outlined as follows:

1. Preparation of Paraffin Solution in Kerosene:

- The procedure for preparing the paraffin solution is identical to that described in section 2.2.

2. Preparation of the Autoclave for the Experiment:

- This step is conducted in the same manner as the autoclave preparation outlined in section 2.2.

3. Experiment on Determining the Nucleation Time of Gas Hydrates:

- To ensure the emulsion's stability and maintain a constant degree of system super saturation, the stirrer speed is set at 200 rpm. This speed is maintained uniformly throughout the experiment.

- Utilizing previously identified equilibrium points of hydrate formation, the intervals for system cooling are determined. The reactor is set to the required system pressure. Subsequently, the liquids and gas within the reactor are cooled from a starting temperature of 22°C to a temperature that is 1°C higher than the hydrate formation temperature over a period of 1.5 hours.

- The system is then further cooled to a temperature 2-3°C below the hydrate formation temperature, based on the specific conditions required. This cooling phase lasts for 30 minutes, after which the system is poised for the holding phase of the experiment.

- The system is maintained under these conditions for 10 hours (although this duration may be adjusted based on empirical observations), during which the pressure and temperature within the system, as well as the torque exerted by the stirrer, are meticulously recorded.

- Throughout the nucleation phase, the system's pressure and temperature are kept constant. Upon

reaching the equilibrium conditions, the hydrate formation process initiates, characterized by a reduction in pressure as the gas transitions into the hydrate molecules, and a notable increase in temperature due to the exothermic nature of hydrate agglomerate formation.

- The duration from the start of the experiment to the nucleation and formation of hydrates is accurately documented. Following this, the autoclave is reheated to room temperature, and the experiment is restarted. This repetition is crucial for reducing the stochastic variability inherent in the formation of gas hydrates. Initially, it is planned to conduct 10 experiments at a single paraffin mass content to calculate the average time required for the nucleation of crystalline hydrates.

3. RESULTS AND DISCUSSION

From these investigations, several key insights were uncovered. Primarily, a methodology for determining the depth of crystalline hydrates and Asphaltene Resin Paraffin Deposits (ARPD) formation within a producing well was introduced. This algorithm facilitates the precise identification of the depth at which organic deposits form, contingent on specific operational parameters of well equipment. Moreover, there exists potential to refine this method by integrating new relationships and differential equations. These equations would account for the gas hydrates' nucleation time in the presence of paraffin within the flow, deriving their foundation from the laboratory studies executed in this research.

Secondly, the acquisition of the equilibrium hydrate formation curve, illustrated in Figure 9, was for a system containing a 5% paraffin content. Additionally, data points for a 2% paraffin content, with all other parameters of the gas-liquid system and the autoclave remaining constant, were also derived. It was determined that paraffin does not influence the thermobaric conditions necessary for the formation of crystalline hydrates.

Thirdly, patterns regarding the nucleation time of hydrate crystals in systems containing paraffin were identified. The impact of paraffin on the nucleation time was specifically assessed for systems with 2 and 5% mass content of paraffin solution in kerosene, revealing nuanced insights into how paraffin presence affects the initial stages of hydrate crystal formation.

Several hypotheses were corroborated through the study, leading to the identification of distinct patterns:

- The nucleation process and the formation of crystalline hydrates exhibit a stochastic nature, with the time required ranging from 30 minutes to several hours, as illustrated in Figures 10 and 11.

- The modeling compound demonstrates a memory effect (57), wherein successive experiments yield nearly identical nucleation times. This memory

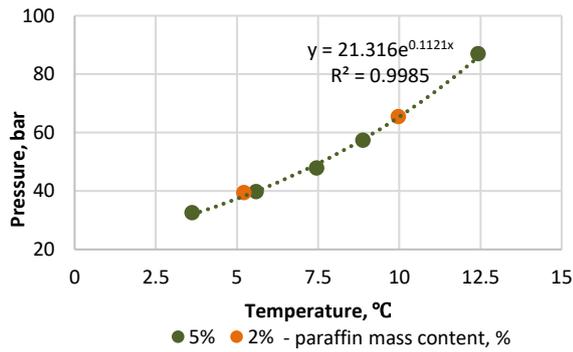


Figure 9. Equilibrium curve of hydrate formation obtained during the experiment

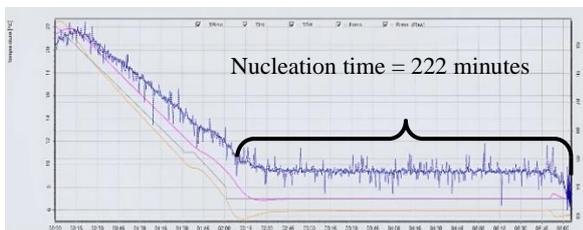


Figure 10. Equilibrium curve of hydrate formation obtained during the experiment

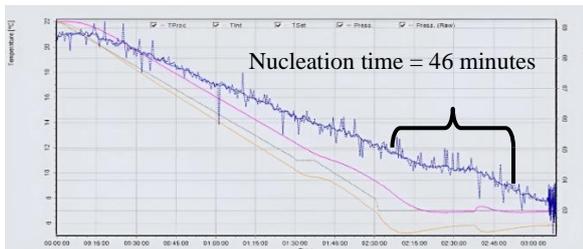


Figure 11. Hydrate nucleation less than one hour

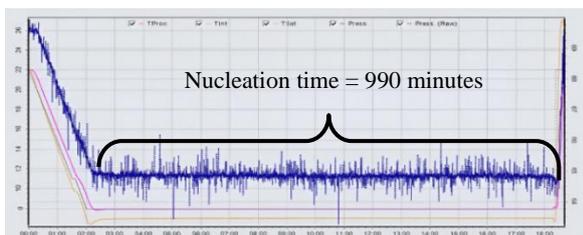


Figure 12. Equilibrium curve of hydrate formation obtained during the experiment

effect also facilitates the accelerated formation of gas hydrates in subsequent experiments, highlighting an intriguing aspect of the compound's behavior under experimental conditions.

- The degree of super cooling has a significant impact on the nucleation time. Decreasing the super cooling temperature from 3 to 2°C led to a fivefold increase in nucleation time, as depicted in Figure 12.

- The presence of wax in the mixture markedly prolongs the nucleation time of gas hydrates. With a 2% wax mass content, the average nucleation time was recorded at 94 minutes, whereas for a wax content of 5%, the nucleation time extended to 826 minutes. This represents a nine-fold increase in average nucleation time. Such variations underscore the role of increased wax content in impeding mass and heat transfer among hydrate molecules; thus inhibiting the formation of sizable gas hydrate agglomerates. This discovery underscores the importance of considering the inhibitory effects of wax on the formation of gas hydrate deposits during oil extraction processes. The outcomes of these investigations are summarized in Tables 1 and 2.

The variability observed in Tables 1 and 2 may be attributed to both the stochastic nature of the nucleation process and the memory effect. Specifically, the water molecules' hydrogen bonds surrounding the guest molecule, following hydrate decomposition, are preserved. Upon subsequent cooling, these bonds more readily serve as crystallization nuclei, thereby shortening the induction period (58).

To rigorously assess the impact of paraffin, a null hypothesis was posited prior to experimentation: "The hydrate nucleation time for a mixture with 2% wax mass content is equivalent to that for a mixture with 5% wax mass content," with a p-value threshold set at 0.05. This hypothesis was examined utilizing the SciPy library within the Python programming language.

The null hypothesis was ultimately refuted, as the statistical analysis yielded a p-value of 0.035, against the preset threshold of 0.05. This outcome indicates a significant difference in the hydrate nucleation times between mixtures with 2 and 5% wax mass content. Consequently, it is inferred that paraffin acts as a natural kinetic inhibitor in the formation of gas hydrates, highlighting its potential utility in managing gas hydrate formation in oil production contexts.

TABLE 1. Results of 10 experiments with 2% paraffin

Hydrates nucleation time	Experiment number									
	1	2	3	4	5	6	7	8	9	10
From equilibrium conditions of hydrate formation, min	222	240	32	43	46	35	19	19	239	40

TABLE 2. Results of 10 experiments with 5% paraffin

Hydrates nucleation time	Experiment number									
	1	2	3	4	5	6	7	8	9	10
From equilibrium conditions of hydrate formation, min	53	240	57	148	2645	1945	27	540	440	2160

4. CONCLUSION

The research delineated in this manuscript pertains to the elucidation of conditions conducive to the simultaneous formation of gas hydrates and asphaltene-resin-paraffin deposits (ARPD) within wellbores. The authors have developed a comprehensive methodology encompassing an extensive range of calculations. This methodology initiates with the determination of the thermobaric conditions prevailing inside the wellbore and culminates in the analysis of equilibrium conditions pertinent to hydrate formation alongside the crystallization temperature of paraffin. This study meticulously considers vital aspects of the reservoir fluid production process, including gas saturation, the distribution and dispersion of free gas within the flow, oil shrinkage, the presence of free water, and the impact of fluid viscosity on the operational parameters of Electric Submersible Pumps (ESP). This approach ensures a holistic understanding of the phenomena under study, integrating crucial factors that influence the formation of these deposits.

The computational algorithm introduced in this article holds promise for enhancing hydrocarbon production methodologies. Additionally, there is scope for the refinement and modernization of this algorithm to better meet the evolving needs of the industry.

During the course of laboratory experiments, it was ascertained that paraffin plays a significant role in influencing the kinetic conditions pertinent to the nucleation of gas hydrates. Notwithstanding, evidence suggests that paraffin embedded within kerosene does not alter the thermobaric conditions requisite for the formation of gas hydrates. Thus, paraffin emerges as a natural kinetic inhibitor, given that the nucleation rate of hydrates at a designated super cooling temperature escalates with an increase in its concentration within kerosene. Consequently, incorporating the nucleation timing in the context of wax presence facilitates the formulation of a temporal operational model for wells. This model, predicated on differential equations, juxtaposes the duration of fluid evacuation from the wellbore against the formation timeline of gas hydrate deposits. The implementation of this model is aimed at establishing operational conditions within the wellbore that preempt the formation of gas hydrates, particularly in regions where combating their accumulation proves most challenging.

Future experimental endeavors aimed at pinpointing the most likely nucleation interval for hydrate formation necessitate a methodological recalibration to mitigate the 'memory effect' observed during the reformation of gas hydrates. Within the ambit of this study, experiments demonstrated a prolongation in nucleation timing when the mixture was preheated to 35-40°C before subsequent re-cooling. It is noteworthy that the memory effect could be contingent upon the provenance of the water and its dwell time subsequent to hydrate dissociation.

The insights garnered from this research elucidate the formation mechanisms of crystalline hydrates and paraffin deposits, paving the way for the innovation of more efficacious strategies to thwart their accumulation in oil wells. This advancement harbors the potential to improve efficiency, curtail maintenance and operational expenditures, and engender more consistent production outcomes for oil entities.

5. REFERENCES

1. Mastronardo E, La Mazza E, Palamara D, Piperopoulos E, Iannazzo D, Proverbio E, et al. Organic salt hydrate as a novel paradigm for thermal energy storage. *Energies*. 2022;15(12):4339. <https://doi.org/10.3390/en15124339>
2. Zhang J, Li C, Shi L, Xia X, Yang F, Sun G. The formation and aggregation of hydrate in W/O emulsion containing different compositions: A review. *Chemical Engineering Journal*. 2022;445:136800. <https://doi.org/10.1016/j.cej.2022.136800>
3. Semenov AP, Stoporev AS, Mendgaziev RI, Gushchin PA, Khlebnikov VN, Yakushev VS, et al. Synergistic effect of salts and methanol in thermodynamic inhibition of sII gas hydrates. *The Journal of Chemical Thermodynamics*. 2019;137:119-30. <https://doi.org/10.1016/j.jct.2019.05.013>
4. Borisova NN, Rozhin II. Method for determining the mass flow for pressure measurements of gas hydrates formation in the well. *Журнал Сибирского федерального университета Серия «Математика и физика»*. 2021;14(2):193-203. <https://doi.org/10.17516/1997-1397-2021-14-2-193-203>
5. Raupov I, Burkhanov R, Lutfullin A, Maksyutin A, Lebedev A, Safiullina E. Experience in the application of hydrocarbon optical studies in oil field development. *Energies*. 2022;15(10):3626. <https://doi.org/10.3390/en15103626>
6. Nikolaichuk L, Ignatiev K, Filatova I, Shabalova A. Diversification of Portfolio of International Oil and Gas Assets using Cluster Analysis. *International Journal of Engineering, Transactions A: Basics*. 2023;36(10):1783-92. <https://doi.org/10.5829/IJE.2023.36.10A.06>
7. Tananykhin D, Struchkov I, Khormali A, Roschin P. Investigation of the influences of asphaltene deposition on oilfield development using reservoir simulation. *Petroleum Exploration*

- and Development. 2022;49(5):1138-49. [https://doi.org/10.1016/S1876-3804\(22\)60338-0](https://doi.org/10.1016/S1876-3804(22)60338-0)
8. Duryagin V, Nguyen Van T, Onegov N, Shamsutdinova G. Investigation of the selectivity of the water shutoff technology. *Energies*. 2022;16(1):366. <https://doi.org/10.3390/en16010366>
 9. Raupov I, Rogachev M, Sytnik J. Design of a polymer composition for the conformance control in heterogeneous reservoirs. *Energies*. 2023;16(1):515. <https://doi.org/10.3390/en16010515>
 10. 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. <https://doi.org/10.31897/PMI.2021.3.15>
 11. 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. <https://doi.org/10.31897/PMI.2020.4.5>
 12. Van TN, Aleksandrov AN, Rogachev MK. An extensive solution to prevent wax deposition formation in gas-lift wells. *Journal of Applied Engineering Science*. 2022;20(1):264-75. <https://doi.org/10.1007/s13202-022-01598-8>
 13. Nguyen VT, Pham TV, Rogachev MK, Korobov GY, Parfenov DV, Zhurkevich AO, et al. A comprehensive method for determining the dewaxing interval period in gas lift wells. *Journal of Petroleum Exploration and Production Technology*. 2023;13(4):1163-79.
 14. Martins JR, da Cunha Ribeiro D, Pereira FdAR, Ribeiro MP, Romero OJ. Heat dissipation of the Electrical Submersible Pump (ESP) installed in a subsea skid. *Oil & Gas Science and Technology–Revue d'IFP Energies nouvelles*. 2020;75:13. <https://doi.org/10.2516/ogst/2020009>
 15. Manabe R, Wang Q, Zhang H-Q, Sarica C, Brill JP, editors. A mechanistic heat transfer model for vertical two-phase flow. *SPE Annual Technical Conference and Exhibition?*; 2003: SPE.
 16. Gao Y, Cui Y, Xu B, Sun B, Zhao X, Li H, et al. Two phase flow heat transfer analysis at different flow patterns in the wellbore. *Applied thermal engineering*. 2017;117:544-52. <https://doi.org/10.1016/j.applthermaleng.2017.02.058>
 17. Serbin DV, Dmitriev AN. Experimental research on the thermal method of drilling by melting the well in ice mass with simultaneous controlled expansion of its diameter. *Записки Горного института*. 2022;257:833-42. <https://doi.org/10.31897/PMI.2022.82>
 18. Wang X, Wang Z, Deng X, Sun B, Zhao Y, Fu W. Coupled thermal model of wellbore and permafrost in Arctic regions. *Applied Thermal Engineering*. 2017;123:1291-9. <https://doi.org/10.1016/j.applthermaleng.2017.05.186>
 19. Merey S, Aydin H, Eren T. Design of electrical submersible pumps in methane hydrate production wells: A case study in Nankai trough methane hydrates. *Upstream Oil and Gas Technology*. 2020;5:100023. <https://doi.org/10.1016/j.upstre.2020.100023>
 20. Mardashov DV, Bondarenko AV, Raupov IR. Technique for calculating technological parameters of non-Newtonian liquids injection into oil well during workover. *Записки Горного института*. 2022;258:881-94. <https://doi.org/10.31897/pmi.2022.16>
 21. Mardashov DV. Development of blocking compositions with a bridging agent for oil well killing in conditions of abnormally low formation pressure and carbonate reservoir rocks. *Записки Горного института*. 2021;251:667-77. <https://doi.org/10.31897/pmi.2021.5.6>
 22. Fu W, Wang Z, Sun B, Ji C, Zhang J. Multiple controlling factors for methane hydrate formation in water-continuous system. *International Journal of Heat and Mass Transfer*. 2019;131:757-71. <https://doi.org/10.1016/j.ijheatmasstransfer.2018.10.025>
 23. Aman ZM, Di Lorenzo M, Kozielski K, Koh CA, Warriar P, Johns ML, et al. Hydrate formation and deposition in a gas-dominant flowloop: Initial studies of the effect of velocity and subcooling. *Journal of Natural Gas Science and Engineering*. 2016;35:1490-8. <https://doi.org/10.1016/j.jngse.2016.05.015>
 24. Zhou S, Chen X, He C, Wang S, Zhao S, Lv X. Experimental study on hydrate formation and flow characteristics with high water cuts. *Energies*. 2018;11(10):2610. <https://doi.org/10.3390/en11100000>
 25. Guimin Y, Hao J, Qingwen K. Study on hydrate risk in the water drainage pipeline for offshore natural gas hydrate pilot production. *Frontiers in Earth Science*. 2022;9:816873. <https://doi.org/10.3389/feart.2021.816873>
 26. Kuzmin AM, Buslaev GV, Morenov VA, Tseneva SN, Gavrilov NA. Improving the energy-efficiency of small-scale methanol production through the use of microturboexpander units. *Записки Горного института*. 2022;258:1026-37. <https://doi.org/10.31897/PMI.2022.104>
 27. Plyushin P, Vyatkin K, Kozlov A. Development of a method for estimating thermal conductivity of organic deposits on the wax flow loop laboratory installation. *International Journal of Engineering, Transactions C: Aspects* 2022;35(6):1178-85. <https://doi.org/10.5829/IJE.2022.35.06C.09>
 28. Rogachev MK, Aleksandrov AN. Justification of a comprehensive technology for preventing the formation of asphalt-resin-paraffin deposits during the production of highly paraffinic oil by electric submersible pumps from multiformation deposits. *Записки Горного института*. 2021;250:596-605. <https://doi.org/10.31897/PMI.2021.4.13>
 29. Petrakov D, Loseva A, Alikhanov N, Jafarpour H. Standards for selection of surfactant compositions used in completion and stimulation fluids. *International Journal of Engineering, Transactions C: Aspects* 2023;36(9):1605-10. <https://doi.org/10.5829/IJE.2023.36.09C.03>
 30. Khaibullina KS, Sagirova LR, Sandyga MS. Substantiation and selection of an inhibitor for preventing the formation of asphalt-resin-paraffin deposits. *Periodico Tche Quimica*. 2020;17(34). https://doi.org/10.52571/ptq.v17.n34.2020.565_p34_pgs_541_551.pdf
 31. Nurgalieva KS, Saychenko LA, Riazi M. Improving the efficiency of oil and gas wells complicated by the formation of Asphalt-Resin-Paraffin deposits. *Energies*. 2021;14(20):6673. <https://doi.org/10.3390/en14206673>
 32. Бельский А, Моренов В, Купавых К, Сандыга М. Электроснабжение станции нагрева нефти в скважине от ветроэлектрической установки. *Энергетика Известия высших учебных заведений и энергетических объединений СНГ*. 2019;62(2):146-54. <https://doi.org/10.21122/1029-7448-2019-62-2-146-154>
 33. Li R, Huang Q, Zhu X, Zhang D, Lv Y, Larson RG. Investigation of delayed formation of wax deposits in polyethylene pipe using a flow-loop. *Journal of Petroleum Science and Engineering*. 2021;196:108104. <https://doi.org/10.1016/j.petrol.2020.108104>
 34. Юдин П, Богатов М. Моделирование процесса выпадения асфальтосмолопарафиновых веществ на внутренней поверхности насосно-компрессорных труб с покрытием и без на лабораторном циркуляционном стенде. *Нефтегазовое дело*. 2021;19(2):97-103. <https://doi.org/10.17122/ngdelo-2021-2-97-103>
 35. Bai J, Jin X, Wu J-T. Multifunctional anti-wax coatings for paraffin control in oil pipelines. *Petroleum Science*. 2019;16(3):619-31. <https://doi.org/10.1007/s12182-019-0309-7>
 36. Kovrigin L, Kukharchuk I. Automatic control system for removal of paraffin deposits in oil well in permafrost region by thermal

- method. *Chemical Engineering Research and Design*. 2016;115:116-21. <https://doi.org/10.1016/j.cherd.2016.09.028>
37. Голубев И, Голубев А, Лаптев А. Практика применения аппаратов магнитной обработки для интенсификации процессов первичной подготовки нефти. *Записки Горного института*. 2020;245:554-60. <https://doi.org/10.31897/PMI.2020.5.7>
 38. Sizikov AA, Vlasov VA, Stoporev AS, Manakov AY. Decomposition kinetics and self-preservation of methane hydrate particles in crude oil dispersions: experiments and theory. *Energy & fuels*. 2019;33(12):12353-65. <https://doi.org/10.1021/acs.energyfuels.9b03391>
 39. Stoporev AS, Manakov AY, Altunina LK, Bogoslovsky AV, Strelets LA, Aladko EY. Unusual self-preservation of methane hydrate in oil suspensions. *Energy & fuels*. 2014;28(2):794-802. <https://doi.org/10.1021/ef401779d>
 40. Wang W, Huang Q, Zheng H, Wang Q, Zhang D, Cheng X, et al. Effect of wax on hydrate formation in water-in-oil emulsions. *Journal of Dispersion Science and Technology*. 2020;41(12):1821-30. <https://doi.org/10.1080/01932691.2019.1637751>
 41. Liu Y, Meng J, Lv X, Ma Q, Shi B, Wang C, et al. Investigating hydrate formation and flow properties in water-oil flow systems in the presence of wax. *Frontiers in Energy Research*. 2022;10:986901. <https://doi.org/10.3389/fenrg.2022.986901>
 42. Zi M, Wu G, Wang J, Chen D. Investigation of gas hydrate formation and inhibition in oil-water system containing model asphaltene. *Chemical Engineering Journal*. 2021;412:128452. <https://doi.org/10.1016/j.cej.2021.128452>
 43. Hang J, Bai L, Zhou L, Jiang L, Shi W, Agarwal R. Inter-stage energy characteristics of electrical submersible pump under gassy conditions. *Energy*. 2022;256:124624. <https://doi.org/10.1016/j.energy.2022.124624>
 44. Bulgarelli NAV, Biazussi JL, Verde WM, Perles CE, de Castro MS, Bannwart AC. Experimental investigation of the electrical submersible pump's energy consumption under unstable and stable oil/water emulsions: A catastrophic phase inversion analysis. *Journal of Petroleum Science and Engineering*. 2022;216:110814. <https://doi.org/10.1016/j.petrol.2022.110814>
 45. Zhou L, Hang J, Bai L, Krzemianowski Z, El-Emam MA, Yasser E, et al. Application of entropy production theory for energy losses and other investigation in pumps and turbines: A review. *Applied Energy*. 2022;318:119211. <https://doi.org/10.1016/j.apenergy.2022.119211>
 46. Zhang M, Jia A, Lei Z, Lei G. A Comprehensive Asset Evaluation Method for Oil and Gas Projects. *Processes*. 2023;11(8):2398. <https://doi.org/10.3390/pr11082398>
 47. Tomescu SG, Mălăel I, Conțiu R, Voicu S. Experimental Validation of the Numerical Model for Oil-Gas Separation. *Inventions*. 2023;8(5):125. <https://doi.org/10.3390/inventions8050125>
 48. Arzhanov MM, Malakhova VV, Mokhov II. Modeling thermal regime and evolution of the methane hydrate stability zone of the Yamal peninsula permafrost. *Permafrost and Periglacial Processes*. 2020;31(4):487-96. <https://doi.org/10.1002/ppp.2074>
 49. Yakushev V, Semenov A, Bogoyavlensky V, Medvedev V, Bogoyavlensky I. Experimental modeling of methane release from intrapermafrost relic gas hydrates when sediment temperature change. *Cold Regions Science and Technology*. 2018;149:46-50. <https://doi.org/10.1016/j.coldregions.2018.02.007>
 50. Semenov A, Mendgaziev R, Stoporev A, Istomin V, Tulegenov T, Yarakhmedov M, et al. Direct Measurement of the Four-Phase Equilibrium Coexistence Vapor-Aqueous Solution-Ice-Gas Hydrate in Water-Carbon Dioxide System. *International Journal of Molecular Sciences*. 2023;24(11):9321. <https://doi.org/10.3390/ijms24119321>
 51. Lekomtsev A, Kozlov A, Kang W, Dengaev A. Designing of a washing composition model to conduct the hot flushing wells producing paraffin crude oil. *Journal of Petroleum Science and Engineering*. 2022;217:110923. <https://doi.org/10.1016/j.petrol.2022.110923>
 52. Liu Y, Wu C, Lv X, Xu X, Ma Q, Meng J, et al. Evolution of morphology and cohesive force of hydrate particles in the presence/absence of wax. *RSC advances*. 2022;12(23):14456-66. <https://doi.org/10.1039/D2RA02266D>
 53. Semenov AP, Medvedev VI, Gushchin PA, Yakushev VS. Effect of heating rate on the accuracy of measuring equilibrium conditions for methane and argon hydrates. *Chemical Engineering Science*. 2015;137:161-9. <https://doi.org/10.1016/j.ces.2015.06.031>
 54. Liu Y, Wu C, Lv X, Du H, Ma Q, Wang C, et al. Hydrate growth and agglomeration in the presence of wax and anti-agglomerant: A morphology study and cohesive force measurement. *Fuel*. 2023;342:127782. <https://doi.org/10.1016/j.fuel.2023.127782>
 55. Tong S, Li P, Lv F, Wang Z, Fu W, Zhang J, et al. Promotion and inhibition effects of wax on methane hydrate formation and dissociation in water-in-oil emulsions. *Fuel*. 2023;337:127211. <https://doi.org/10.1016/j.fuel.2022.127211>
 56. Liu Z, Wang Z, Chen L, Chen L, Li X, Sun B. Experimental and modeling investigations of hydrate phase equilibria in natural clayey-silty sediments. *Chemical Engineering Journal*. 2022;449:137557. <https://doi.org/10.1016/j.cej.2022.137557>
 57. Gao Q, Zhao J, Guan J, Zhang C. Influence of the memory effect during CO₂/CH₄ mixed gas hydrate reformation process. *Fuel*. 2023;353:129249. <https://doi.org/10.1016/j.fuel.2023.129249>
 58. Wen Z, Yao Y, Luo W, Lei X. Memory effect of CO₂-hydrate formation in porous media. *Fuel*. 2021;299:120922. <https://doi.org/10.1016/j.fuel.2021.120922>

COPYRIGHTS

©2024 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 از این کار آنالیز زمان هسته زایی مولکول های هیدرات گاز در جریان نفت است. این تحقیق با هدف بررسی تأثیر ذرات پارافین بر زمان تشکیل رسوبات هیدرات در طی تولید مکانیزه نفت انجام شده است. درک و مدیریت صحیح تشکیل رسوبات آلی در فضای چاه می تواند هزینه های نگهداری تجهیزات را به میزان قابل توجهی کاهش دهد، ایمنی و پایداری تولید را افزایش دهد و بازده اقتصادی تولید هیدروکربن را بهبود بخشد. بخش اول مقاله روشی برای تعیین عمق تشکیل هیدرات های گازی و رسوبات آسفالت-رزین-پارافین (ARPD) در تولید چاه های نفت با حل سیستم های معادلات دیفرانسیل ترمویاریک ارائه می کند. مطالعات آزمایشگاهی بیشتری برای تجزیه و تحلیل زمان هسته زایی هیدرات های گازی در حضور پارافین انجام شد. این آزمایش در یک اتوکلاو فشار بالا ویژه انجام شد که در آن امکان ایجاد شرایط ترمویاریک لازم وجود دارد. همزن داخل اتوکلاو پراکندگی لازم سیستم را برای شبیه سازی جریان در چاه ایجاد می کند. در نتیجه آزمایش ها مشخص شد که ذرات پارافین مانع از تشکیل رسوبات هیدرات گاز شده و روند هسته زایی آنها را کند می کند. همچنین مشخص شد که افزایش محتوای پارافین در مخلوط به میزان 3 درصد، زمان هسته زایی هیدرات های گازی را 9 برابر افزایش می دهد. بر اساس تجزیه و تحلیل داده های موجود، نتیجه گیری می شود که برای توسعه مدل زمانی عملیات چاه تولیدی تحت شرایط هیدرات گاز و تشکیل ARPD به مطالعات عمیق تری نیاز است.



Design and Optimization of the Dual-Stator Axial-Field Flux-Switching Permanent Magnet Motor with High-Torque Density and Low-cost

F. Farrokh^a, A. Vahedi^b, H. Torkaman^c, M. Banejad^a

^a Faculty of Electrical Engineering, Shahrood University of Technology, Shahrood, Iran

^b Department of Electrical Engineering of Iran University of Science & Technology, Tehran, Iran

^c Faculty of Electrical Engineering, Shahid Beheshti University, Tehran, Iran

PAPER INFO

Paper history:

Received 27 November 2023

Received in revised form 06 January 2024

Accepted 10 January 2024

Keywords:

Dual-Stator Axial Field Flux-Switching Machine

High-Torque Density

High-efficiency

Low-cost

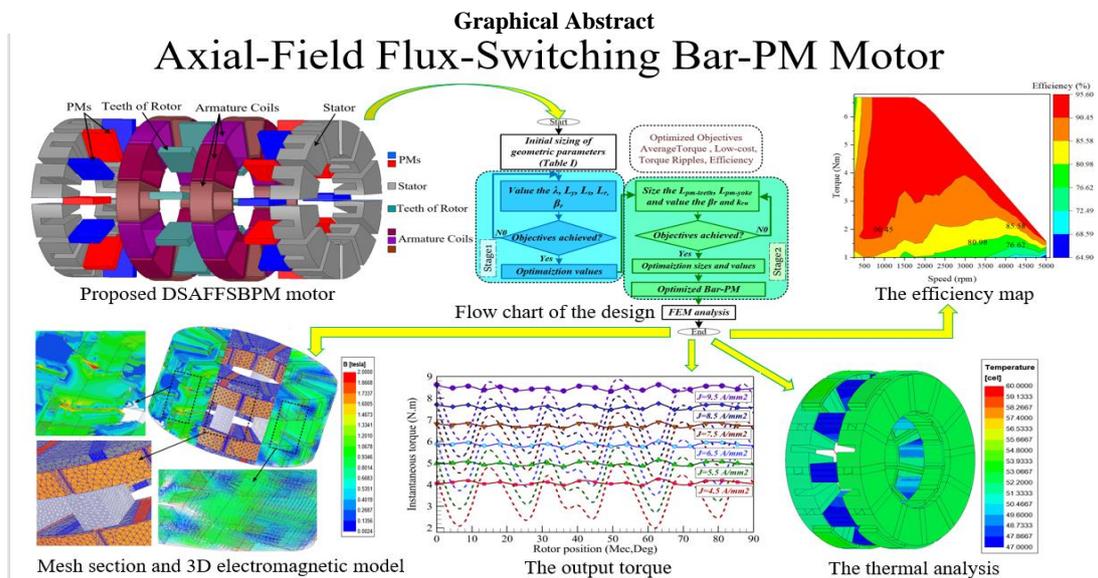
Thermal Stress

Electric Vehicles

ABSTRACT

In this paper, a new dual-stator axial field flux-switching permanent magnet (DSAFFSPM) motor has been proposed to improve the torque density and cost of the machine. In this topology, the 12-pole dual-stator has been located on both sides of one 10-pole inner-toothed rotor. The dual-stator has hosted permanent magnet (PM) type of Bar-PM and the coils. The novelty of this study is development of a technique that can be implemented on PM of the DSAFFSPM structure. In this regard, the proposed analytical design with a sizing equation has been presented and multi-objective optimization is employed to achieve the optimum size by Multi-Objective Genetic Algorithm (MOGA) method. The machine characteristics are acquired and analyzed utilizing the 3D finite element method (3D-FEM). A comparative study has been done to prove the superiority of the performance indices. This topology demonstrates the high-power density and the low vibration and noise due to lower torque ripple and cogging torque. Meanwhile, the Bar-PM topology has lower core loss and thermal stress due to high-efficiency. Consequently, the proposed model provides high torque density and low cost, specifically designed for electric vehicle (EVs) applications.

doi: 10.5829/ije.2024.37.07.a.14



*Corresponding Author Email: F.farrokh94@yahoo.com (F. Farrokh)

Please cite this article as: Farrokh F, Vahedi A, Torkaman H, Banejad M. Design and Optimization of the Dual-Stator Axial-Field Flux-Switching Permanent Magnet Motor with High-Torque Density and Low-cost. International Journal of Engineering, Transactions A: Basics. 2024;37(07):1357-68.

1. INTRODUCTION

Different types of technology machines include axial and radial flux machines. Studies showed that axial flux machines (AFPM) have higher torque and power density (1-3), as well as efficiency (4-6). Axial field flux-switching machines (AFFSPM) have limited torque density due to the location of PMs and coils, but still have higher torque and efficiency than radial flux machines and can reduce costs. New motor designs have been proposed in recent publications (7-10). One used non-rare-earth PM for high-torque density (11). While another achieves even higher torque density with rotor-excited PM (12). The AFFSPM design with E-core reduces cost and PM volume, but also has lower average torque compared to U-core and C-core designs (13). Kim et al. (14), AFFSPM have presented with a rotor middle and external dual-stator. Topologies try to maintain a phase shift of $\pm\pi$ electric radians, topology 1 (shifting stator 2) has the highest output torque compared to topology 2 (shifting rotor poles of stator 2) and topology 3 (shifting both stator 2 and rotor poles of stator 2).

Newly, three conventional DSAFFSPM structures with internal dual-rotor, internal single teeth rotor (ISTR), and external dual-rotor have been compared together for EV application. The findings show that, the internal single-rotor topology is a suitable candidate for EV application because it has high efficiency, and high torque density (15). Nevertheless, it has troubled higher cogging torque and torque ripple.

Own to reduction the reserve of rare-earth PM materials (such as NdFe35) and market monopoly have affected price fluctuation for further application of PM motors in EVs. Some researchers suggested non-rare-earth PM (such as ferrite) types (11). The results showed that compared with rare-earth PM materials have low performance. To further make it better torque density a type of hybrid excitation (both DC and AC windings) machine has been proposed (16-19). However, this leads to low torque, power densities, provision of the field excitation source, an increase in cost, and loss. Reducing the consumption of rare-earth PMs compromises the AFFSPM motor performance because the main flux structure is supplied by PM. For this reason, has made the issue of PM reduction less attention by researchers. Recently, reducing the consumption of rare-earth PMs and cost material has become a hot topic, particularly in PM machines (20-22).

The main objective of present work is to design a dual-stator axial-field flux-switching Bar-PM (DSAFFSBPM) motor with high-torque density and low-cost for EV application. Employing a technique that is aimed at the reducing value of rare-earth PM in the DSAFFSBPM, so that it can be implemented on the conventional ISTR-DSAFFSPM structure. The main contribution novelties of this paper are a novel technique

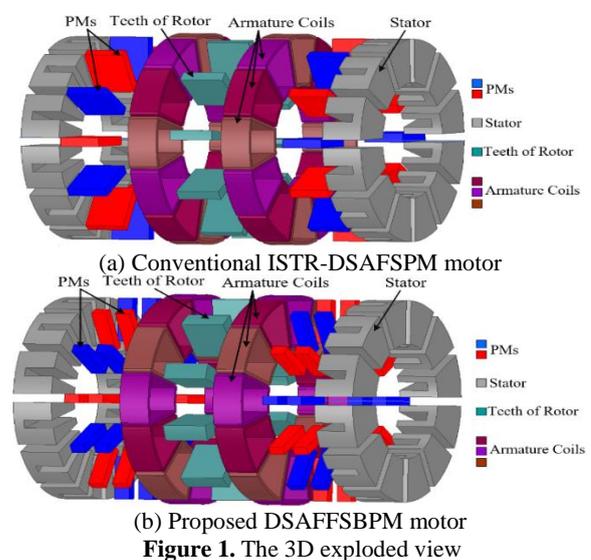
that has been implemented on PMs for the first time. Then, the technique aims to achieve multi-objective optimization, which involves reducing the increase in torque density, lowering costs, and decreasing thermal temperature using air channels. The challenges addressed by this paper are:

- 1) To be provided the high-torque density and the low-cost simultaneously.
- 2) Due to their doubly salient structure, the cogging torque is an unfavourable effect.
- 3) Due to their high air-gap flux density caused by flux focusing effects be suffered torque ripple.

2. THE STRUCTURE AND OPERATION PRINCIPLE

2. 1. Proposed New Structure Figure 1 shows the 3D exploded model. The configuration of the conventional ISTR-DSAFFSPM motor is illustrated in Figure 1. The DSAFFSBPM motor is also presented in Figure 1.b. In this structure, unlike the former types of AFPMs, no magnet exists in the rotor. Therefore, the stator hosted Bar-PM and coils. The only difference is that one slot comprises segments of rare-earth magnets as Bar-PM, which are circumferentially magnetized, and embedded precisely in the stator yoke and stator teeth of these slots, and tightened. The stator is created of 12 poles on each side. 12 concentric-coils twist the poles in two 3-phase complies alternately. The rotor of this topology is of toothed type.

2. 2. The Operation Principle To demonstrate the operational principles of this structure, a 2D model is proposed as shown in Figure 2. It illustrates four rotor positions for PM flux. These four special positions rotor a, b, c, and d comply with the four points of maximum



positive, zero, maximum negative, and zero flux of phase A, respectively. Figure 2(a); the rotor tooth T2 is placed almost in front of the stator tooth, which the winding A1 wraps. The magnetic direction of the PM is to the right, as a result, the direction of the magnetic flux is upward in winding A1. The magnetic flux passes the rotor tooth T1 and the air-gap, which enters the stator tooth S1. The flux-linkage A1 is positive maximum; this is dubbed the positive position. When the rotor rotates $\theta_r=9^\circ$ Mec in Figure 2(b), the rotor tooth T2 and the PM are aligned, so the flux in the winding becomes zero. This position is dubbed the zero position of the rotor. When the rotor rotates another at $\theta_r=9^\circ$ Mec in Figure 2(c), the rotor tooth T2 is almost in front of the stator tooth S2. The flux-linkage A1 is negative maximum; this is dubbed the negative position. In Figure 2(d), As the rotor rotates at $\theta_r=27^\circ$ Mec, the rotor tooth T2 is located in front of the slot; it means, it is placed in the second zero position; this process repeats to generate the periodic flux-linkage, and the back-EMF.

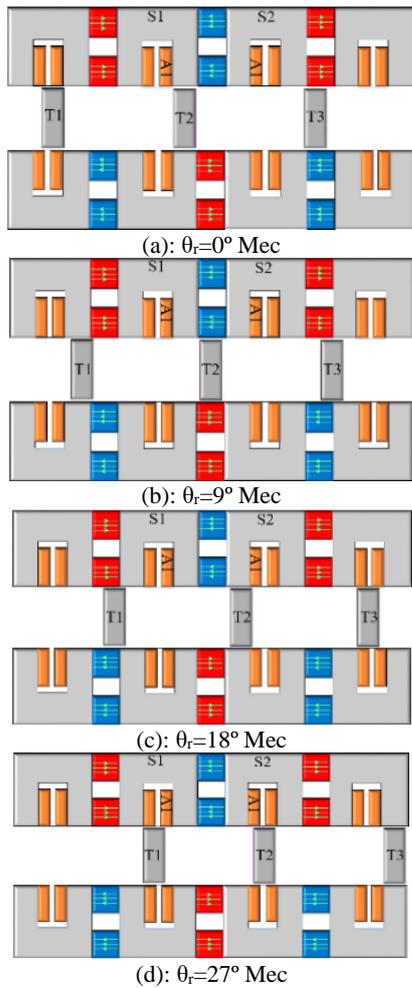


Figure 2. Four specific positions rotor of the 3-phase DSAFFSBPM motor

Inferred to Equation 1, when a sinusoidal current has followed the equations, can result in the back-EMF voltage is formed, hence electromagnetic torque was generated.

$$I_a = I_{max} \sin(\theta) \tag{1-a}$$

$$E_a = E_{max} \sin(\theta)$$

$$I_b = I_{max} \sin(\theta - 2\pi/3) \tag{1-b}$$

$$E_b = E_{max} \sin(\theta - 2\pi/3)$$

$$I_c = I_{max} \sin(\theta + 2\pi/3) \tag{1-b}$$

$$E_c = E_{max} \sin(\theta + 2\pi/3)$$

3. DESIGN PROCEDURE AND MULTI-OBJECTIVE OPTIMIZATION

3. 1. The Sizing Equations of DSAFFSBPM Motor

The output power equation of the motor for EVs, and the dimensions of the motor is deduced according to Farrokh et al. (15), the output power of the DSAFFSBPM motor be expressed as follows:

$$P_{out} = \frac{m}{T} \int_0^T e(t).i(t) dt \tag{2}$$

According to the current and the phase PM flux-linkage are sinusoidal, the output power writes as follows:

$$P_{out} = \frac{m}{T} \int_0^T E_m \sin(\frac{2\pi}{T}t) I_m \sin(\frac{2\pi}{T}t) dt \tag{3}$$

$$P_{out} = \frac{m}{2} E_m I_m \tag{4}$$

In which E_m is maximum voltage, and I_m is the current magnitude. The output power is derived by adding the motor efficiency is η as stated below:

$$P_{out} = \frac{m}{2} E_m I_m \eta \tag{5}$$

According to Equation 1, the EMF in the coils (e), when the armature is open-circuit and only the flux PM exists, is given as follows:

$$e = -\frac{d\psi_m}{dt} = -N_{ph} \frac{d\phi}{d\theta} \frac{d\theta}{dt} = -N_{ph} \frac{d\phi}{d\theta} \omega_r \tag{6}$$

where ψ_m term is the flux linkage of the phase. The term N_{ph} is the number of the winding turns per phase, ϕ_m term will be the flux-linkage of one turn, θ and ω_r are the rotor position and the angular speed of rotor, respectively. According to Equation 1, the flux-linkage is expressed as follows:

$$\phi_m = \phi_R \cos(P_r \theta) \tag{7}$$

where ϕ_R and P_r are the flux magnitude and the number of the rotor poles, respectively.

Substituting the Equation 7 into Equation 6, the back-EMF, yields as follows:

$$e = N_{ph} \omega_r P_r \varphi_R \sin(P_r \theta) = E_m \sin(P_r \theta) \quad (8)$$

According to Figure 2(a), this position is the maximum flux passing, therefore the air-gap area (A_g) placed between the stator tooth and the rotor pole is inferred as stated as follows:

$$A_g = \beta_r \times \frac{(D_{so} - D_{si})}{2} \quad (9)$$

where β_r is the pole pitch which is shown in Figure 3. D_{so} and D_{si} are the outer and inner diameters of the DSAFFSB-PM motor, respectively. τ_s is stator pole pitch. The average diameter of air-gap is given below:

$$\tau_s = \frac{\pi D_{so}}{P_s} \quad (10)$$

$$\beta_r = \beta_{st} + \beta_{pm} + \beta_t = \frac{\tau_s}{4} \quad (11)$$

where P_s is the number of the stator poles. By inserting Equations 10, and 11 into Equation 9, it can be determined as:

$$A_g = \frac{\pi (D_{so}^2 - D_{si}^2)}{8 P_s} \quad (12)$$

From Equation 13, the flux magnitude is calculated as:

$$\varphi_R = K_l K_g B_{gmax} \beta_i \frac{\pi}{8 P_s} (D_{so}^2 - D_{si}^2) \quad (13)$$

In which K_l is a leakage flux factor and K_g is the factor of the air-gap flux density distribution. B_{gmax} is the peak value of the air-gap flux density in the no-load mode. β_i is the area ratio of the stator tooth and stator tooth-slot unit. Substituting Equation 13 into Equation 8, E_m will be stated as follows:

$$E_m = N_{ph} \omega_r P_r K_l K_g B_{gmax} \beta_i \frac{\pi}{8 P_s} (D_{so}^2 - D_{si}^2) \quad (14)$$

Due to the concentration of the maximum electrical load in the smallest radius of axial-flux machines, the limit of this parameter is considered in the inner radius of the machine. Therefore, the electrical loading, denoted as A_s , is highest at this point. A_s is given below:

$$A_s = 2m N_{ph} \frac{I_{rms}}{\pi D_{si}} \quad (15)$$

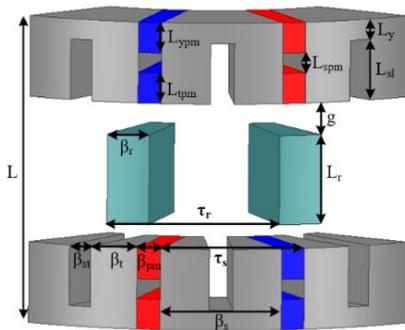


Figure 3. Design Parameters of the DSAFFSBPM

The maximum of sinusoidal current is given as:

$$I_m = \sqrt{2} I_{rms} = \frac{\sqrt{2}}{2} \frac{\pi A_s D_{si}}{m N_{ph}} \quad (16)$$

By converting the rotor speed to rpm infer to $\omega_r = 2\pi n_r / 60$ and by Substituting Equations 14 and 16 into Equation 5 yields:

$$P_{out} = \frac{\sqrt{2} \pi^3}{480} \frac{P_r}{P_s} K_l K_g \lambda (1 - \lambda^2) A_s B_{gmax} \beta_i D_{so}^3 n_r \eta \quad (17)$$

where λ is the ratio of the inner and outer stator diameters. A_s is the maximum current density and n_r is the speed of the rotor.

The output torque is calculated as follows:

$$T_{out} = \frac{3\pi^2 \sqrt{2}}{48} \frac{P_r}{P_s} K_l K_g \lambda (1 - \lambda^2) A_s B_{gmax} \beta_i D_{so}^3 \eta \quad (18)$$

A fixed outer diameter of the stator is a design limiting factor in an electrical vehicle, which impacts machine performance. Therefore, adopting the value optimal λ lead to the best electrical and magnetic loading. The λ as the objective function from the torque equation estimated as follows:

$$T_{out}(\lambda) = \lambda - \lambda^3 \quad (19)$$

$$\lambda_{max} = \frac{1}{\sqrt{3}} = 0.57 \quad (20)$$

Equation 21 is applied in opting for the internal diameter thus, maximize the average output torque. According to Zhao et al. (17), the outer diameter of the stator can be expressed as follows:

$$D_{so} = \sqrt[3]{\frac{480 P_{out} P_s}{\sqrt{2} \pi^3 P_r K_l K_g \lambda (1 - \lambda^2) A_s B_{gmax} \beta_i n_r \eta}} \quad (21)$$

The primary dimensions of the designed motors are listed in Table 1 and illustrated in Figure 3.

TABLE 1. Primary parameters of investigated DSAFFSBPM motor

Parameter	Symbol	Proposed Bar-PM	Optimized Bar-PM
Nominal power (kW)	P_n		1
Rated speed (rpm)	n_r		1500
Rated current (A/mm ²)	J		7.5
No. of phase	m		3
Slots / rotor poles	P_s / P_r		12/10
Stator outer diameter (mm)	D_{so}		140
Axial length (mm)	L		64
Air gap length (mm)	g		1
Stator pole pitch (deg)	β		30
Rotor pole pitch (deg)	β_r	12.5	15.5

PM pitch (deg)	β_r	6	
Rotor yoke with (mm)	L_y	5	4.5
Teeth magnet length (mm)	L_{tpm}	7.5	6
Yoke magnet length (mm)	L_{ypm}		7
Machine volume (L)	-		0.98

3. 2. Multi-Objective Optimization There are many multi-objective algorithms, some struggle to balance solution quality and diversity. However, the MOGA adjusts this balance by combining domain structure and evolutionary operation. MOGA considers both the quality and diversity of optimal solutions during optimization and has been applied successfully in various fields, including transportation (23-25). The MOGA be is utilized to optimize the electromagnetic performance of the DSAFFBPM motor based on objective function (OF). The GA adopts initial design parameters of the rotor, stator as well as PMs such as λ , L_y , L_s , L_r , $L_{pm-teeth}$, $L_{pm-yoke}$, and β_r to comprehensive analysis; when all design objectives are simultaneously and multi-dimensionally. Optimized objectives such as average torque (T_{av}), low-cost (*cost*), torque ripples (T_{ri}), and efficiency follow to find out a convergent response as follows.

$$\begin{cases} \text{Objective function:} \\ \max(T_{av}, \eta) \text{ and } \min(\text{cost}, T_{ri}) \end{cases} \quad (22)$$

$$\begin{cases} \text{Constraint:} \\ T_{av} \geq 5.87N.m, \eta \geq 92.96\% \\ \text{cost} \leq \$60, T_{ri} \leq 45.31\% \end{cases} \quad (23)$$

$$\begin{cases} \text{Variable stage1} = [\lambda, \beta_r, L_y, L_s, L_r] \\ \text{Boundary condition stage1:} \\ 0.55 \leq \lambda \leq 0.57, 6.28 \leq \beta_r \leq 10.46 \text{ deg} \\ 4.5 \leq L_y \leq 6.5 \text{ mm}, 20 \leq L_s \leq 21 \text{ mm} \\ 20 \leq L_r \leq 22 \text{ mm} \\ \text{Variable stage2} = [\beta_r, L_{ypm}, L_{tpm}, k_{cu}] \\ \text{Boundary condition stage2:} \\ 10.46 \leq \beta_r \leq 15.5 \text{ deg}, 7 \leq L_{ypm}, L_{tpm} \leq 8 \text{ mm} \\ 0.68 \leq k_{cu} \leq 0.7 \end{cases} \quad (24)$$

The GAO employs a multi-objective function to optimize its design. Eight design parameters can affect the performance of the genetic algorithm, as outlined by Equation 24 and Figure 4. As shown in Figure 5, after approximately 40 generations of genetic optimization, the two stages converge to the optimal solution. The GAO then optimizes eight key parameters to achieve four optimized objectives, resulting in improved performance. These parameters and associated performance improvements are listed in Table 2.

The design procedure of the DSAFFSBPM motor is illustrated in the form of a flowchart as shown in Figure

4. The parameters are purposefully optimized in two stages, which can provide an accurate response in a minimum time. In stage 1, multi-dimensional optimization is done with the parameters denoted and without reducing the PM value such as λ , L_y , L_s , L_r , $L_{pm-teeth}$, $L_{pm-yoke}$, and β_r are optimized. In stage 2, two affecting parameters β_r and slot fill factor (k_{cu}) are applied to optimize by reducing the PM value. The derivative results of MOGA have been visible in the optimized model as shown in Figure 5, which are marked with green bubble points.

In Equation 23, the variables T_{av} , η , and T_{ri} are analyzed using the conventional ISTR model. The *cost* of optimizing and reducing PMs is found to be lower when using the conventional ISTR model than the proposed Bar-PM model. In Equation 24, the parameters are limited in stage 1 based on the conventional ISTR model, and then in stage 2, they are restricted based on the considered cost.

The effect of MOGA on electromagnetic performance and both initial and optimized parameters are discussed. For investigation and comparison, both initial and optimized performances are recorded in Table 2. Based on electromagnetic performances, the analysis represents that T_{av} is improved by as much as 5.14%, decreased T_{ri} as much as 70.44% enhanced η by 6.37%, and reduced cost by 10.52%. As well as, based on MOGA, initial and optimized design parameters received are registered, in Table 1. The final optimized model is investigated further in the following sections.

4. FEM ANALYSIS RESULTS AND ELECTROMAGNETIC PERFORMANCE

After the design of the optimized structural model, the electromagnetic performance of the motors is assessed using 3D-FEM and numerical analysis.

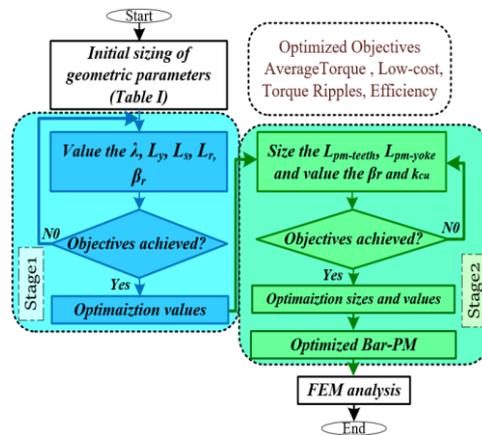
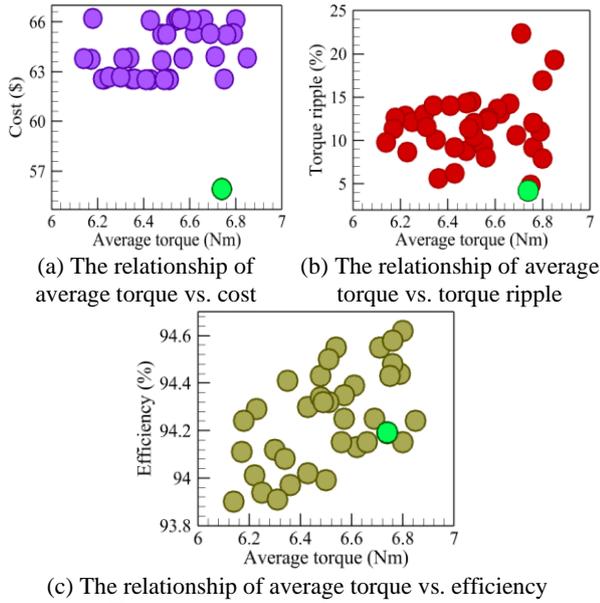


Figure 4. Flow chart of the design procedure for DSAFFSBPM motor

TABLE 2. Electromagnetic performance analysis of and optimized models

Model, Parameter	Initial design Bar-PM	Optimized design Bar-PM
T_{av} [Nm]	6.41	6.74
T_{ri} [%]	14.04	4.15
Cost [\$]	62.51	55.98
η [%]	94.13	94.19

**Figure 5.** The optimization results of the four optimization objectives

4. 1. Investigation on Open-Circuit Magnetic Field Distribution

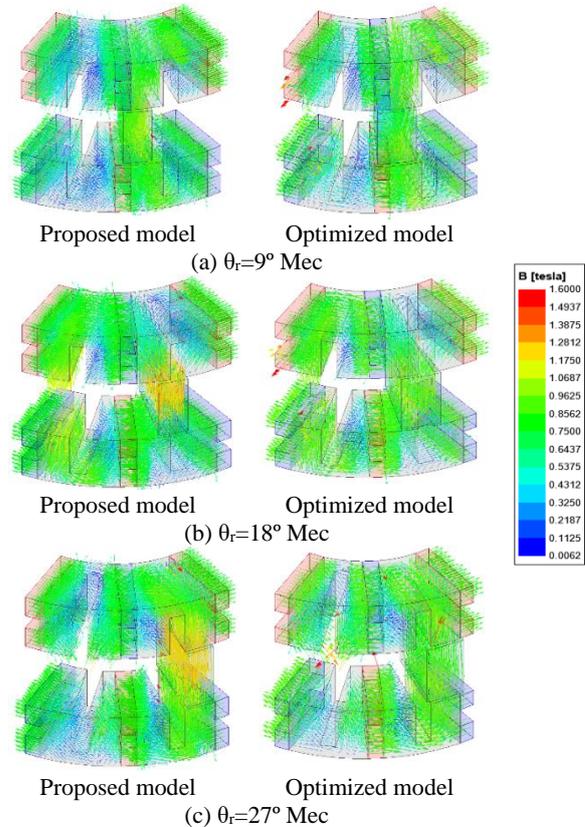
When the stator windings are in open-circuit status, the open-circuit field distribution is only performed through the Bar-PMs. Rotor rotation against 3-phase winding leads to the polarity of the flux and flux focusing. Finally, the concept of “switching flux” is employed in the motor. The $\theta_r=9^\circ$ Mec and $\theta_r=27^\circ$ Mec shown in Figures 6(a) and 6(c) address the status when the rotor teeth are located in the front of the stator teeth. In this status, the maximum flux linkage with the coil takes place respectively in negative and positive values with a magnetic field density of roughly 1.33T for the proposed Bar-PM model. Figure 6(b) illustrate $\theta_r=18^\circ$ Mec rotor position in which the coils ‘linkage flux achieves zero. The maximum flux density is for the optimized Bar-PM model that did not exceed 1.1T in the worst-case scenario. The concentration of flux is higher in the stator yoke than in the stator teeth. As a result, the length of the PM on the side of the stator yoke is slightly longer than the PM on the side of the stator teeth. This ensures proper distribution of flux in these areas. The

maximum flux concentration in the PMs in three different rotor modes did not exceed 1.06 T.

4. 2. Analysis of the Flux Linkage and Induced Voltage Characteristics

According to Equation 7, Figure 7, depicts the three-phase open-circuit flux-linkage in the DSAFFSPM models. The maximum flux-linkage amplitude is 0.047 Wb for the optimized Bar-PM model; then, the proposed Bar-PM model is 0.048 Wb. A 2.08% decrease linkage-flux in the optimized Bar-PM model compared to proposed Bar-PM model is due alter in slots. Figure 7 specifies where the phase difference of $2\pi/3$ of the electrical degree of 3-phase flux-linkage sinusoidal waveform. Therefore, confirms the operation of DSAFFSPM as an AC brushless machine.

In Figure 8, 3-phase induced back-EMF sinusoidally in open-circuit voltage is compared for the DSAFFSPM models at @1500 rotor speed. The back-EMF amplitudes of the proposed and optimized Bar-PM models are 77.57, and 73.89 V, respectively. According to Equations 7 and 8, it is expected that the amplitude of flux-linkage is less than 2.08% in the optimized Bar-PM model. Consequently, its amplitude of the induced voltage range is reduced by 4.74%.

**Figure 6.** The 3D open-circuit magnetic field distribution in three fundamental rotor positions of Bar-PM models

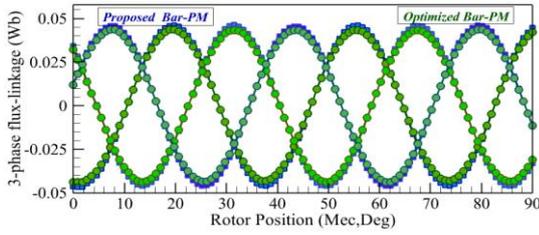


Figure 7. The flux-linkage sinusoidal waveform in the 3-phase DSAFFSPM motor

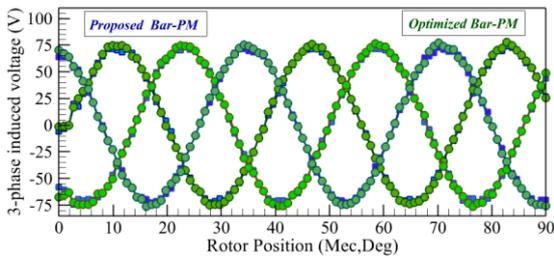


Figure 8. The back-EMF in the 3-phase DSAFFSM-B motor

4. 3. The Induced Voltage Harmonic Spectrum

The λ parameter and the magnet’s thickness must be adopted at a low-cost with promising harmonics spectrum. By considering $\lambda=0.57$, β_r increases amount of 24% and L_{pm} decreases 25% in the magnetization direction, the improvable voltage harmonics spectrum has opted for a low-cost. Based on the normalized domain values of the harmonic spectrum of the three presented models are manifested in Figure 9. The harmonic spectrum of the voltage is improved compared to the conventional ISTR. The harmonic order 5 of the conventional ISTR model equals 5%, while the harmonic order value is 1.2% for the proposed model. The optimized Bar-PM model is equal to 1.1%, which has decreased by 8.33%. The harmonic order 7 of the conventional ISTR model equals 2.7%. Meanwhile, the proposed and optimized model’s harmonic order value is under 1%. From the sight of the harmonics spectrum, the proposed motor has an acceptable condition.

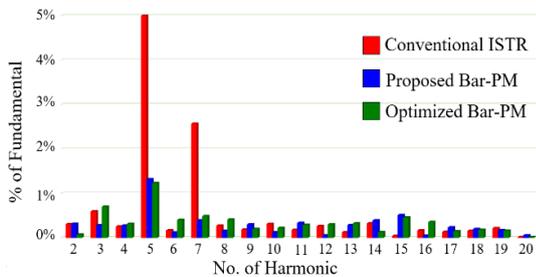


Figure 9. The harmonic spectrum of the back-EMF based on the normalized domain of the harmonics

4. 4. The Cogging Torque and Output Torque Analysis

In Figure 10, the DSAFFSPM models compare the cogging torque and output torque. The conventional ISTR and proposed Bar-PM have peak-to-peak cogging torque values of 2.52 N.m and 1.02 N.m, respectively. These values represent 44.21% and 15.91% of the average torque in the rated current, respectively. The optimized Bar-PM model has a peak-to-peak cogging torque value of 0.266 N.m, which is only 3.94% of the average torque in the rated current.

Equation 25 expresses the output torque ripple:

$$\%T_R = \frac{T_{max} - T_{min}}{T_{avg}} \times 100\% \tag{25}$$

where T_{avg} is the average torque, T_{min} is minimum torque, and T_{max} is the maximum torque, respectively. Using this technique, the reduction of cogging torque in an optimized Bar-PM motor is almost negligible. At a rated current density of 7.5 A/mm², the conventional ISTR, the proposed Bar-PM, and the optimized Bar-PM models exhibit average torque values of about 5.87, 6.41, and 6.74 N.m, respectively. In addition, these models have torque ripple values of 45.31%, 14.04%, and 4.15%, respectively.

Figure 11 demonstrates the output torque of DSAFFSPM models in the rated, the lower and the higher armature current for the rotor position. Comparing the average torques shows that the optimized Bar-PM model has higher average torque. Meanwhile, it has negligible torque ripple compared to the conventional ISTR.

Figure 12 depicts a relative operational index in rated ratio per each armature current density. It has defined as the torque average and percentage of torque ripple. It shows comparison of this index and illustrates that the optimized Bar-PM model has been located under the points. Thus, the optimized Bar-PM motor is designed for the best operational condition.

5. MODEL COMPARISON

5. 1. Comparison of the Mass and Costs The price of each material expressed in \$ and the quantity of

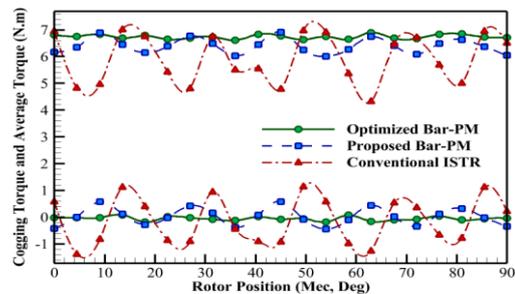


Figure 10. The cogging torque and output torque for the rotor mechanical position in three models

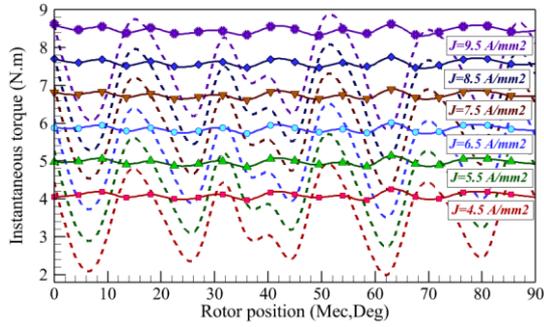


Figure 11. The output torque for the different armature currents as a function of the rotor position at @1500 rpm

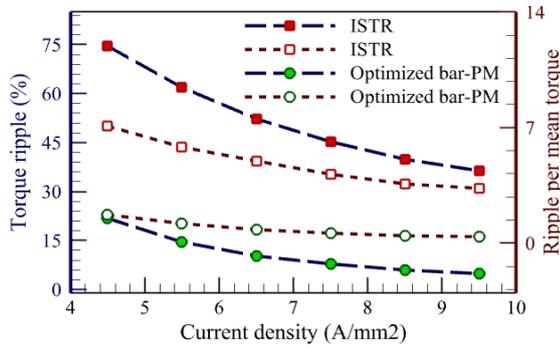


Figure 12. The torque ripple for the armature current density

materials utilized in kg are used to determine each machine's cost. Upon reviewing Table 3, it is clear that both the proposed and optimized Bar-PM models require significantly less PM material compared to the conventional ISTR models. Specifically, the proposed model uses 26.56% less while the optimized model uses 35.93% less. As a result, both models have a lower total cost, with the proposed model being up to 23.87% less expensive and the optimized model being up to 31.83% less expensive than the conventional ISTR models. To

TABLE 3. The mass and cost of the models

Model	ISTR [11]	Optimized ISTR [11]	Bar-PM	Optimized Bar-PM
Mass and Cost				
Magnet weight [kg]	0.64	0.41	0.47	0.41
Steel weight [kg]	2.27	2.6	2.42	2.6
Copper weight [kg]	0.79	0.83	0.79	0.83
Total weight [kg]	3.7	3.84	3.68	3.84
Magnet price [\$]	74.31	47.6	54.57	47.6
Steel price [\$]	1.8	2.06	1.93	2.06
Copper price [\$]	6.01	6.32	6.01	6.32
Total price [\$]	82.12	55.98	62.51	55.98

ensure a fair comparison, the consumables amount and cost of the optimized ISTR model were also optimized based on the optimized Bar-PM model.

5. 2. Comparison of the Performance Indices

The operational characteristics have been compared between the conventional ISTR, optimized ISTR, and the proposed and optimized Bar-PM models. The comparison results at 1500 rpm are listed in Table 4. The torque density of the optimized Bar-PM model is higher by as much as 14.88% and 5.04% compared to the conventional ISTR and proposed Bar-PM, respectively. While the torque density of the optimized ISTR model is lower by as much as 7.71% compared to the optimized Bar-PM model.

The optimized Bar-PM model, which has a torque-to-weight ratio of 1.75, has been enhanced by 9.71%, and 0.57% compared to the conventional ISTR and proposed model, respectively. Thus, the optimized model highlights the torque-to-cost ratio of 0.12 has been improved by as much as 69.01% and 17.64% compared to the conventional ISTR; and proposed Bar-PM, respectively. In the optimized Bar-PM model, the PM has been reduced by 35.93%, and the power-to-weight ratio of 0.089 has grown by 709.09% compared to the conventional ISTR model. Despite having a better torque-to-cost ratio than the ISTR and proposed bar-PM models, the optimized ISTR model still has issues with higher cogging torque and torque ripple. This affects the performance of the DSAFFSPM motor design, but the Bar-PM model has been developed to address these challenges.

Nonetheless, it has been revealed that the cogging torque value is 0.266 Nm, which is much lower compared to the conventional ISTR and proposed models, which are 90.08% and 75.23%, respectively. In addition, the torque ripple status for the optimized model is as much as 4.15% compared to conventional ISTR, and the proposed models decreased as much as 90.84%, and 70.44%, respectively. It has been declined that the cogging torque is as much as 90.32% compared to the optimized ISTR. While the torque ripple, it is suppression by as much as 89.79% compared to the optimized ISTR model. Machine efficiency is as follows:

$$\% \eta = \frac{T_{avg} \omega_r}{T_{avg} \omega_r + P_{cu} + P_{core}} \times 100\% \tag{26}$$

The losses are one of the main reasons for the low efficiency of the conventional ISTR model of much as 1.3% compared to the optimized model. in TABLE 4 represents the efficiency of the models as 92.96%, 93.48%, 94.13%, and 94.19%, respectively, regarding the efficient structure.

Comparing this section, it can be demonstrated that the proposed model with increased rotor pole pitch, and

TABLE 4. Performance indices comparison of the DSAFFSPM models

Parameter	Model	ISTR [11]	Optimized ISTR [11]	Bar-PM	Optimized Bar-PM
Rated Speed [rpm]				1500	
Current Density [A/mm ²]				7.5	
Cogging torque [Nm]	2.52	2.75	1.02	0.266	
Average torque [Nm]	5.87	6.22	6.41	6.74	
Torque ripple [%]	45.31	40.67	14.04	4.15	
Cogging torque/ Mean torque [%]	42.93	44.21	15.91	3.94	
Power [kW]	0.922	0.976	1.006	1.058	
Power density [kW/kg]	0.248	0.254	0.271	0.275	
Power/Cost [kW/\$]	0.011	0.0174	0.016	0.089	
Torque density [Nm/L]	5.98	6.34	6.54	6.87	
Torque/Total weight [Nm/kg]	1.58	1.61	1.74	1.75	
Torque/ Total cost [Nm/\$]	0.071	0.11	0.102	0.12	
Torque/PM cost [Nm/\$]	0.078	0.13	0.11	0.14	
PM Weight [% of Total weight]	17.29	10.67	12.77	10.67	
Total losses [W]	69.82	68.45	62.63	65.21	
Efficiency [%]	92.96	93.48	94.13	94.19	

reduced PM length technique had a superior performance in terms of high-torque density which was improved by 9.36% compared to that of the conventional. Although increasing the rotor pole pitch reduces flux concentration, but locating the PMs strategically generates higher torque with fewer PMs. To compensate, the rotor pole pitch width is slightly increased. While the conventional model has an overall cost of approximately \$82.12, its proposed and optimized models reduce costs by 23.87% and 31.83% respectively. Therefore, it should be noted that the optimized model is with the lowest torque ripples and cogging torque. Although the target of this paper is on high-torque density and low-cost, suppression of the torque ripples and cogging torque are important objectives, as neglecting them can deteriorate the performance of the motor. Therefore, implementing the Bar-PM technique on the conventional ISTR of the optimized model can save rare-PM consumption and costs, in addition to improving performance. Finally, the proposed model satisfies approximately all the required objectives of EVs.

5. 3. The CPSR and Efficiency Map of the Model

The key parameters of designing AFFSBPM motors, high-torque density, high-efficiency, low-cost, torque ripple, and thermal stress are investigated in detail for EV applications. This section emphasizes the constant power speed range (CPSR), the losses, and the efficiency map. Considering the previous part, the three models are compared in the same condition; indices' results mean that the optimized Bar-PM model is highlighted because of the significant operational characteristics.

3D-FEM results of torque-versus-speed and power-versus-speed are plotted in Figure 13. The EV applications have required a reasonable CPSR of up to 3-4 times the base speed. Although the standard AFPM machine has as well as shorter CPSR of about 2 times the rated speed [23]. Using concentric-coils in the AFFSBPM design is extend the CPSR, so that a proper CPSR of about 3 times the rated speed is achieved. The optimized motor falls within the CPSR standard range due to its high output torque at speeds higher than the rated speed. In the optimized model, the iron loss and copper losses are raised by an increment of speed, and the total loss map is shown in Figure 14(a). However, it can be claimed that the efficiency has risen by an increment in the speed of a wide CPSR due to its approximately high-power density in Figure 14(b).

5. 4. Thermal Analysis The DSAFFSPM machines split active sources between two stators, resulting in

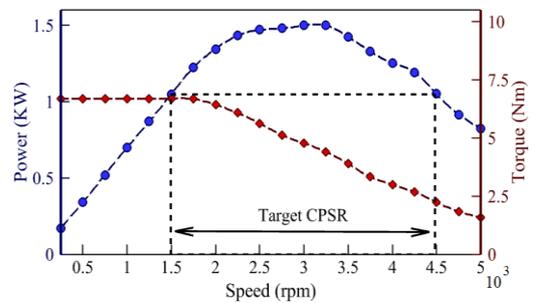
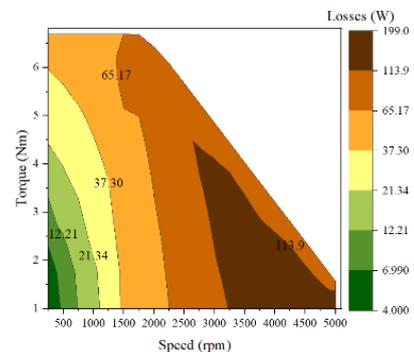
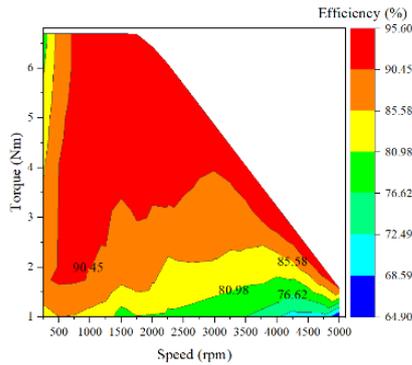


Figure 13. The torque and power curves vs. speeds



(a) The losses map



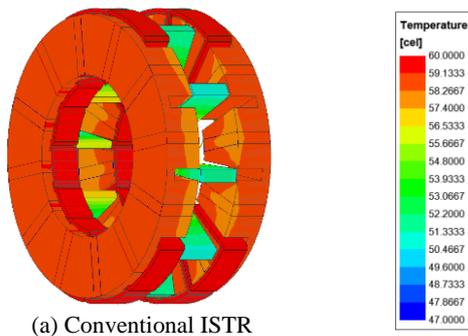
(b) The efficiency map

Figure 14. The efficiency map and torque-speed curves of the optimized Bar-PM model

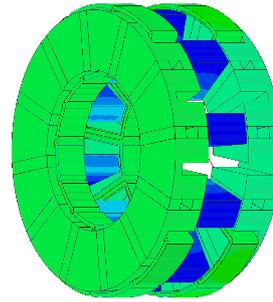
lower temperatures than other PM machines. New design, with air channel between PMs, expected to perform better than conventional ISTR. The thermal analysis of the DSAFFSPM models under rated current and @1500 rotor speed is predicted by FEM. Figure 15 can be seen the temperature distribution in the DSAFFSPM models. In this condition, the calculated temperature decreases at the optimized Bar-PM compared with the conventional ISTR model. In the winding, high temperature has been caused by the high thermal conductivity of copper and the electric current which flows through the winding. The temperature of the winding in the optimized model is 56°C, which is 15.15% lower than the conventional ISTR.

The maximum temperature of the stator for the optimized Bar-PM model is 55.94°C, which decreased by 11.20% compared to conventional ISTR. The optimized Bar-PM model has the lowest inner/outer rotor temperature at 51.25°C up to 47.14°C, followed by the conventional ISTR at 60.93°C up to 52°C. Also, the temperature of the PM for the optimized Bar-PM model is 54.77°C compared to conventional ISTR model, which decreased by as much as 13.69%, while the PM temperature in the inner part of the conventional ISTR model reached 66°C.

The proposed motor topology is different from the conventional topology. In the flux-switching structure,



(a) Conventional ISTR



(b) Optimized Bar-PM

Figure 15. The thermal analysis in the DSAFFSPM models

all the active sources are in the stator. However, the proposed topology includes an air channel between the Bar-PMs and between the teeth of the segment-stator. This feature enables the temperature of the motor to be controlled and reduced by using the air channel.

6. CONCLUSION

This paper presents a technique that can be implemented on the PM of the motor with the aim of reducing the cost and increasing the torque density. The operational characteristics of the proposed motor have been improved through analytical design, sizing equations and multi-objective optimization. The main features of the motor are its high torque density and low cost due to the minimal use of PM in its structure. The cost coefficient for torque/PM is significantly higher in the optimized motor compared to the other one. The back EMF harmonic orders value for the optimized Bar-PM model is less than 2%. Therefore, the optimized model appears almost superior in the torque ripple and cogging torque indices. The proposed motor produces satisfactory power and torque densities and has flux-weakening capabilities for a CPSR up to 3 times the rated speed. The proposed model can be cost-effectively designed and successfully commercialized, making it suitable for use in EVs.

7. REFERENCES

1. Farrokh F, Vahedi A, Torkaman H, Banejad M, Faradonbeh VZ. A 2D hybrid analytical electromagnetic model of the dual-stator axial-field flux-switching permanent magnet motor. *IET Electric Power Applications*. 2024;18(2):252-64. 10.1049/elp2.12385
2. Nasiri-Zarandi R, Mohammadi Ajamloo A, Abbaszadeh K. Cogging torque minimization in transverse flux permanent magnet generators using two-step axial permanent magnet segmentation for direct drive wind turbine application. *International Journal of Engineering, Transactions A: Basics*. 2021;34(4):908-18. 10.5829/ije.2021.34.04a.17
3. Patel A, Suthar B. Cogging torque reduction of sandwiched stator axial flux permanent magnet brushless dc motor using magnet notching technique. *International Journal of Engineering, Transactions A: Basics*. 2019;32(7):940-6. 10.5829/ije.2019.32.07a.06

4. Ebrahimi H, Torkaman H, Javadi H. Simultaneous improvement of cogging torque and torque density in axial flux-switching permanent magnet motor. *IET Electric Power Applications*. 2023. 10.1049/elp2.12390
5. Patel A, Suthar B. Design optimization of axial flux surface mounted permanent magnet brushless dc motor for electrical vehicle based on genetic algorithm. *International Journal of Engineering, Transactions A: Basics*. 2018;31(7):1050-6. 10.5829/ije.2018.31.07a.07
6. Kazemain M, Ebrahimi-Nejad S, Jaafarian M. Experimental investigation of energy consumption and performance of reverse osmosis desalination using design of experiments method. *International Journal of Engineering, Transactions A: Basics*. 2018;31(1):79-87. 10.5829/ije.2018.31.01a.12
7. Mohammadi Ajamloo A, Ghaheri A, Shirzad H, Afjei E. Non-linear analytical modelling and optimisation of a 12/8 rotor excited flux-switching machine. *IET Electric Power Applications*. 2020;14(9):1592-603. 10.1049/iet-epa.2019.0732
8. Farahzadi M, Abbaszadeh K, Mirnikjoo S. Electromagnetic-Thermal Analysis of a Hybrid-Excited Flux Switching Permanent Magnet Generator for Wind Turbine Application. *IEEE Transactions on Energy Conversion*. 2023. 10.1109/TEC.2023.3269038
9. Ghaheri A, Afjei E, Torkaman H. A novel axial air-gap transverse flux switching PM generator: Design, simulation and prototyping. *IET Electric Power Applications*. 2023;17(4):452-63. 10.1049/elp2.12277
10. Zarghani A, Torkaman H, Arbab N, Toulabi MS. Lumped parameter thermal network for thermal analysis of a rotor-excited axial flux switching machine with electromagnetic-thermal design. *Measurement*. 2022;193:110971. 10.1016/j.measurement.2022.110971
11. Zhao W, Lipo TA, Kwon B-I. A novel dual-rotor, axial field, fault-tolerant flux-switching permanent magnet machine with high-torque performance. *IEEE Transactions on Magnetics*. 2015;51(11):1-4. 10.1109/TMAG.2015.2445926
12. Torkaman H, Ghaheri A, Keyhani A. Design of rotor excited axial flux-switching permanent magnet machine. *IEEE Transactions on Energy Conversion*. 2018;33(3):1175-83. 10.1109/TEC.2018.2807804
13. Zhang W, Liang X, Lin M. Analysis and comparison of axial field flux-switching permanent magnet machines with three different stator cores. *IEEE Transactions on Applied Superconductivity*. 2016;26(7):1-6. 10.1109/TASC.2016.2595588
14. Kim JH, Li Y, Sarlioglu B. Novel six-slot four-pole axial flux-switching permanent magnet machine for electric vehicle. *IEEE Transactions on Transportation Electrification*. 2016;3(1):108-17. 10.1109/TTE.2016.2620169
15. Farrokh F, Vahedi A, Torkaman H, Banejad M. Design and comparison of dual-stator axial-field flux-switching permanent magnet motors for electric vehicle application. *IET Electrical Systems in Transportation*. 2023;13(2):e12074. 10.1049/els2.12074
16. Yıldırım E, Güleç M, Aydın M. An innovative dual-rotor axial-gap flux-switching permanent-magnet machine topology with hybrid excitation. *IEEE Transactions on Magnetics*. 2018;54(11):1-5. 10.1109/TMAG.2018.2848878
17. Zhao J, Quan X, Lin M. Model predictive torque control of a hybrid excited axial field flux-switching permanent magnet machine. *IEEE Access*. 2020;8:33703-12. 10.1109/ACCESS.2020.2973360
18. Zhang W, Liang X, Yu F. Fault-tolerant control of hybrid excitation axial field flux-switching permanent magnet machines. *IEEE Transactions on Magnetics*. 2018;54(11):1-5. 10.1109/TMAG.2018.2827948
19. Han J, Zhang Z. Design and optimization of a low-cost hybrid-pole rotor for spoke-type permanent magnet machine. *IEEE Transactions on Magnetics*. 2021;58(2):1-5. 10.1109/TMAG.2021.3090655
20. Ullah W, Khan F, Hussain S, Alturise F. Investigation of low-cost dual port co-rotating dual rotor field excited flux switching generator for low-power rooftop wind turbine integrated with multi-port DC micro grid. *IET Renewable Power Generation*. 2022;16(14):3108-18. 10.1049/rpg2.12561
21. Muhammad A, Khan F, Ullah B, Milyani AH, Azhari AA. Design and FEM analysis of high-power density C-core permanent magnet transverse flux generator with reduced PM volume. *IET Renewable Power Generation*. 2023;17(4):885-93. 10.1049/rpg2.12642
22. Lin M, Hao L, Li X, Zhao X, Zhu Z. A novel axial field flux-switching permanent magnet wind power generator. *IEEE Transactions on Magnetics*. 2011;47(10):4457-60. 10.1109/TMAG.2011.2157908
23. Capponi FG, De Donato G, Caricchi F. Recent advances in axial-flux permanent-magnet machine technology. *IEEE Transactions on Industry Applications*. 2012;48(6):2190-205. 10.1109/TIA.2012.2226854
24. Arehpanahi M, Kheiry E. A New Optimization of Segmented Interior Permanent Magnet Synchronous Motor Based on Increasing Flux Weakening Range and Output Torque (Research Note). *International Journal of Engineering*. 2020;33(6):1122-7. 10.5829/ije.2020.33.06c.09
25. Chen Y, Zhuang J, Ding Y, Li X. Optimal design and performance analysis of double stator multi-excitation flux-switching machine. *IEEE Transactions on Applied Superconductivity*. 2019;29(2):1-5. 10.1109/TASC.2019.2891899

COPYRIGHTS

©2024 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 این مقاله، یک موتور شار میدان محوری شار سوئیچینگ مغناطیس دائم دو استاتور (DSAFFSPM) برای بهبود چگالی گشتاور و هزینه دستگاه پیشنهاد شده است. در این ساختار، استاتور دوگانه ۱۲ قطبی در دو طرف یک روتور ۱۰ قطبی دندانه داخلی قرار گرفته است. دو استاتور میزبان آهنربای مغناطیس دائم (PM) نوع Bar-PM و سیم پیچ‌ها است. نکته جدید این مطالعه ارائه تکنیکی است که می‌تواند بر روی آهنرباهای ساختار DSAFFSPM اجرا شود. در این راستا، طرح تحلیلی پیشنهادی با یک معادله اندازه ارائه شده است و از بهینه‌سازی چند-هدفه برای دستیابی به اندازه بهینه با روش الگوریتم ژنتیک چند هدفه (MOGA) استفاده شده است. ویژگی‌های الکترومغناطیسی ماشین با استفاده از روش المان محدود سه-بعدی (3D-FEM) به دست آمده و تجزیه و تحلیل می‌شوند. مطالعه تطبیقی برای اثبات برتری شاخص‌های عملکرد انجام شده است. ساختار پیشنهادی به دلیل ریبیل گشتاور و گشتاور دندانه‌ای کم دارای چگالی توان بالا، ارتعاش و نویز کم‌تری می‌باشد. در همین حال، ساختار پیشنهادی تنش حرارتی کم‌تری به دلیل راندمان بالا و تلفات هسته کم دارد. در نتیجه، ساختار پیشنهادی چگالی گشتاور بالا و هزینه کم را ارائه می‌کند که به طور خاص برای کاربردهای خودروهای الکتریکی (EVs) طراحی شده است.



Topology and Thickness Optimization of Concrete Thin Shell Structures Based on Weight, Deflection, and Strain Energy

M. Vatandoost^a, M. Golabchi^a, A. Ekhlasi^{*b}, M. Rahbar^b

^a Department of Architectural Technology, Faculty of Architecture, College of Fine Arts, University of Tehran, Tehran, Iran

^b School of Architecture and Environmental Design, Iran University of Science and Technology, Tehran, Iran

PAPER INFO

Paper history:

Received 30 October 2023

Received in revised form 21 December 2023

Accepted 30 December 2023

Keywords:

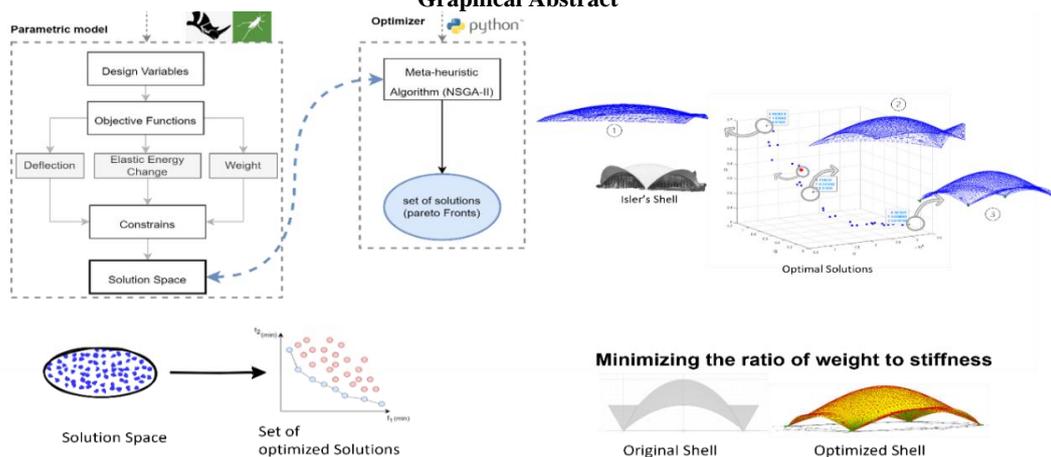
Concrete Shell Structures
Structural Optimization
Shape Optimization
Topology Optimization
Multi-objective Optimization
Non-dominated Sorting

ABSTRACT

This study presents the optimized shape and thickness of thin continuous concrete shell structures, minimizing their weight, deflection, and elastic energy change while meeting the performance requirements and minimizing material usage. Unlike previous studies that focused on single-objective optimization, this research focuses on multi-objective optimization (MOO) by considering three objective functions. This combination of objective functions has not been reflected in previous research, distinguishing this study. The computational design workflow incorporates a parametric model, multiple components for measuring objective functions in the grasshopper of Rhino, and a metaheuristic algorithm, the non-dominated sorting multi-objective genetic algorithm (NSGA-II), as the search tool, which was coded in Python. This workflow allows us to perform form-finding and optimization simultaneously. To demonstrate the effectiveness of this metaheuristic algorithm in structural optimization, we applied it in a case study of a well-known shell designed using the physical prototyping hanging model technique. Interpretations of samples of optimized results indicate that although solution 1 weighs nearly the same as solution 2, it has less deflection and strain energy. Solution 3, with a three-fold mass, has significantly less deflection and strain energy than solution 1 and solution 2, with deflection reductions of over 50 and 17%, respectively. Solutions 3 and 4 show better deflection and strain energy performance. Furthermore, a comparison of the MOO results with the Isler shell revealed that this method found a solution with less weight and deflection while being stiffer, confirming its practicality. The study found that MOO is a reliable method for form-finding and optimization, generating accurate and reasonable results.

doi: 10.5829/ije.2024.37.07a.15

Graphical Abstract



*Corresponding Author Email: ekhlasi@iust.ac.ir (A. Ekhlasi)

Please cite this article as: M. Vatandoost, M. Golabchi, A. Ekhlasi, M. Rahbar. Topology and Thickness Optimization of Concrete Thin Shell Structures Based on Weight, Deflection, and Strain Energy. International Journal of Engineering, Transactions A: Basics. 2024;37(07):1369-83.

1. INTRODUCTION

1. 1. Problem Statement

On the one hand, concrete's ubiquitous availability, high strength, durability, and versatility make it an essential and irreplaceable building material, and due to its wide accessibility of necessary raw materials; it is the most often utilized artificially made material (1-3). In addition, concrete shell structures are among the most efficient structural systems (4, 5). On the other hand, Carbon emissions and global warming threaten our environment, according to reports of World Green Building Council (6), and Global status report for buildings and construction (7), and as indicated by Zhong et al. (8) the building sector is responsible for more than 35% of global energy consumption and more than 40% of carbon emissions (9, 10). In response, architects and researchers constantly revise their ways of designing and building structures. If we find ways to design and build our structures more efficiently, less material will be used, and we can reduce and control the environmental impact (11-13).

Since the emergence of computer-aided design in architectural fields, the appearance of complex free forms has materialized. Although computational design has paved the way for the design of complex forms, these structures are required to be designed in their optimum state. Historically, shell structures were form-founded and designed by physical prototyping and hanging models. Conversely, the optimal design for complex free-form thin shell structures requires advanced methods. Structural optimization plays a crucial role in achieving optimal design solutions that meet performance requirements while minimizing material usage and overall weight. In recent years, metaheuristic algorithms have gained prominence as effective tools for solving complex optimization problems. This paper focuses on the application of a metaheuristic algorithm, specifically the non-dominated sorting multi-objective, multi-criteria genetic algorithm (MOGA), in the context of structural optimization. Hence, in this study, we provide shape and thickness optimization of thin concrete shell structures by employing metaheuristic algorithms and by minimizing the shell structure's weight, deflection, and elastic energy change.

1. 2. Related Studies

Historically, shell structures had been designed through physical prototyping, including hanging chains and hanging models (5, 14). Through the introduction of computational methods and advances in Computer-Aided Design (CAD) methods, researchers have developed advanced design methods; one such advanced method is utilizing metaheuristic algorithms in the design process. Metaheuristic algorithms are able to find

accurate solutions for complex design problems (15, 16). A vast and diverse solution can be generated using metaheuristic algorithms, known as 'solution space.' These algorithms are able to find the best (fittest) solution(s) considering the defined criteria without evaluating each solution in the solution space (17, 18). Based on knowledge of the domain and simulation of Natural or physical laws, these algorithms can find an accurate answer to a complex problem.

Shell structure optimization is reflected in several previous research that emphasizes this topic's importance. We have classified previous studies based on the objective function of the optimization in two major classes; the objective functions of the related studies are summarized in Table 1.

Pugnale et al. (19) utilized a genetic algorithm to optimize a shell structure, while the objective function was deflection in the Z direction. Additionally, Santhosh et al. (20) optimized a Gridshell structure by genetic algorithm (GA), while objective functions were shell vertical displacement, and they studied different geometry patterns for grids. Besides, the total weight of the structure was considered the optimization criteria in research by Basso et al. (21). A comparable study with the same objective function but a different optimization technique was conducted by Baghdadi et al. (22). Moreover, in the study by Ansola et al. (23) the strain energy was considered as the optimization criteria. Furthermore, In a study by Kimura and Ohmori (24), the topology and thickness distribution of concrete shells were studied by considering strain energy as the objective function. Also, Yang et al. (25) utilized a particle swarm optimization algorithm to minimize the strain energy in a free-form shell. In addition, Tomas and Marti (26) optimized the design of a shell by considering the total structure weight as a single objective optimization, repeated that for stress level, and studied it based on strain energy. Additionally, strain energy change in a Gridshell was considered by Feng and Ge (27) as the objective for optimization. Moreover, in the study by Hassani et al. (28), the objective was structure weight. The total weight was minimized in a similar study by Richardson et al. (29). In addition, maximization of the stiffness was the optimization criteria in the study by Shimoda and Liu (30). Besides, Kaveh et al. (31) provided cross-section optimization for Grid shells, and Tomei et al. (32) developed the method for grid shells constructed by pre-stressed rods. Another example is the research by Ansola et al. (33) which optimize thickness and topology of the shell structures by considering displacement as the objective. Furthermore, Gythiel and Schevenels (34) formulated the optimization of a single layer reticulated shell by minimizing the elements cross section (total mass) of the structure as the objective; In this study, size, shape and topology of the structure were

variables while considering multi-criteria in the design. Teimouri and Asgari (35) utilized BESO for topology optimization while maximizing stiffness of the structure. In one of the recent published research, Zhang et al. (36) proposed collapse-resistant optimization for a single-layer grid shell.

Emami et al. (37) optimized a perforated concrete shell concerning daylight and energy performance; research with similar objectives was conducted by Liuti et al. (38) and Emami (39), where the objectives were structure and natural light. In another study, Pugnale (40) considered structure and acoustic as the objective function for optimization. Turrin et al. (41) also provided optimization based on daylight and solar heat gain. Furthermore, In the study by Puppa and Trautz (42), optimization criteria were buckling load, strain energy, and sensitivity to imperfection. Optimization based on the strain energy and total weight of the members is reflected in a study by Nagata and Honma (43) and Wang et al. (44). Henriksson et al. (45) presented the multi-objective optimization considering total mass and

deflection. Moreover, in a recognized building, ‘NESTHiLo’ researchers from the block research group at the ETH Zurich (46, 47) designed and built a thin concrete shell as a roof and optimized it based on the area of glazing, the formwork deviation, and elastic energy change. In another study, Zhao et al. (48) found the optimal design for a Gridshell based on the length and cross-section of elements. Vargas et al. (49) employed The differential evolution (DE) algorithm, and considered penalty function in the workflow to handle constrains in the optimization of different structures, while the objective functions were weight and displacement. Mirra and Pugnale (50) employed a multi-objective genetic algorithm to minimize displacement and shell footprint from the target shape while maximizing the shell height. Nishei and Fujita (51) considered strain energy in combination with the collapse load factor. In Table 1, previous studies are classified based on the objective functions for optimizing the thin shell structures.

TABLE 1. Classifying research based on objective functions

	Objective function	Reference	Objective functions	Reference
Single objective optimization	deflection	Pugnale et al. (19), Gokul et al. (20) Ansola et al. (33)	f ₁ : daylight f ₂ : Energy performance	Emami et al. (37)
	strain energy	Kimura and Ohmori. (24) Feng and Ge (27), Yang et al. (25), Wang (52) Ansola et al. (23)	f ₁ : structure f ₂ : acoustic (SPL)	Pugnale (40)
	Material usage (Weight of the structure)	Basso et al. (21), Hassani et al. (28), Baghdadi et al. (22), Kostura et al. (53) Richardson et al. (29)	f ₁ : daylight f ₂ : Solar heat gain	Turrin et al. (41)
	Stiffness	Shimoda and Liu (30)	f ₁ : structure f ₂ : Natural light	Liuti et al. (38), Emami (39)
	stress level weight, strain energy, (separately)	Tomas and Marti (26)	f ₁ : strain Energy f ₂ : surface curvature	Jiang (54)
	Cross-section	Tomei et al. (32), Kaveh et al. (31) Gythiel and Schevenels (34)	f ₁ : strain energy f ₂ : Total weight of the members	Nagata and Honma (43), Wang et al. (44)
Multi-objective optimization			f ₁ : total mass f ₂ : deflection	Henriksson et al. (45)
			f ₁ : strain energy f ₂ : collapse load factor	Nishei and Fujita (51)
			f ₁ : elements length f ₂ : cross-section	Zhao et al. (48),

f ₁ : strain Energy	
f ₂ : Geometric index	Cao et al. (55)
f ₃ : Economic index	
f ₁ : displacement	
f ₂ : shell height	Mirra and Pugnale (50)
f ₃ : Footprint deviation	
f ₁ : Imperfection sensitivity	
f ₂ : buckling load	Pappu et al. (42)
f ₃ : strain energy	
f ₁ : Elastic energy change	
f ₂ : Glazing surface area	Veenendaal (46), Veenendaal et al. (47)
f ₃ : and formwork deviation	

1. 3. Research Objectives Based on provided literature review, most of these studies considered a single criterion for optimization, a single-objective optimization; however, in some research, two objectives were considered as the multi-objective optimization problem. Nevertheless, considering three or more objective functions to optimize the shell structures is rare. The goal of this research and our contribution is summarized below.

- Finding the optimal design solution for the continuous concrete shell based on structural performance.
- In the design of concrete shell structures, despite most previous research mentioned in the literature review, which only considered one objective, single-objective optimization, or two objectives for the optimization, in this research, we intended to perform multi-objective optimization by considering three objective functions.
- This combination of the three considered objective functions has not been reflected in previous research, distinguishing this study.
- Multi-objective optimization by considering three objective functions: the total weight of the structure, the maximum deflection, and elastic energy change were competing toward the final optimal solution.

Apply the optimization workflow in a case study to demonstrate the effectiveness of the metaheuristic algorithm (NSGA II), and find the optimal solutions compared to the hanging model simulation.

2. METHOD

To conduct a multi-objective optimization by employing metaheuristic algorithms, we need to integrate three parts: a parametric model, functions to measure objectives, and a solver (56, 57). In this study, the parametric model is created in the grasshopper of the

Rhino 3D. By altering variables, this parametric model generates many candidate solutions known as ‘solution space.’ Moreover, multi-components are added to the model to measure and record each solution rank related to the considered objective functions. In this study, the Karamba 3D plug-in (58) is used to calculate the structural performance of the parametric shell. Lastly, a solver, the MOGA (the metaheuristic algorithm), is employed in the workflow by a series of custom ‘Python’ programming language codes, which can generate multiple candidate solutions and rank them, and through mutation, crossover and simulation of ‘natural selection’ law, find the best (fittest) solution(s) (59). The concept of utilizing a metaheuristic algorithm as a search method is depicted in Figure 1. Also, the flowchart of the multi-objective genetic algorithm (the non-dominated sorting genetic algorithm, NSGA-II) is depicted in Figure 1.

2. 1. Optimization Formulation Optimization in engineering is formulated by Mirjalili and Dong (60), Rao (61) as follows:

$$Find X = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{Bmatrix}, \text{ which minimizes } F(x), \tag{1}$$

X is an array of variables, $X = \{x_1, x_2, \dots, x_m\}^T$, and in the multi-objective optimization, $F(X)$ is the combination of multiple functions which we intended to minimize simultaneously; therefore, we have:

$$F(X) = \{f_1, f_2, \dots, f_n\} \tag{2}$$

while satisfying inequality and equality constraints:

$$g_r(x) \leq 0 \tag{3}$$

$$h_r(x) = 0 \text{ r} = 1 \text{ to k} \tag{4}$$

Figure 2 depicts the concept of utilizing metaheuristic algorithms as a search tool.

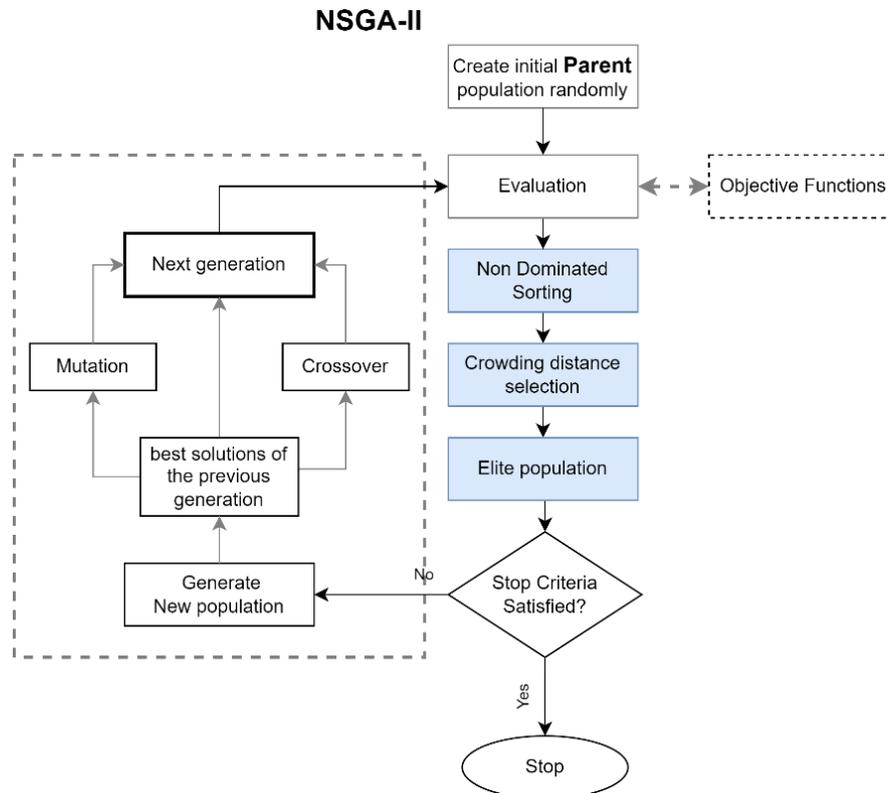


Figure 1. The flowchart of the MOGA-NSGA-II algorithm

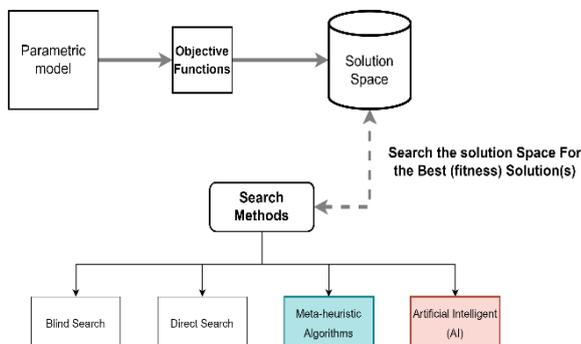


Figure 2. Utilizing metaheuristic algorithms as the search method to find the fittest solution

In multi-objective optimization, we intended to minimize or maximize competing objective functions and find a combination of variables that minimize this objective function, which is a complex and multifaceted problem (62). This is because a set of variables that minimize or maximize one objective function (for instance, f_1) might result in an entirely low-rank solution while considering another objective (for instance, f_2, f_3) (49).

2. 2. Objective Functions

In this study, the

structure's total weight, the shell's maximum deflection, and the elastic energy change are three competing objectives employed as the main design criteria, summarized in Table 2.

The total weight of the structure is related to the shell surface and thickness, and the deflection is defined as the maximum displacement of the shell in the Z-direction, while the shell is under the load combination of the self-weight and a 1KN equal distribution of the live load. The Strain Energy (Elastic energy change) of the structure is defined as the potential energy that is accumulated within a structural element due to its elastic deformation (63). The strain energy formula is given below:

$$U = \frac{\sigma^2}{2E \times V} \tag{5}$$

where, σ = stress, E = Young's modulus, V = volume of the body (Structure).

Therefore, by measuring Strain Energy, we can compare different structures; the structure with the lowest elastic energy change is stiffer.

These three objectives are competing toward the best solution. At the same time, a shell with a more significant raise-to-span ratio will be stiffer with less deflection and less elastic energy but will have more surface, and more material will be used and, therefore, will weigh more.

TABLE 2. Three objective functions of this study

	Objectives	Unit
f_1	Total weight	Kg
f_2	Deflection	mm
f_3	Strain Energy (Elastic energy change)	KNm

2. 3. MOGA, NSGA-II

Multi-objective genetic algorithm, MOGA, is one of the fast and reliable metaheuristic algorithms (64); that has been inspired by nature. The GA is based on the theory of 'natural selection' of Charles Darwin, and Goldberg formalized it to be used in engineering optimization (65). The Non-dominated sorting genetic algorithm (NSGA-II) is the subset of GA, which enables us to find a set of optimal solutions known as the 'Pareto Frontier' by mutation, crossover, and selection functions, based on the defined objective function (criteria).

2. 3. 1. Non-domination Sorting

Non-dominated sorting is a technique that selects and stores the solutions not dominated by other solutions in the solution space. Domination is defined as Equations 5 and 6 (66).

Solution 1 dominates Solution 2 if and only if, for all the objective functions (i), the value of the considered objective function for solution 1 is equal to or less than solution 2, or there is a solution in which the value of the considered objective function for that solution is equal to or less than solution 2. Additionally, to select diverse solutions from the Pareto front, a fitness value named 'crowding distance' is defined and assigned to each solution in the Pareto Frontier, which relates the solution density to ranking, which lets the diverse solution be chosen. The solution with the heist crowding distance means the solutions far from each other will be chosen in the Pareto Frontier (67), depicted in Figure 3.

Solution1 dominates, solution 2 if and only if:

All	For each objective, (i)	$\forall_i : f_i^{solution1} \leq f_i^{solution2}$	(Universal quantification)	(6)
Any	There is a solution that:	$\exists_i : f_i^{solution1} < f_i^{solution2}$	(Existential quantification)	(7)

Crowding distance is calculated by Equation 7. For boundary solutions, the solutions with the heists and lowest quantities, infinite distance ($d_s = \infty$), are assigned to make sure this is involved in the Pareto frontier (68).

$$d_s = \sum_{i=1}^l \frac{f_i(s+1) - f_i(s-1)}{f_i^{max} - f_i^{min}} \tag{8}$$

S = 1 to k

S: Number of solutions in the Pareto frontier

Therefore, NSGA-II, in each cycle, ranks and sorts the solutions in a set called the 'Pareto Frontier.' The solutions in the Pareto Frontier are the solutions that are superior to other solutions but have no advantages over each other, and every solution in the 'Pareto Frontier'

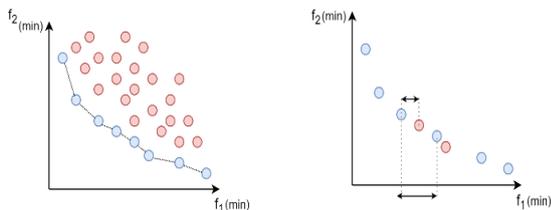


Figure 3. The non-dominated sorting arranges solutions in hierarchies (fronts); the solutions in the same Front are superior to other solutions while not dominated by each other (left). The Crowding distance fitness value enables to elimination of similar solutions and having a more diverse set of solutions in the Pareto front (right)

founded by NSGA-II can be considered as the final design solution (69).

2. 4. Unified Workflow

As described before, the workflow of MOO consists of three parts: a parametric model, objective functions calculator, and a metaheuristic algorithm. This integration workflow is shown in Figure 4.

2. 4. 1. Stopping Criteria

In contrast to the Exhaustive search method, which considers all the solutions known as the 'Brut force' method, the metaheuristic algorithms will find the fitness solution without measuring every solution, making it computationally faster. While utilizing metaheuristic algorithms, either it will converge to a satisfactory solution or infinite looping. To avoid infinite looping, a stopping criterion for the search operation must be defined. We have defined the stopping criteria when either of these four occurs (Figure 5): a) an acceptable number of iterations reached, b) algorithm convenient cycles, c) slow or no progress encountered, and d) a solution does not exist.

3. APPLICATION

To show the effectiveness of the MOGA algorithm in structural optimization, we have employed it in a case study. In this case study, we have chosen to optimize one

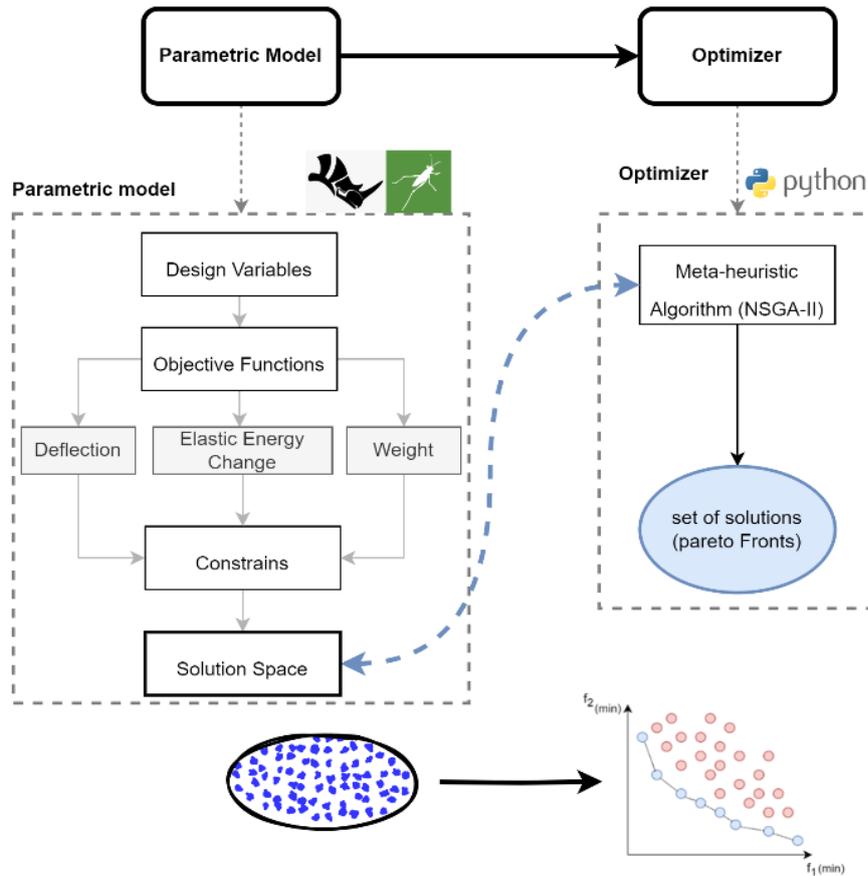


Figure 4. The integrated workflow of multi-objective optimization

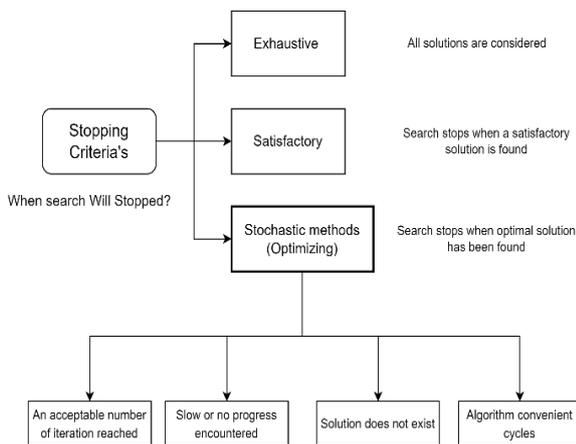


Figure 5. Stopping criteria in the operation of the metaheuristic algorithm

of the well-known previously built concrete thin shell structures. As described before, these shell structures were form-funded and designed by physical modeling and hanging models. We intended to optimize it based on the mentioned objective functions and compare the results.

3. 1. Intro to the Case Study The ‘Wyss Garden,’ designed and built by Isler in 1962, is selected as the case study to be optimized. The details of the base model are extracted from literature (70) (see Figure 6). According to Chilton and Isler (2), this was the first Isler’s ‘free-form’ shell with 650 m² of surfaces generated by circular curves. At that time, CAD modeling and optimization techniques were unavailable, and this shell was designed based on the hanging model. Therefore, we will compare the optimized shell results with Isler’s built shell.

Throughout the computational design process, various potential solutions were developed. For each of these solutions, the objective function(s) are calculated, and the resulting data for each solution should be saved. Due to the necessity for repetition across various potential solutions. This procedure is time-consuming and computationally costly. Therefore, it is crucial to establish a parametric model that is both simple and accurate in the first stages of the design (71).

This study employs networks of Non-uniform rational basis splines (NURBS), a widely used method in the field for representing surfaces and curves, to mathematically represent and model shells (72). It provides exceptional

adaptability and accuracy when it comes to manipulating modeled and analytical shapes that are defined by standard mathematical formulas.

By considering Isler's 'Wyss Garden' as the base, we have defined the parametric model with a set of eight boundary NURB curves, based on Isler's 'Wyss Garden'. By adjusting the position of the control points on these curves, we can create different topologies. This modification allows for 16 variables to be manipulated, resulting in various shell shapes. Figure 7 illustrates the specific information. Moreover, Table 3 establishes the limits for the variables. In this case study, there are eight control points. The problem variables are the x or y positions of these points, together with the z position that determines the shell topology. In Figure 7, the position of 'n1' can range from zero to 5.00 meters in the y-direction and from zero to 10.00 meters in the Z-direction.

Variables

To have a parametric model and be able to generate

NSGA-II hyper parameters

The algorithm hyper-parameters are summarized in Table 4.

The main material for the shell is lightweight concrete C50/60, and the material properties are summarized in Table 5. Concrete Design Properties is provided according to EN1992-1-1 ($\gamma_c = 1.50$, $f_{yk} = 500$ MPa)¹.

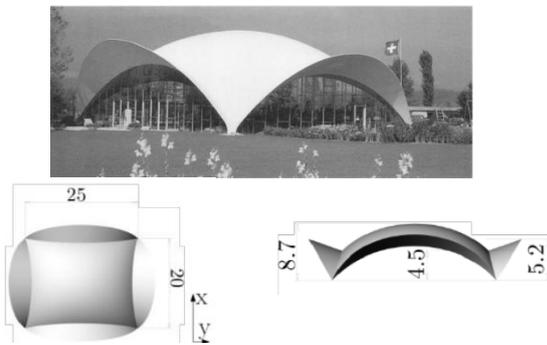


Figure 6. The 'Wyss Garden' by Isler, source: (2, 70, 73)

TABLE 3. Variable bounds

Variable	Lower bound (m)	Upper Bound (m)
x_0, x_2, x_4, x_6	0.	5.00
x_1, x_3, x_5, x_7	-7.00	0.
z_0, z_1, \dots, z_7	0.	10.00
t	0.08	0.20

¹ f_{yk} : Steel characteristic yield strength; γ_c : Concrete partial material safety factor

TABLE 4. NSGA-II hyperparameters in the case study

NSGA-II hyper parameters	
Number of populations	50
Number of generations	10
Number of Genes	4
Mutation rate %	0.2
Crossover rate %	0.8
Number of Objectives	3

TABLE 5. Lightweight Concrete, Material Properties

Compressive strength (f_c)	50 MPa (~500 kg/cm ²)
Tensile strength	4.07 MPa
Modulus of Elasticity (E)	37278 MPa
Density	2400 kg/m ³ (150 lb/ft ³)
Poisson's ratio	0.2
Shear strength	6 MPa
Thickness (t)	8 - 20 cm

4. RESULTS AND DISCUSSIONS

Thin shell concrete structures can adequately withstand forces in compression but are sensitive to tension forces. The hanging model method is based on the form resulting by using fabric and other tensile materials under the self-weight with the defined boundary conditions (loading and supports); the fabric will fall in the equilibrium position in pure tension. By reversing the resulted form, we will have a structured act in pure compression (5). Historically, shell structures were designed based on hanging models. Thus, first, we generate the result by computer simulation of the hanging model using Kangaroo physics (74) in the grasshopper of Rhino to provide a general idea and a good guess about the acceptable result based on the structural performance. In addition, in our parametric model, by setting out the parameters from the built model, we provided the structural performance of the 'Wyss Garden' for Evaluation and compared the results of the Isler shell with the MOO results; these results are summarized in Figure 10, and Table 6.

The result by MOO is provided in Figure 8; the parallel plot of all considered solutions in the optimization search is depicted. Each vertical coordinate represents a variable (17 in this case study, refer to Table 3) and corresponding objective function values (three in this case study, refer to Table 2). In this case study, 498 solutions were considered

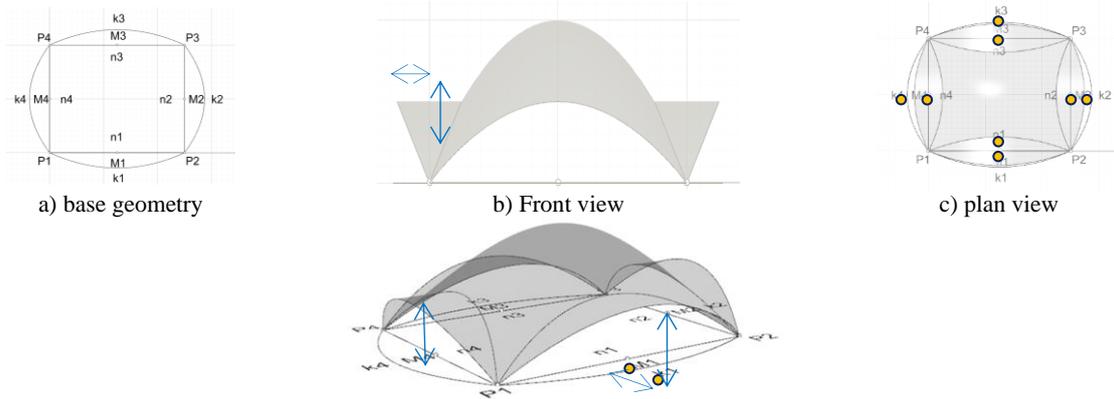


Figure 7. The defined parametric model in this Case study

to find the optimized solutions. The solutions in the Pareto Frontier that are non-dominated and have superiority over other solutions are shown in Figure 9. Each blue dot represents a candidate solution that can be chosen as the final solution.

In addition, in Figure 9, the objective functions were compared two by two; The solutions distribution in the

Pareto frontier indicates that f_1 (total weight) and f_2 (deflection) are competing toward the best solution when f_1 is reduced, f_2 has increased and vice versa. Similarly, f_2 (deflection) and f_3 (strain energy) have an equivalent relation and are in conflict toward finding fitness solutions. However, based on the provided charts, f_2 and f_3 have a linear relation; when f_2 reduces, f_3 reduces based

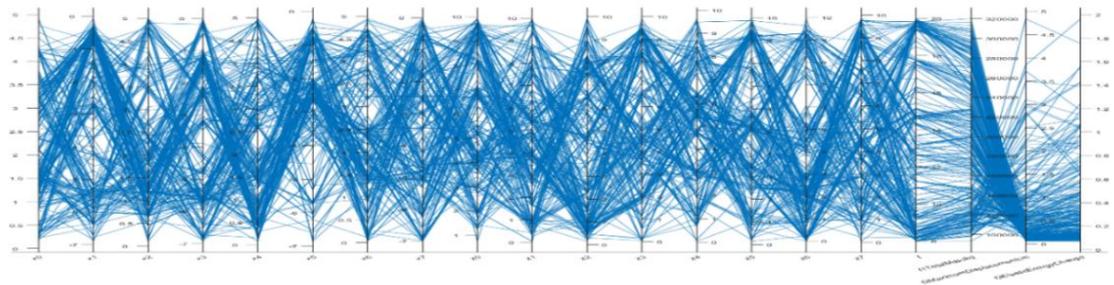


Figure 8. Parallel plot of the solution space; 498 solutions were considered in this case study

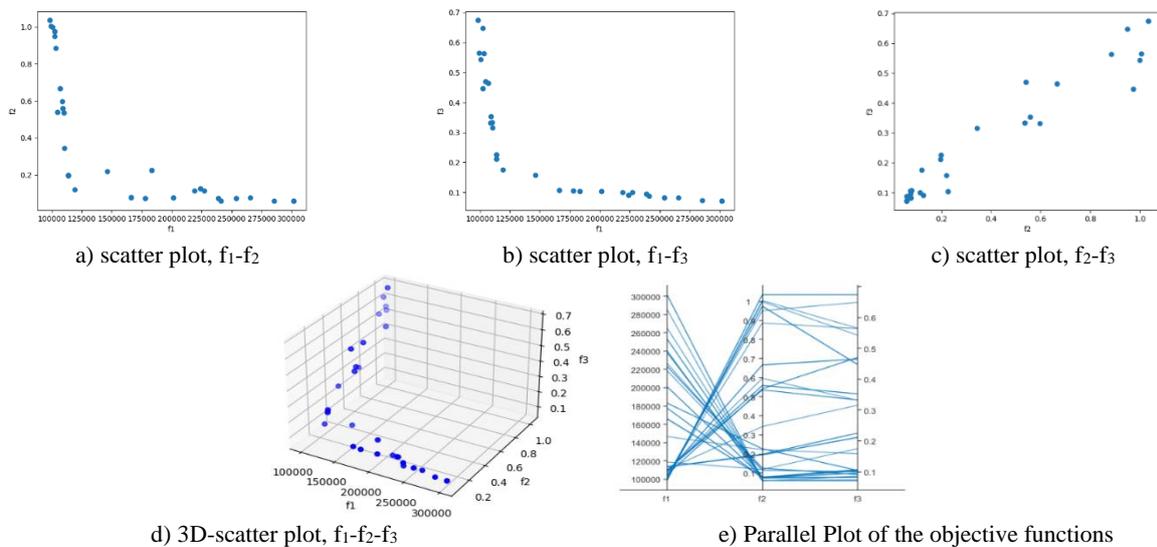


Figure 9. Pareto Frontier of the Case study I, f_1 : Total weight, f_2 : Deflection, f_3 : Elastic Energy Change

on similarly. In Addition, the Parallel plot is depicted in Figure 9 indicates how these three objectives are competing.

In a similar study conducted by Turrin et al. (41), they utilized genetic algorithm to find the optimal design solution based on structural performance. In their research, the design alternative could be generated by GA. In another study by Sassone and Pugnale (75), the maximum displacement was the objective function for structural. Despite the above mentioned related studies, in this research we have provided multi-objective optimization of the lighthweigh continious thin shell structures, by combining three objective function, which is scarce in the literature. The provided workflow, enables us to find diverse set of candidate solutions based on the defined criteria. Additionally, Veenendaal (46), Veenendaal et al. (47), in their built research project, which is a roof of a building, provided the Pareto Frontier of a thin shell based on three objective functions, Elastic energy change, glazing surface area and formwork deviation, which validates this method for finding the optimal design solution. The design alternative was chosen based on the head clearance and architectural requirments, and developed.

Furthermore, the results of this study which is depicted in Figures 10 and 11 and Table 6 consist of the resulting shell based on hanging model simulation and the Isler shell remodeling, along with five solutions from

the Pareto frontier as samples. Either of these five solutions or any other solutions from the Pareto Frontier (Figure 12) can be chosen by the Decision Maker (architect) as the final solution for further design development. However, a reliable scientific method for this selection is required, and it must be developed.

By comparing two samples of MOO's optimized results with the Isler shell, we can conclude: the total mass of solution 1 is less than Isler shell. At the same time, solution 2 weighs nearly the same as Isler, but the deflection and strain energy of solution 2 is reduced compared to Isler.

Furthermore, based on the provided results in Figure 10 and Table 6, although the total mass of the structure in the solution 3 is about three times the mass of the solutions 1 and 2, the deflection and strain energy of this solution is much less than these two candidate solutions; the deflection in comparison to solutions 1 and 2 is reduced more than 50 and 17%, respectively. Besides, solutions 3 and 4 have better deflection and strain energy performance than solutions 1 and 2.

Based on the provided results, and observing that there are some candidate solutions among the best candidate solution in Pareto Frontier that are similar to the Isler model, we can conclude that physical prototyping and hanging model methods are valid methods for form-finding of thin shell structures, and they can give the architecture a good starting point and

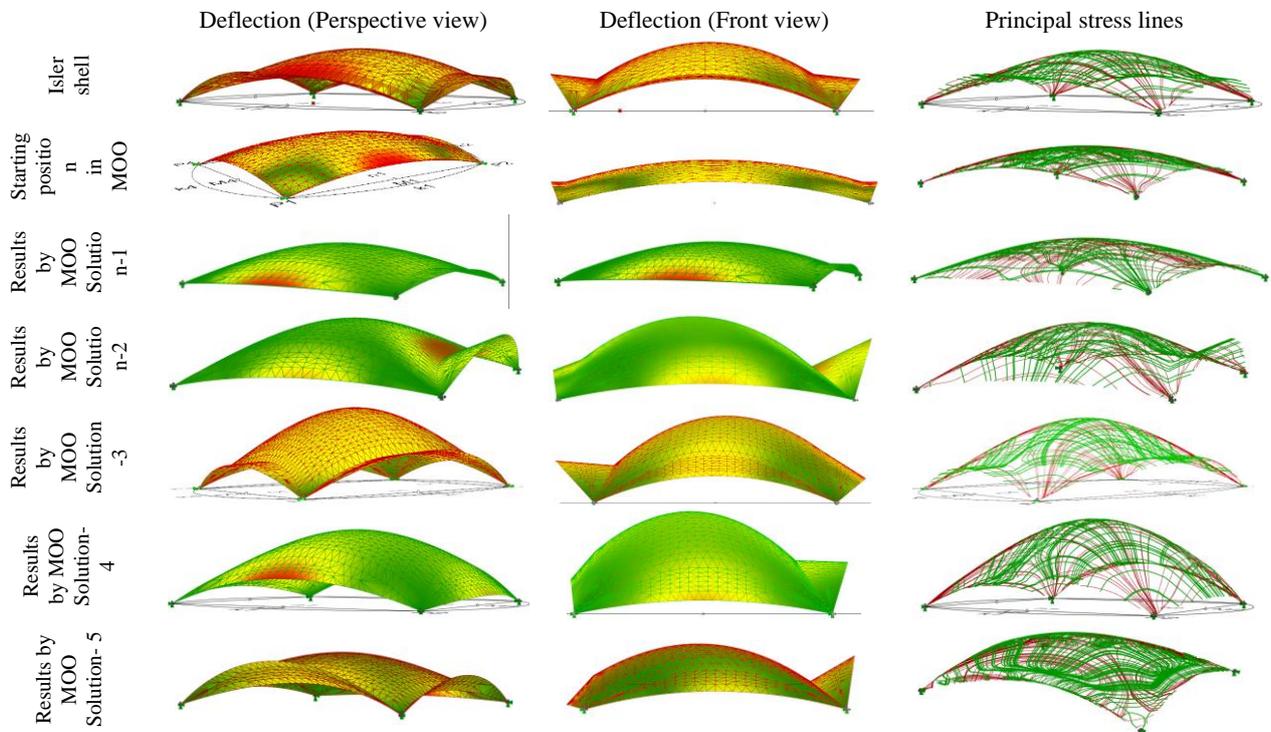


Figure 10. The results of the case study, selected sample of the optimized solutions are depicted



Figure 11. Hanging model simulation

TABLE 6. Comparison of the objective functions

Objective functions	Isler shell	Starting position	Optimized Solutions				
			Solution-1	Solution-2	Solution-3	Solution-4	Solution-5
t Thickness (cm)	8	8	8	8	19	17	8
f_1 Total mass (kg)	110022.8248	93214.154691	98080.815803	110570.452284	301837.474533	253604.394806	109828.6
f_2 Deflection (cm)	0.5357	0.711381	1.035417	0.344302	0.058893	0.074598	0.535816
f_3 Strain energy (kNm)	0.39055	0.643162	0.67408	0.31626	0.070759	0.082373	0.333121

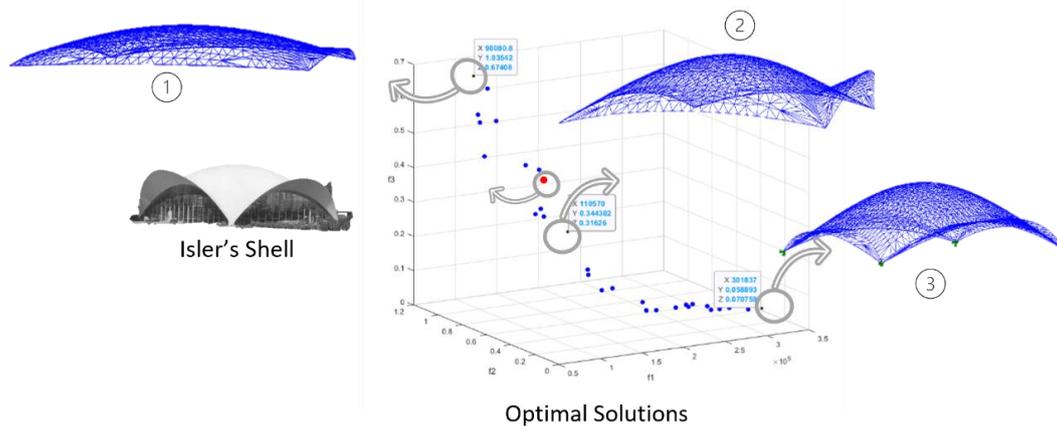


Figure 12. Optimized solution in Pareto Frontier (blue dots) and Isler model (red dot)

basic understanding of the load paths; however, these methods are not able to consider other criteria in the design, such as acoustic, energy, environmental and other criteria.

It is worth note that however, if we consider other criteria in design, such as the structure's acoustic, energy, and environmental performance, the result provided by the hanging method will be different from the fittest solution.

4. CONCLUSION

This study presented the optimal design solution based on structural performance by minimizing three objective functions. The topology and thickness of the thin, lightweight shell structure were optimized by minimizing the structure's total weight, deflection, and strain energy.

We have utilized the non-dominated sorting multi-

objective, multi-criteria genetic algorithm in the computational design of thin shell structures. In our workflow, we combined a parametric model, multiple components for measuring defined objectives, and the metaheuristic algorithm (NSGA-II) which was coded in Python. Utilizing metaheuristic algorithms in the design will provide diverse optimal solutions instead of one single solution. To show the effectiveness of this metaheuristic algorithm in structural optimization, we have employed it in a case study for topology and thickness optimization of a shell, which was form-founded based on the hanging model method. By utilizing MOGA, we were able to find a solution with less weight and less deflection while being stiffer, which confirms the practicality of this method. However, a comparison of the results of the Isler concrete thin shell as the base model and the optimized model reveals that since our objective function was the structural performance of the shell, there are a set of diverse

solutions in the provided Pareto front that are similar to the Isler solution (for instance, see solutions 2 and 5 in Table 5 and Figure 12). This result indicates that MOO is a reliable method for form-finding and optimization and is able to generate accurate and good results.

5. FUTURE RESEARCH DIRECTION

Based on the existing limitation in current methods, suggestions for the future research direction could be listed as follows:

- Utilizing more criteria and objective functions in the design of concrete thin shell structures, such as 'Embodied Carbon assessment,' 'life cycle assessment,' cost, measuring energy performance, acoustic performance of the structure, and considering fabrication methods as an objective and other criteria.
- Utilizing other metaheuristic algorithms in the design of thin shell structures that have not been employed before, such as Particle swarm optimization algorithm, Graywolf optimization algorithm, Dolphin Echolocation algorithms, and other algorithms, to compare the convergence rate, accuracy, and speed of these algorithms.
- Developing a scientific method to select a final solution from the set of optimal solutions in the Pareto Frontier.
- Utilizing Machine Learning techniques (76) to find the optimal design solutions, constructing reliable Datasets, and training algorithms to find optimal solutions (prediction) even without computing (77), (78).

6. STATEMENTS AND DECLARATIONS

Funding

Not applicable

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Generative AI and AI-assisted Technologies in the writing process

The authors did not use AI or AI-assisted technologies except for the English language grammar and punctuation check's purpose.

7. AUTHORSHIP CONTRIBUTION STATEMENT

MV: Conceptualization, Methodology, Software, Data curation, Validation, Visualization, Writing - original

draft. **MG:** Project administration, Supervision, Writing - Review & Editing. **AE:** Supervision, Writing - Review & Editing. **MR:** Software, Validation, Writing - Review & Editing.

8. AVAILABILITY OF DATA AND MATERIALS

The corresponding author, upon responsible request, will provide some or all data, models, or codes that support the findings of this study.

9. ACKNOWLEDGMENT

The authors thank the support provided by the Department of Architectural Technology, Faculty of Architecture, College of Fine Arts, at the University of Tehran, Iran.

10. APPENDIX

TABLE 7. Abbreviations Table

SOO	Simple objective optimization	DM	Decision Maker
MOO	Multi-objective optimization	GA	Genetic Algorithm
NURBS	Non-Uniform Rational B-spline	PSO	Particle Swarm Optimization
NSGA-II	Non-dominated Sorting Genetic Algorithm	DE	differential evolution (DE) algorithm
MOGA	Multi-objective Genetic Algorithm	PF	Pareto Optimal Front
MOGWO	Multi-objective gray wolf optimizer	BESO	Bi-directional Evolutionary Structural Optimization algorithm

11. REFERENCES

1. Adesina A. Recent advances in the concrete industry to reduce its carbon dioxide emissions. *Environmental Challenges*. 2020;1:100004.
2. Chilton J, Isler H. *Heinz isler*: Thomas Telford; 2000.
3. Goaiz HA, Shamsaldeen HA, Abdulrehman MA, Al-Gasham T. Evaluation of steel fiber reinforced geopolymer concrete made of recycled materials. *International Journal of Engineering*. 2022;35(10):2018-26.
4. Farshad M. *Design and analysis of shell structures*: Springer Science & Business Media; 2013.
5. Adriaenssens S, Block P, Veenendaal D, Williams C. *Shell structures for architecture: form finding and optimization*: Routledge; 2014.
6. Council WGB. *New report: The building and construction sector can reach net zero carbon emissions by 2050*. World Green Building Council Toronto, ON, Canada; 2019.

7. IEA U, editor Global status report for buildings and construction 2019/20: IEA Paris, France.
8. Zhong X, Hu M, Deetman S, Steubing B, Lin HX, Hernandez GA, et al. Global greenhouse gas emissions from residential and commercial building materials and mitigation strategies to 2060. *Nat Commun.* 2021;12(1):6126. 10.1038/s41467-021-26212-z <https://www.ncbi.nlm.nih.gov/pubmed/34675192>
9. Block P, Schlueter A, Veenendaal D, Bakker J, Begle M, Hischer I, et al. NEST HiLo: Investigating lightweight construction and adaptive energy systems. *J Build Eng.* 2017;12:332-41. 10.1016/j.job.2017.06.013
10. Hosseinian SM, Faghani M. Assessing the effect of structural parameters and building site in achieving low carbon building materialization using a life-cycle assessment approach. *J Build Eng.* 2021;44:103318. 10.1016/j.job.2021.103318
11. Pawar AJ, Suryawanshi SR. Comprehensive Analysis of Stress-strain Relationships for Recycled Aggregate Concrete. *International Journal of Engineering.* 2022;35(11):1-10. 10.5829/ije.2022.35.11b.05
12. Sreeman D, Kumar Roy B. Optimization study of isolated building using shape memory alloy with friction pendulum system under near-fault excitations. *International Journal of Engineering.* 2022;35(11):2176-85.
13. Talib HY, Al-Salim NHA. Improving Punching Shear in Flat Slab by Replacing Punching Shear Reinforcement by Ultrahigh Performance Concrete. *International Journal of Engineering.* 2022;35(8):1619-28. 10.5829/ije.2022.35.08b.18
14. Kotnik T, Schwartz J. The architecture of heinz isler. *Journal of the international association for shell and spatial structures.* 2011;52(3):185-90.
15. Kaveh A. *Advances in metaheuristic algorithms for optimal design of structures*: Springer; 2014.
16. Talbi E-G. *Metaheuristics: from design to implementation*: John Wiley & Sons; 2009.
17. Kaveh A, Ghazaan MI. *Meta-heuristic algorithms for optimal design of real-size structures*: Springer; 2018.
18. Mirjalili S, Faris H, Aljarah I. *Evolutionary machine learning techniques*: Springer; 2019.
19. Pugnale A, Sassone M. Morphogenesis and structural optimization of shell structures with the aid of a genetic algorithm. *Journal of the International Association for Shell and Spatial Structures.* 2007;48(3):161-6.
20. Gokul Santhosh S, Singh A, Abdul Akbar M. Parametric Studies and Optimization of Grid Shell Structures Using Genetic Algorithm. *Recent Trends in Civil Engineering: Select Proceedings of ICRAE 2021*: Springer; 2022. p. 109-22.
21. Basso P, Del Grosso AE, Pugnale A, Sassone M. Computational morphogenesis in architecture: Cost optimization of free-form grid shells. *Journal of the international association for shell and spatial structures.* 2009;50(3):143-50.
22. Baghdadi A, Doerrie R, Kloft H, editors. Application of different optimisation techniques for material minimisation of individual structural elements. *Proceedings of IASS Annual Symposia*; 2020: International Association for Shell and Spatial Structures (IASS).
23. Ansola R, Canales J, Tarrago JA, Rasmussen J. Combined shape and reinforcement layout optimization of shell structures. *Struct Multidiscip O.* 2004;27(4):219-27. 10.1007/s00158-004-0399-7
24. Kimura T, Ohmori H. Computational morphogenesis of free form shells. *Journal of the International Association for Shell and Spatial structures.* 2008;49(3):175-80.
25. Yang F, Yu T, Bui TQ, editors. Morphogenesis of free-form surfaces by an effective approach based on isogeometric analysis and particle swarm optimization. *Structures*; 2023: Elsevier.
26. Tomás A, Martí P. Shape and size optimisation of concrete shells. *Engineering Structures.* 2010;32(6):1650-8.
27. Feng RQ, Ge JM. Shape optimization method of free-form cable-braced grid shells based on the translational surfaces technique. *Int J Steel Struct.* 2013;13(3):435-44. 10.1007/s13296-013-3004-3
28. Hassani B, Tavakkoli SM, Ghasemnejad H. Simultaneous shape and topology optimization of shell structures. *Struct Multidiscip O.* 2013;48(1):221-33. 10.1007/s00158-013-0894-9
29. Richardson JN, Adriaenssens S, Coelho RF, Bouillard P. Coupled form-finding and grid optimization approach for single layer grid shells. *Engineering Structures.* 2013;52:230-9. 10.1016/j.engstruct.2013.02.017
30. Shimoda M, Liu Y. A non-parametric free-form optimization method for shell structures. *Struct Multidiscip O.* 2014;50(3):409-23. 10.1007/s00158-014-1059-1
31. Kaveh A, Amirsoleimani P, Eslamlou AD, Rahmani P, editors. Frequency-constrained optimization of large-scale dome-shaped trusses using chaotic water strider algorithm. *Structures*; 2021: Elsevier.
32. Tomei V, Grande E, Imbimbo M. Design optimization of gridshells equipped with pre-tensioned rods. *J Build Eng.* 2022;52:104407. 10.1016/j.job.2022.104407
33. Ansola R, Canales J, Tarrago J, Rasmussen J. On simultaneous shape and material layout optimization of shell structures. *Struct Multidiscip O.* 2002;24:175-84.
34. Gythiel W, Schevenels M. Gradient-based size, shape, and topology optimization of single-layer reticulated shells subject to distributed loads. *Struct Multidiscip O.* 2022;65(5):144. 10.1007/s00158-022-03225-w
35. Teimouri M, Asgari M. Developing an efficient coupled function-based topology optimization code for designing lightweight compliant structures using the BESO algorithm. *Optim Eng.* 2023;1-29. 10.1007/s11081-023-09808-w
36. Zhang ZJ, Kanno Y, Feng RQ. Collapse-resistance optimization of fabricated single-layer grid shell based on sequential approximate optimization. *Struct Multidiscip O.* 2023;66(7):153. 10.1007/s00158-023-03604-x
37. Emami N, Giles H, von Buelow P. Structural, daylighting, and energy performance of perforated concrete shell structures. *Automat Constr.* 2020;117:103249. 10.1016/j.autcon.2020.103249
38. Liuti A, Pugnale A, Erioli A, editors. Computational morphogenesis in architecture: Structure and light as a multi-objective design/optimization problem. *Proceedings of the 2nd International Conference on Structures and Architecture (ICSA 2013)*, Guimaraes, Portugal; 2013.
39. Emami N, editor Structure and daylighting performance comparisons of Heinz Isler's roof shell based on variations in parametrically derived multi opening topologies. *Proceedings of IASS Annual Symposia*; 2015: International Association for Shell and Spatial Structures (IASS).
40. Pugnale A. Integrating Form, Structure and Acoustics: A Computational Reinterpretation of Frei Otto's Design Method and Vision. *Journal of the International Association for Shell and Spatial Structures.* 2018;59(1):75-86. 10.20898/j.iass.2018.195.901.
41. Turrin M, von Buelow P, Stouffs R. Design explorations of performance driven geometry in architectural design using parametric modeling and genetic algorithms. *Adv Eng Inform.* 2011;25(4):656-75. 10.1016/j.aei.2011.07.009
42. Puppa GD, Trautz M, editors. Multi-objective shape optimization of shell structures based on buckling forms. *Proceedings of IASS Annual Symposia*; 2015: International Association for Shell and Spatial Structures (IASS).

43. Nagata K, Honma T, editors. Multi-objective Optimization for Free Form Surface Shell Using Swarm Intelligence with Decent Solutions Search. Proceedings of IASS Annual Symposia; 2016: International Association for Shell and Spatial Structures (IASS).
44. Wang XY, Zhu SJ, Zeng Q, Guo XN. Improved multi-objective Hybrid Genetic Algorithm for Shape and Size Optimization of Free-form latticed structures. *J Build Eng.* 2021;43:102902. <https://doi.org/10.1016/j.job.2021.102902>
45. Henriksson V, Lomax S, Richardson JN, editors. Multi-Objective Optimization of a Concrete Shell. Proceedings of IASS Annual Symposia; 2019: International Association for Shell and Spatial Structures (IASS).
46. Veenendaal D. DESIGN AND FORM FINDING OF FLEXIBLY FORMED CONCRETE SHELL STRUCTURES: ETH Zurich; 2017.
47. Veenendaal D, Bakker J, Block P. Structural Design of the Flexibly Formed, Mesh-Reinforced Concrete Sandwich Shell Roof of Nest Hilo. *Journal of the International Association for Shell and Spatial Structures.* 2017;58(1):23-38. [10.20898/j.iass.2017.191.847](https://doi.org/10.20898/j.iass.2017.191.847)
48. Zhao ZW, Wu JJ, Qin Y, Liang B. The strong coupled form-finding and optimization algorithm for optimization of reticulated structures. *Adv Eng Softw.* 2020;140:102765. [10.1016/j.advengsoft.2019.102765](https://doi.org/10.1016/j.advengsoft.2019.102765)
49. Vargas DEC, Lemonge ACC, Barbosa HJC, Bernardino HS. Differential evolution with the adaptive penalty method for structural multi-objective optimization. *Optim Eng.* 2019;20(1):65-88. [10.1007/s11081-018-9395-4](https://doi.org/10.1007/s11081-018-9395-4)
50. Mirra G, Pugnale A, editors. Comparison between human-defined and AI-generated design spaces for the optimisation of shell structures. *Structures*; 2021: Elsevier.
51. Nishei T, Fujita S, editors. Structural optimization of latticed shells considering strain energy and collapse load factor. Proceedings of IASS Annual Symposia; 2020: International Association for Shell and Spatial Structures (IASS).
52. Wang H, Li PC, Wang J. Shape optimization and buckling analysis of novel two-way aluminum alloy latticed shells. *J Build Eng.* 2021;36:102100. [10.1016/j.job.2020.102100](https://doi.org/10.1016/j.job.2020.102100)
53. Kostura Z, Clark MC, Olson JR, editors. Design of a 500,000 m² Iconic Roof for the New International Airport for Mexico City. Proceedings of IASS Annual Symposia; 2018: International Association for Shell and Spatial Structures (IASS).
54. Jiang B, Zhang J, Ohsaki M, editors. Shape optimization of free-form shell structures combining static and dynamic behaviors. *Structures*; 2021: Elsevier.
55. Cao Z, Wang Z, Zhao L, Fan F, Sun Y. Multi-constraint and multi-objective optimization of free-form reticulated shells using improved optimization algorithm. *Engineering Structures.* 2022;250:113442.
56. Desai R. Structural Optimization Using Genetic Algorithm. *Handbook of AI-based Metaheuristics*: CRC Press; 2021. p. 73-116.
57. Reisinger J, Zahlbruckner MA, Kovacic I, Kán P, Wang-Sukalia X, Kaufmann H. Integrated multi-objective evolutionary optimization of production layout scenarios for parametric structural design of flexible industrial buildings. *J Build Eng.* 2022;46:103766.
58. Preisinger C. Karamba3D. Vienna: Bollinger und Grohmann ZT GmbH. 2018.
59. Vatandoost M, Yazdanfar S, Ekhlassi A. Computer-Aided Design: Classification of Design problems in Architecture. *Journal of Information and Computational Science.* 2019;9(10):605-25.
60. Mirjalili S, Dong JS. Multi-objective optimization using artificial intelligence techniques: Springer; 2020.
61. Rao SS. Engineering optimization: theory and practice: John Wiley & Sons; 2019.
62. Yi YK, Tariq A, Park J, Barakat D. Multi-objective optimization (MOO) of a skylight roof system for structure integrity, daylight, and material cost. *J Build Eng.* 2021;34:102056. [10.1016/j.job.2020.102056](https://doi.org/10.1016/j.job.2020.102056)
63. Smith G, Rivlin R. The strain-energy function for anisotropic elastic materials. *Collected Papers of RS Rivlin: Volume I and II.* 1997:541-59.
64. Al-Tashi Q, Mirjalili S, Wu J, Abdulkadir SJ, Shami TM, Khodadadi N, et al. Moth-Flame Optimization Algorithm for Feature Selection: A Review and Future Trends. *Handbook of Moth-Flame Optimization Algorithm.* 2022:11-34.
65. Vatandoost M, Ekhlassi A, Yazdanfar S. Computer-aided architectural design: classification and application of optimization algorithms. *Journal of Information and Computational Science.* 2020;10(6):18-43.
66. Chatterjee S, Sarkar S, Dey N, Ashour AS, Sen S. Hybrid non-dominated sorting genetic algorithm: II-neural network approach. *Advancements in Applied Metaheuristic Computing*: IGI Global; 2018. p. 264-86.
67. Sayyaadi H. Modeling, assessment, and optimization of energy systems: Academic Press; 2020.
68. Deb K, Agrawal S, Pratap A, Meyarivan T, editors. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. *Parallel Problem Solving from Nature PPSN VI: 6th International Conference Paris, France, September 18–20, 2000 Proceedings 6*; 2000: Springer.
69. Yang X-S. Nature-inspired optimization algorithms: Academic Press; 2020.
70. Baghdadi A, Heristchian M, Kloft H, editors. Displacement-based seismic assessment of Heinz Isler's shell structures. Proceedings of IASS Annual Symposia; 2020: International Association for Shell and Spatial Structures (IASS).
71. Vu-Bac N, Duong TX, Lahmer T, Zhuang X, Sauer RA, Park HS, et al. A NURBS-based inverse analysis for reconstruction of nonlinear deformations of thin shell structures. *Comput Method Appl M.* 2018;331:427-55. [10.1016/j.cma.2017.09.034](https://doi.org/10.1016/j.cma.2017.09.034)
72. Zhang XK, Jin CN, Hu P, Zhu XF, Hou WB, Xu JT, et al. NURBS modeling and isogeometric shell analysis for complex tubular engineering structures. *Comput Appl Math.* 2017;36(4):1659-79. [10.1007/s40314-016-0312-1](https://doi.org/10.1007/s40314-016-0312-1)
73. Baghdadi A, Heristchian M, Kloft H. Structural assessment of remodelled shells of Heinz Isler. *International Journal of Advanced Structural Engineering.* 2019;11(4):491-502.
74. Piker D. Kangaroo Form Finding with Computational Physics. *Archit Design.* 2013;83(2):136-7. [10.1002/ad.1569](https://doi.org/10.1002/ad.1569)
75. Sassone M, Pugnale A, editors. Evolutionary Structural Optimization in Shell Design. *Advanced Numerical Analysis of Shell-like Structures: Special Workshop, Zagreb, 26-28 September 2007*; 2007.
76. Ramavath S, Suryawanshi S. Optimal Prediction of Shear Properties in Beam-Column Joints Using Machine Learning Approach. *International Journal of Engineering.* 2024;37(1):67-82.
77. El Alaoui M, Chahidi LO, Rougui M, Lamrani A, Mechaqrane A. Prediction of Energy Consumption of an Administrative Building using Machine Learning and Statistical Methods. *Civil Engineering Journal.* 2023;9(5):1007-22.
78. Bessa M, Pellegrino S. Design of ultra-thin shell structures in the stochastic post-buckling range using Bayesian machine learning and optimization. *International Journal of Solids and Structures.* 2018;139:174-88.

COPYRIGHTS

©2024 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 این مطالعه بهینه‌سازی شکل و ضخامت سازه‌های پوسته بتنی پیوسته نازک ارائه شده است که وزن، انحراف قائم و تغییر انرژی الاستیک آنها را به حداقل می‌رساند در حالی که الزامات عملکردی را برآورده می‌کند و مصرف مصالح را به حداقل می‌رساند. برخلاف مطالعات قبلی که بیشتر بهینه‌سازی تک هدفه مورد توجه قرار گرفته است، این تحقیق با در نظر گرفتن سه تابع هدف، بر بهینه‌سازی چندهدفه تمرکز دارد. این ترکیب از توابع هدف در تحقیقات قبلی منعکس نشده است که این مطالعه را متمایز می‌کند. گردش کار طراحی محاسباتی شامل یک مدل پارامتریک، مؤلفه های متعدد برای اندازه گیری توابع هدف در گرس‌هایمپیر راینر و یک الگوریتم فراابتکاری، الگوریتم ژنتیک چندهدفه مرتب‌سازی غیرغالب (NSGA-II)، به عنوان ابزار جستجو است که در زبان برنامه نویسی پایتون کدنویسی شده است. این گردش کار به ما امکان می‌دهد تا فرم‌یابی و بهینه‌سازی را به طور همزمان انجام دهیم. برای نشان دادن موثر بودن این الگوریتم فراابتکاری در بهینه‌سازی سازه‌ها، ما آن را در یک مطالعه موردی، در یک پوسته شناخته شده طراحی شده با استفاده از تکنیک شبیه‌سازی فیزیکی مدل‌آویخته به کار بردیم. تفاسیر چند نمونه از نتایج بهینه شده نشان می‌دهد که اگرچه وزن پاسخ-۱ تقریباً مشابه پاسخ-۲ است، اما انرژی کرنشی و انحراف قائم کمتری دارد. پاسخ-۳، با جرم سه برابر، انحراف قائم و انرژی کرنشی به طور قابل توجهی کمتری از پاسخ ۱ و پاسخ ۲ دارد که به ترتیب بیش از ۵۰٪ و ۱۷٪ کمتر است. پاسخ-۳ و پاسخ-۴ عملکرد سازه‌ای بهتری در انحراف قائم و انرژی کرنشی نشان می‌دهند. علاوه بر این، مقایسه نتایج بهینه‌سازی چندهدفه با پوسته طرح شده/ایسلر Isler نشان داد که این روش پاسخ سازه‌ای با وزن و انحراف قائم کمتر و در عین حال سازه‌ای سخت‌تر پیدا کرد و کاربردی بودن این روش را تأیید کرد. این مطالعه نشان داد که بهینه‌سازی چندهدفه روشی قابل اعتماد برای فرم‌یابی و بهینه‌سازی پوسته‌ها است که نتایج دقیق و قابل قبولی ایجاد می‌کند.



A Model for Analysis of Social Media in Adoption of Mobile Banking

H. Hamidi*, M. Mohammadi

Department of Industrial Engineering, Information Technology Group, K. N. Toosi University of Technology, Tehran, Iran

PAPER INFO

Paper history:

Received 26 December 2023

Received in revised form 01 February 2024

Accepted 05 February 2024

Keywords:

Banking

Social Media

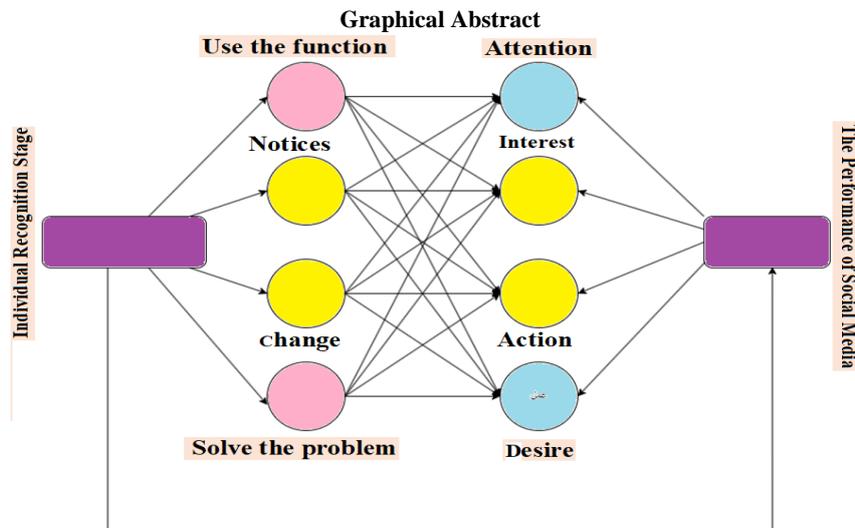
Mobile Banking Acceptance

Mobile Banking

ABSTRACT

Banks can design more efficient methods for customer acquisition by utilizing social media platforms. By monitoring information from social media platforms, banks can analyze customers' reactions to offer by competitors and adopt appropriate strategies to increase customer satisfaction and attract new customers. Present research focuses on analyzing the role of social media in the acceptance of mobile banking at Bank Melli of Qazvin Province. This research is of a survey and applied nature. The data collection tool is a survey, and the data collection method is fieldwork. The study population includes all managers and deputies of Bank Melli of Qazvin Province in various positions. The sample size is calculated using the Cochran formula due to the limited population. In this study, the indicators were identified, and hypotheses were formulated based on these indicators. Finally, using statistical techniques, all the proposed hypotheses were proven and the impact of all indicators was confirmed. The results have confirmed the effect of social media performance on the individual recognition of mobile banking consumers.

doi: 10.5829/ije.2024.37.07a.16



1. INTRODUCTION

Social media networks have become a powerful force in shaping various aspects of business (1). This influence can be evaluated from various perspectives. The movement of customers towards new media, the declining effectiveness of traditional advertising, the

importance of word-of-mouth advertising, especially in the virtual space, access to a vast audience and markets, unique opportunities of the virtual space in establishing and developing communications, the capabilities of social media platforms in product presentation and sales, brand strengthening, and more, are among the most significant impacts of social media in the realm of

*Corresponding author email: h.hamidi@kntu.ac.ir (H. Hamidi)

commerce (2, 3). Banks, as one of the key players in the business and commerce sector, are not exempt from this rule, as we witness their active presence on social media platforms every day. Depending on the adopted strategies, they show special attention to social media as a new paradigm in engaging with their customers (4, 5).

The use of social media has become so widespread that banks have no choice but to join these platforms for their marketing activities. Strategic use of social media can impact not only the marketing of products and services but also risk management, product and service design, market prediction, competitor analysis, and customer education. Banks and financial institutions that can quickly align their business operations with appropriate social media strategies will be able to better and more effectively meet customer needs and provide them with the best experiences. According to research conducted, banks have recognized the power of social media as a channel for two-way communication with customers, resulting in a significant decrease in face-to-face interactions in bank branches (6, 7). Banks utilize social media to identify customer needs and preferences, provide quick and timely responses, introduce new services and products, address inquiries and issues, engage with customers, develop and promote their brand (8). However, social media platforms have not reached a level of maturity where they are used for financial transactions. Financial service companies that use social media for marketing purposes and have successfully facilitated customer engagement on these platforms have gained a competitive advantage in terms of increased customer loyalty to their brand compared to their competitors. To fully harness the potential of social media channels, banks need to view it as more than just a means of enhancing brand credibility. In this regard, banks need a better understanding of their customers by utilizing social analytics to gain deep insights into their behaviors, emotions, and needs. Banks also need to integrate their traditional systems with modern social analytics tools and methods (9, 10). When these tools and methods are integrated with a strategy that encompasses all customer communication channels, social analyses can transform the current customer relationship management system into a social customer relationship management system. The pace of changes in the banking industry over the past decade has been so high that predicting the future of banking over the next five years seems very challenging. The driving force behind these changes is not only technological advancement but also intense competition among banks to attract and retain customers (11). Most people prefer banking services that revolve around their lifestyle and how they live: services that are simple to use, available around the clock, and at the same time ensure the security of our data. Banks have been talking about "customer-centricity" for almost a decade (12, 13). Now is the time for them to act on their

words because if they do not, intense competition in the market and the ease of switching banks have made it easy for them to lose their customers. While no one can predict how banking will be tomorrow, the path of this journey is clear. The banks of the future face some challenges that they need to address. Banks may appear more future-oriented, but their key aspects will remain the same. According to Eric Linders, the executive director of retail sales at BBA Banking, it is very straightforward and transparent: you protect people's assets, provide them with credibility, and enable them to transfer their money from A to B. In the next ten years, we will continue to do the same things, but new technologies will enable us to provide these services in different methods.

This study aimed to propose a model for analysis of social media in the adoption of mobile banking. In the following, the research literature is introduced in section 2. In section 3, research methodology is presented, and then in section 4, the analysis of the information is given. Finally, in section 5 the conclusion is presented.

2. LITERATURE REVIEW

The rapid growth of technology and the entry of new competitors in the banking sector necessitate the use of innovative service management. In this field, banks will succeed if they are flexible through changes and strengthening of human resources, create value for customers. Information technology in banks has undergone various changes over different periods. Technology, as a powerful tool in human communications, has multiple functions and its impact on societal conditions is significant (14-16). Today, banks will be able to turn their conditions from competitive equality with other banks into a competitive advantage by having a high level of innovation. From the perspective of resource-based theories and dynamic capabilities, achieving sustainable competitive advantage requires capabilities that are not easily obtained. Additionally, these capabilities are sometimes quite expensive. Relying solely on operational and recognized capabilities, such as advertising or improving service quality, will only guarantee competitive equality for banks active in the Iranian market. Technology has transformed the banking system and led to the transformation of monetary and financial services. For survival and progress in the competitive space, banks need to provide innovative services. Banks must be market-oriented, customer-centric, and engage in internal and external innovations because modern banking aims to facilitate and increase the speed of service delivery. Innovative banking services are among the important factors in creating a competitive advantage for banks in attracting and increasing customer satisfaction (17, 18). The most important capability that our banks need today

to achieve above-average industry performance or maintain these performances is the power of innovation. Innovation and technology in the banking industry are intertwined. It is evident that in this competitive environment, banks and institutions will be more successful if they move from traditional and experienced approaches to a continuous and unstoppable movement towards newer approaches. They should develop fundamental capabilities and enhance competitive advantages in progress, utilizing new technologies and electronic systems. Electronic banking is considered one of the fundamental and important principles of the progress of any bank in the modern era (19, 20).

Electronic banking, online banking, or internet banking refers to providing facilities for employees to increase their speed and efficiency in providing banking services at branch locations, as well as inter-branch and inter-bank processes worldwide. It also involves offering hardware and software capabilities to customers, allowing them to perform their desired banking operations securely and confidently through safe communication channels at any hour of the day (24/7) without the need for physical presence in a bank. In other words, electronic banking involves using advanced network-based software and hardware technologies for the electronic exchange of financial resources and information, eliminating the need for customers' physical presence in branches. It allows customers to conduct economic transactions on a secure website, such as retail banking or virtual banking, financial and credit institutions, or construction companies, at any time of the day. Electronic banking requires a forward-looking approach in various areas, and to progress in electronic banking, it is necessary to use future-oriented principles and dimensions. Following old principles is no longer viable, and new structures are needed. Our ancestors believed that the future is similar to the past. However, with a superficial look at the speed of changes and advancements in various fields today, we realize that the future is no longer the same as the past. Humanity has always desired to dominate nature and change it according to its will, seeking tools to achieve this. The modern banking system in the field of information technology is no exception to this rule, and comprehensive foresight in the field of information technology plays a fundamental role in the progress of banks in technology. One of the innovations that has experienced significant growth due to extensive developments in information and communication technology is e-commerce.

3. METHODOLOGY

Banks can design more efficient methods for customer attraction by utilizing social media. Social media and mobile-accessible applications can help banks penetrate geographical areas where physical presence has been challenging, introducing their products and services. By monitoring information from social media, banks can analyze customer reactions to offer by competitors and

adopt appropriate strategies to increase customer satisfaction and attract new customers. Social media marketing is also a useful tool for attracting young people, as they are the largest followers of bank pages on social media and can persuade older customers to use bank products and services. To attract more customers, entertaining features unrelated to banking topics can be presented on the social media pages of banks. Some banks support non-governmental organizations active in education, public health, and sustainable livelihoods and engage in charitable, humanitarian, and volunteer activities on social media to attract attention and increase visits to their social media pages. Recently, social networks have become a focal point for banks and are considered a key area in changing the service and customer communication approaches of banks. Social media is an acceptable option for banks to differentiate themselves from competitors. The primary function of social media for banks is as a tool for marketing, advertising, and providing information to customers to encourage them to visit a branch or advertise services. However, social media is not just a marketing tool; but, it is a novel opportunities can be brought about using these networks for banks. In fact, with the continuous expansion of social media, a significant percentage of bank customers are members of social networks and use these platforms in their daily lives. Therefore, social networks can be used as an important platform for promoting and even selling the products and services of various businesses. This platform creates a suitable opportunity for banks to align themselves with the emerging social business model and markets. Moreover, by analyzing the large volume of customer data, this platform can assess customer behavior and contribute to provide better, personalized, and new products and services. Additionally, this platform can serve as a new channel for providing financial services to target customers. This article attempts to highlight the importance of using social media in the banking industry, explore the opportunities and advantages of employing these media in the banking industry, and outline the steps and some implementation challenges of social banking (21-24).

While the banking industry traditionally seeks to maintain customer relationships, the needs and desires of customers have evolved significantly over the past decade. The emergence of social networks has rapidly changed customer expectations and their interaction with banks. With the expansion of social media platforms such as Facebook, Twitter, Telegram, etc., we observe a growing number of individuals incorporating these platforms into their daily lives. As the number of people using social media continues to rise, customers expect banks to provide their services through these platforms. They anticipate banks to utilize social media channels for quicker and more effective service and product offerings,

personalized financial advice, sharing financial proposals and future plans, addressing banking regulations, and establishing mechanisms for customer feedback on their banking products and services (10, 25, 26).

For banks to continue their existence in the 21st century, retain customers, and gain a competitive advantage, they must pay special attention to this new service channel. By analyzing the vast amount of data available on social media, banks can identify key factors that enable them to offer better products and services, improve customer service, enhance marketing strategies, manage risks, and increase business efficiency.

Today, communication methods within organizations have evolved, and banks, in turn, must act more customer-centric and innovative in the field of communications. Social media is a direct result of these changes. Alt and Puschmann (27) believed that four drivers will significantly transform banks in the coming years, and all these drivers are directly and indirectly related to the phenomenon of social media:

1. Financial challenges
2. Changes in customer banking behavior
3. Rapid dissemination of new information technology solutions
4. Emergence of non-banking institutions

Currently, more than 80% of internet users worldwide use social media networks for their daily activities. Therefore, it is not surprising that a significant portion of these users also perform their banking activities through social media platforms. Banks must pay special attention to this growing space. Gartner defines social media-based banking as an emerging approach in retail banking that makes lending processes and relationships between lenders and financial institutions transparent. Social media banking includes a package of well-defined services similar to traditional banking services, but this time on social networks. It also comes with value-added services that banks previously offered in a limited way to depositors. With the capabilities provided by online networks and tools, banks can offer value-added services to customers at any level. Banks can use social networks for informational purposes, expose their products and services to reviews and critiques, adapt and adjust their services programs to customer tastes and expectations based on user feedback. Overall, customer identification, data acquisition, understanding behavior, examining preferences, interaction, increasing loyalty, and increasing the number of customers are some of the advantages of paying attention to social media networks in the banking system. Considering the effectiveness of these media, banks can adopt innovative strategies in various functional dimensions, including marketing, customer service, communications, online banking, etc. (28-30). This article aims to explore the importance of using social media in the banking industry, the opportunities and advantages of employing these media

in the banking sector, and outlines the steps for implementing social banking and some of the challenges it presents. Additionally, the article describes the gradual evolution of the use of social media and the impact of customer and bank maturity (in using social media) on choosing an appropriate social strategy.

3. 1. Research Objectives • Analysis of the role of social media in accepting banking with Bank Melli of Qazvin province

- Investigating the effect of each of the mentioned indicators
- Providing solutions to solve existing problems.

3. 2. The Main Hypothesis • The performance of social media has an impact on the stage of personal recognition of mobile banking consumers.

3. 3. Sub Hypothesis • The performance of social media has an effect on the use of mobile banking by consumers.

- The performance of social media has an impact on informing mobile banking consumers.
- The performance of social media has an effect on solving the problem of mobile banking consumers.
- The performance of social media has an effect on the change in the use of mobile banking by consumers.
- The stage of personal recognition has an effect on the attention of mobile banking consumers.
- The level of personal recognition has an effect on consumers' interest in mobile banking.
- The stage of personal recognition has an effect on the willingness of mobile phone banking consumers.
- The stage of personal recognition has an effect on the action of mobile banking consumers.

4. DATA ANALYSIS

This research is of applied and survey type. The method of collecting information is in the field, and library studies are also used. In this research, the structural equation technique is used to analyze the data, and the software used in this research is Laserl.

In the following, other descriptive indices such as mean and standard deviation for all research variables are discussed. Table 1 demonstrates descriptive statistics including mean and standard deviation and Kolmogorov-Smirnov test for each of the research variables including dimensions of social media performance (function use, information, problem solving, change) and individual recognition stage (attention, interest, desire, action). In choosing a statistical test for research, we must decide whether to use parametric tests or non-parametric tests. One of the main criteria for this selection is the Kolmogorov-Smirnov test. The Kolmogorov-Smirnov

test shows the non-normality of the data distribution. That is, it compares the distribution of an attribute in a sample (for example, age among 100 nurses) with the distribution that is assumed for society (for example, the age of all nurses). If the Kolmogorov-Smirnov test is rejected, the data have a normal distribution, and it is possible to use parametric statistical tests for research. On the contrary, if the Kolmogorov-Smirnov test is confirmed, it means that the data does not have a normal distribution, so we should use non-parametric tests in the research.

When checking the normality of the data, we test the null hypothesis that the distribution of the data is normal at the 5% error level. Therefore, if the larger test statistic equal to 0.05 is obtained, then there will be no reason to reject the null hypothesis based on the fact that the data is normal. In other words, the data distribution will be normal. For the normality test, the statistical assumptions are set as follows:

H0: The distribution of data related to each of the variables is normal.

H1: The distribution of data related to each of the variables is not normal.

Based on the results listed in Table 1, all the test statistic values are less than 0.05, as a result, the distribution of measurement data of each dimension is abnormal. Therefore, non-parametric tests can be used.

4. 1. Partial Least Squares Technique and Test of Research Hypotheses

The theoretical foundations related to the partial least squares method (PLSM) will be presented in detail. PLSM is one of the second-generation structural equation modeling approaches and has advantages compared to the first-generation methods that were based on covariance. Here, we have used the second-generation method of structural equation modeling of component-based methods, which was later changed to the partial least squares (PLS) method, which was invented by Alt and Puschmann (27). The PLS method consists of two main steps:

1) Examining the fit of measurement models, structural model.

2) Testing relationships between structures.

- Confirmatory factor analysis of research variables

Before entering the stage of testing hypotheses and conceptual research models, it is necessary to ensure the accuracy of independent variable and dependent variable measurement models. Therefore, in the following, the measurement models of these two variables are presented in order, which was done by first-order confirmatory factor analysis. Confirmatory factor analysis is one of the oldest statistical methods that is used to investigate the relationship between existing variables (obtained factors) and observed variables (questions) and represents the measurement model. Figures 1 and 2 show the confirmatory factor analysis of research variables. The factor loadings of the model in non-standard and standard estimation mode show the influence of each of the variables or items in explaining the variance of the variable or main factor scores. In other words, the factor load indicates the degree of correlation of each observed variable (questionnaire) with the underlying variable (factors). If the value of the variable significance test statistic is higher than 1.96, it indicates that the considered item is significant, otherwise the item is removed.

According to Figure 1, it shows that the influence of the factor (hidden variable) and the observable variable is shown by the factor load. Factor load is a value between zero and one. If the factor load is less than 0.4, the relationship is considered weak and is ignored. A factor between 0.4 and 0.6 is acceptable, and if it is greater than 0.6, it is very desirable.

According to Figure 2, it shows that when the correlation of variables is identified, a significance test should be performed. To check the significance of observed correlations, bootstrap or jackknife cross cutting methods are used. In this study, the self-adjustment method was used, which gives the t statistic. At the 5% error level, if the bootstrapping t-value is greater than 1.96, the observed correlations are significant.

According to Table 2, the stated variables are used to examine the measurement model or (confirmatory factor

TABLE 1. Descriptive statistics

Variable		Meaningful	Test statistics	Standard deviation	Average	Result
The performance of social media	Use the function	0.001	0.025	0.325	4.81	abnormal
	Notices	0.002	0.023	0.256	2.32	abnormal
	Solve the problem	0.002	0.062	0.424	3.64	abnormal
	change	0.003	0.052	0.425	3.52	abnormal
Individual recognition stage	Attention	0.004	0.064	0.456	4.12	abnormal
	interest	0.006	0.078	0.512	3.15	abnormal
	desire	0.023	0.075	0.532	4.25	abnormal
	Action	0.036	0.062	0.789	3.13	abnormal

TABLE 2. Summary of the results of confirmatory factor analysis (measurement model) of the used scale

Object	Load factor	Meaningful	Object	Load factor	Meaningful
Q1	707.0	808.4	Q18	725.0	902.9
Q2	738.0	786.6	Q19	862.0	823.6
Q3	758.0	491.6	Q20	623.0	850.6
Q4	768.0	5.823	Q21	723.0	320.5
Q5	869.0	3.708	Q22	845.0	5.745
Q6	725.0	3.678	Q23	742.0	702.2
Q7	702.0	2.823	Q24	0.702	102.3
Q8	826.0	2.798	Q25	755.0	4.145
Q9	732.0	3.797	Q26	802.0	652.7
Q10	770.0	3.532	Q27	755.0	6.102
Q11	707.0	3.144	Q28	745.0	5.725
Q12	0.789	3.602	Q29	702.0	6.742
Q13	0.820	14.864	Q30	764.0	6.722
Q14	0.822	2.264	Q31	756.0	6.426
Q15	0.804	20.074	Q32	702.0	9.123
Q16	0.823	5.653			
Q17	0.952	4.752			

preferences of mobile banking is equal to 0.802. The stage of personal recognition on the attention of consumers of mobile banking is equal to 0.611; the stage of personal recognition on the interest of consumers of

mobile banking is equal to 0.532. The stage of personal recognition on the desire of mobile phone banking consumers is 0.924. The stage of personal recognition on the action of mobile banking consumers is 0.93. These coefficients are confirmed, according to the influence coefficient obtained. It shows that the higher the value of the relations of the mechanism of variables expressed in the organization is higher than 0.4, the better the quality of acceptance of banking with Bank Melli of Qazvin province.

According to Figure 5, it shows that the stated assumptions, the coefficient of significance is as follows: between two variables, the performance of social media on the stage of individual recognition of mobile banking consumers is equal to 3.789. The performance of social media on the use from the performance of mobile banking users, it is equal to 3.444. The performance of social media on informing mobile banking users is equal to 12.315, and the performance of social media on solving the problems of mobile banking users is equal to 10.474. The performance of social media on changing the use of mobile banking consumers is equal to 8.464, the level of personal recognition on the attention of mobile banking consumers is equal to 7.444. The level of personal recognition on the interest of mobile banking consumers is equal to 256/5, there is a significant relationship between the level of personal recognition on the desire of mobile phone banking consumers equal to 9.729, the level of personal knowledge on the action of mobile phone banking consumers equal to 2.77. These coefficients are confirmed; considering the effect coefficient obtained shows that the higher the value of the relations of the mechanism of the variables expressed in

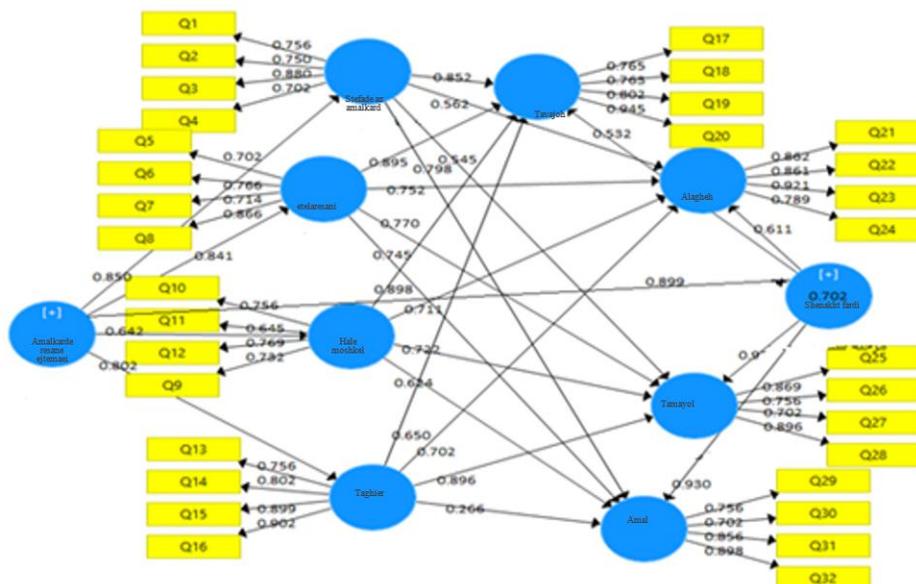


Figure 3. Coefficient of factor analysis of the role of social media in the acceptance of banking with the Bank Melli of Qazvin province

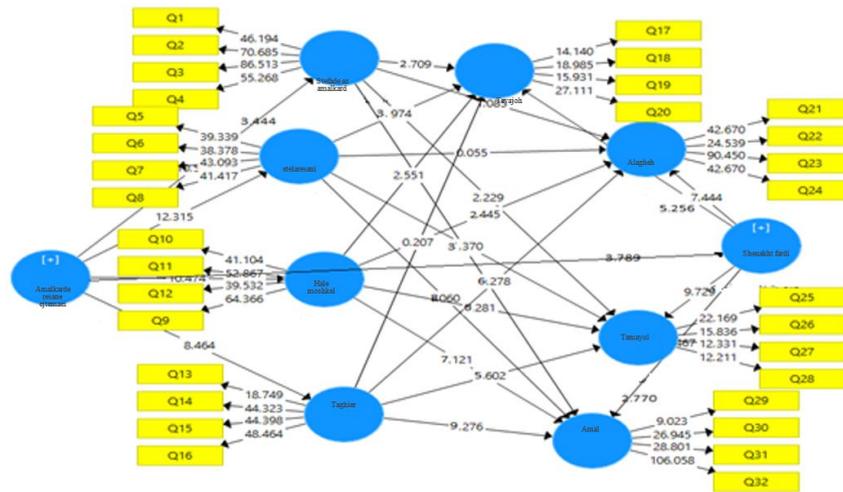


Figure 4. The t-value coefficient of the analysis of the role of social media in the acceptance of banking with the Bank Melli of Qazvin province

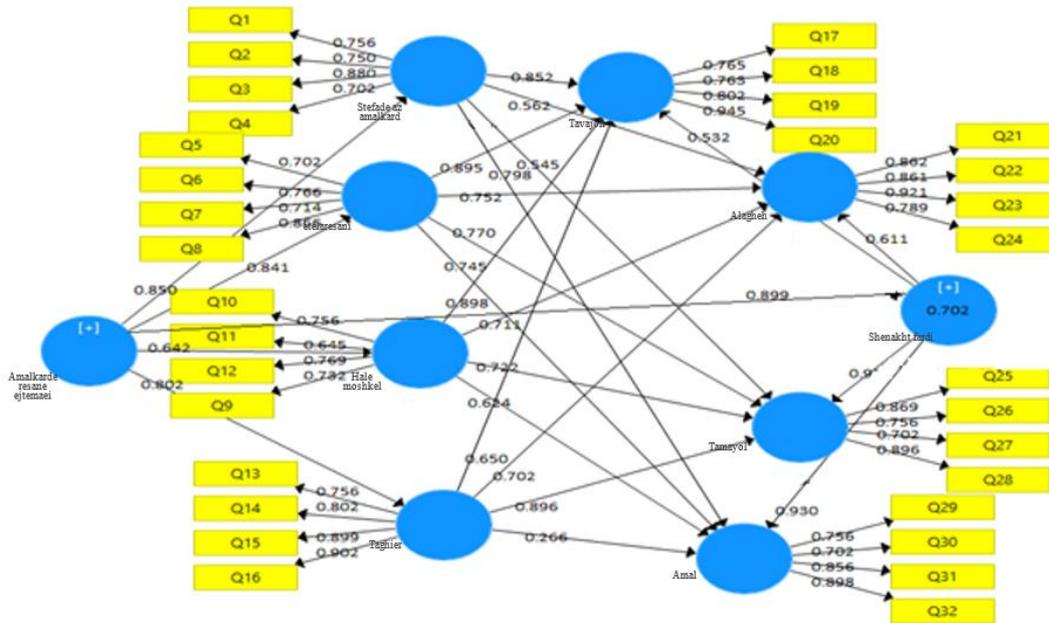


Figure 5. Factor loading coefficient of the analysis of the role of social media in the acceptance of banking with the Bank Melli of Qazvin province

the organization is higher than 1.96, the better the quality of acceptance of banking with Bank Melli of Qazvin province.

It is a measure that is used to connect the measurement part and the structural part of structural equation modeling and it shows the effect that an exogenous variable has on an endogenous variable. The essential point here is that R^2 is calculated only for the endogenous (dependent) structures of the model, and in the case of exogenous structures, the value of this criterion is zero. China, 1998, three values of 0.67-0.33-

0.19 are determined as weak, medium and strong model, which we have a dependent variable in this research is the level of individual recognition with a coefficient of 0.702.

4. 2. Redundancy Criterion

This criterion is obtained from the product of communality values of the structures and their corresponding R^2 values indicates the amount of variability of the indices of an endogenous structure that is affected by one or more exogenous structures (see Table 3).

TABLE 3. Redundancy criterion

Factors	Communality
Use the function	0.415
Notices	0.432
Solve the problem	0.603
change	0.745
Attention	0.702
interest	0.845
desire	0.750
Action	0.657

4. 3. Overall Fit of the Model (GOF Criterion)

This criterion is related to the general part of structural equation models. This means that by this criterion, the researcher can control the fit of the overall part after checking the fit of the measurement part and the structural part of the overall research model. The

goodness-of-fit (GOF) criterion is calculated according to the following formula:

$$GOF = \sqrt{Communalities \times R^2} \tag{1}$$

Mohammadi and Hamidi [18] have introduced three values of 0.01, 0.025-0.36 as weak, medium and strong values for GOF.

$$Communalities = \frac{(0.415 + 0.432 + 0.603 + 0.702 + 0.845 + 0.750 + 0.657)}{8} = 0.526$$

Individual recognition stage $R^2 = 0.702$

$$(General Model Fit) : GOF = \sqrt{0.526 \times 0.702} = 0.826$$

And the obtained GOF value of the project's success indicates the strength and correctness of the structural model and measurement in confirming the assumptions. As you can see in Table 4, the interpretation of the relationship and the degree of confirmation of the assumptions has been discussed:

TABLE 4. Summary of the results of confirmatory factor analysis (structural model) of the used scale

Research assumptions	Meaningful	Operational burden	Conclusion
The performance of social media has an effect on the stage of personal recognition of mobile banking consumers	3.789	899.0	Accepted
The performance of social media has an effect on the use of mobile banking by consumers.	4443	8500	Accepted
The performance of social media has an impact on informing mobile banking consumers.	315.12	841.0	Accepted
The performance of social media has an effect on solving the problem of mobile banking consumers.	474.10	0.642	Accepted
The performance of social media has an effect on the change in the use of mobile banking by consumers.	4648	0.802	Accepted
The stage of personal recognition has an effect on the attention of mobile banking consumers.	444.7	0.611	Accepted
The stage of personal recognition has an effect on consumers' interest in mobile banking.	2565	0.532	Accepted
The stage of personal recognition has an effect on consumers' willingness to use mobile banking.	729.9	0.924	Accepted
The stage of personal recognition has an effect on the action of mobile banking consumers.	7702	0.930	Accepted

5. CONCLUSION

Considering that the compilation and design of the planning document of new technologies is considered as one of the upstream documents of an organization, it is necessary to realize this important in the first stage as one of the goals of the Central Bank of the country. It should be considered.

Considering that information management technologies such as organization resource planning system are new in the country and by nature in government organizations and one of the most important

damages of information technology at the level of government institutions is the lack of transparency and The maturity of the legal and legal policies, and on the other hand, the trained personality of the organization's experts is the result of the legal and regulatory citation mission, the importance of informing and clarifying the existing laws and regulations of the country in the field of information management, has a great effect. It has a high level of acceptance and the threshold of organizational tolerance, and it is necessary for the legal committee of the organization's resource planning project to collect laws, regulations, approvals, instructions,

protocols, by-laws, etc. form and act on the planned injection and expertise of the mentioned cases in the layer of organizational ideas. It is necessary to mention that the legal notification should be done on time and according to the progress phases of the project in order to achieve the necessary impact.

One of the influential damages in the way of realizing the organization's resource planning system in the Central Bank of Iran was the concern of security risks. Unfortunately, the cognitive weakness of the user layers of the organization regarding the information security standards in the physical and digital space and the mental and emotional crisis caused by the lack of knowledge of the mentioned subject have had a great impact on the decision-making of this organization and it is necessary for the information protection team. By using experienced specialists, the organization should inform and culture the information security standards, especially the ISO 27001 standard, at the level of different organizational layers.

Unfortunately, the country's banking education system, like many government organizations, pursues statistical goals and increasing the welfare level of employees more than producing intellectual resources needed by the organization. Therefore, it is necessary to plan and implement organization trainings at both technical and non-technical levels. At the technical level, how and at the non-technical level, the why of using an educational and information technology is discussed. Considering the existence of high intellectual potentials in the central bank, providing practical and not theoretical standards in educational activities can have a tremendous effect on the mobility of different layers of the organization.

Considering the nature of risk-taking of organization's resource planning projects and the need to increase the courage of managers and reduce natural stress caused by new ideas, it is necessary to teach the standards of risk management, value management, stress management, conflict management and conflict management. organizations at the level of planning and operational managers.

Considering that the expert opinions of the organization do not accept the focus on a specific issue due to the high diversity of assigned missions, it is necessary to have a committee in parallel with the design and implementation team of the organization's resource planning in order to provide continuous information and Step-by-step clarification of operational steps, with the approach of information bombardment, will induce the formation and appropriate mentality in order to support and accompany different organizational layers.

The resource planning system of the organization in the central bank of the whole country is more than the production of tools, it is the amount of thought and the optimal use of thoughts. But it is necessary, considering

the wide spectrum of active experts in this organization, the organization's proposal system with the aim of collecting points of view, experiences and pure ideas existing in experienced and matured organizational thoughts, planning, launching and start working

6. REFERENCES

1. Merryta D, Tuga M. Artificial Intelligence Model as an Early Warning System for Fraudulent Transactions in Mobile Banking. *ICIC Express Letters Part B: Applications*. 2023;14(07):747. <https://doi.org/10.24507/icicelb.14.07.747>
2. Madhuri TS, Babu ER, Uma B, Lakshmi BM. Big-data driven approaches in materials science for real-time detection and prevention of fraud. *Materials Today: Proceedings*. 2023;81:969-76. <https://doi.org/10.1016/j.matpr.2021.04.3232214-7853>
3. Mytnyk B, Tkachyk O, Shakhovska N, Fedushko S, Syerov Y. Application of artificial intelligence for fraudulent banking operations recognition. *Big Data and Cognitive Computing*. 2023;7(2):93. <https://doi.org/10.3390/bdcc7020093>
4. Hamidi H, Vafaei A, Monadjemi SA. Evaluation and checkpointing of fault tolerant mobile agents execution in distributed systems. *Journal of Networks*. 2010;5(7):800. 10.4304/jnw.5.7.800-807
5. Seeja K, Zareapoor M. Fraudminer: A novel credit card fraud detection model based on frequent itemset mining. *The Scientific World Journal*. 2014;2014. <http://dx.doi.org/10.1155/2014/252797>
6. Padhi B, Chakravarty S, Biswal B, editors. Anonymized credit card transaction using machine learning techniques. *Advances in Intelligent Computing and Communication: Proceedings of ICAC 2019; 2020: Springer*.
7. Moradi R, Hamidi H. A New Mechanism for Detecting Shilling Attacks in Recommender Systems Based on Social Network Analysis and Gaussian Rough Neural Network with Emotional Learning. *International Journal of Engineering, Transactions B: Application*. 2023;36(2):321-34. <https://doi.org/10.5829/ije.2023.36.02b.12>
8. Sharma D, Kang SS, editors. Hybrid model for detection of frauds in credit cards. *2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N); 2022: IEEE*.
9. Zhang Y, Chen X, Li J, Wong DS, Li H, You I. Ensuring attribute privacy protection and fast decryption for outsourced data security in mobile cloud computing. *Information Sciences*. 2017;379:42-61. <https://doi.org/10.1016/j.ins.2016.04.015>
10. Nilchi AN, Vafaei A, Hamidi H, editors. Evaluation of security and fault tolerance in mobile agents. *2008 5th IFIP International Conference on Wireless and Optical Communications Networks (WOCN'08); 2008: IEEE*. 10.1109/WOCN.2008.4542509
11. Hashemi SK, Mirtaheri SL, Greco S. Fraud detection in banking data by machine learning techniques. *IEEE Access*. 2022;11:3034-43. 10.1109/ACCESS.2022.3232287
12. Nkomo BK, Breetzke T, editors. A conceptual model for the use of artificial intelligence for credit card fraud detection in banks. *2020 Conference on Information Communications Technology and Society (ICTAS); 2020: IEEE*. 10.1109/ICTAS47918.2020.233980
13. Hamidi H, Mohammadi K. Modeling fault tolerant and secure mobile agent execution in distributed systems. *International Journal of Intelligent Information Technologies (JIIT)*. 2006;2(1):21-36. 10.4018/jiit.2006010102

14. Aljawameh S. Cloud security engineering concept and vision: Concept and vision. Cyber security and threats: Concepts, methodologies, tools, and applications: IGI Global; 2018. p. 93-101.
15. Khan N, Al-Yasiri A. Cloud security threats and techniques to strengthen cloud computing adoption framework. Cyber security and threats: Concepts, methodologies, tools, and applications: IGI Global; 2018. p. 268-85.
16. Hamidi H, Vafaei A. Evaluation of fault tolerant mobile agents in distributed systems. International Journal of Intelligent Information Technologies (IJIT). 2009;5(1):43-60. <http://doi.org/10.4018/jiit.2009010103>
17. Balagolla E, Fernando W, Rathnayake R, Wijesekera M, Senarathne A, Abeywardhana K, editors. Credit card fraud prevention using blockchain. 2021 6th international conference for Convergence in Technology (I2CT); 2021: IEEE. 10.1109/I2CT51068.2021.9418192
18. Mohammadi K, Hamidi H, editors. Evaluation of fault-tolerant mobile agents in distributed systems. 2005 1st IEEE and IFIP International Conference in Central Asia on Internet; 2005: IEEE. 10.1109/CANET.2005.1598208
19. Kirar JS, Kumar D, Chatterjee D, Patel PS, Yadav SN, editors. Exploratory Data Analysis for Credit Card Fraud Detection. 2021 International Conference on Computational Performance Evaluation (ComPE); 2021: IEEE. 10.1109/ComPE53109.2021.9751922
20. Hamidi H, Vafaei A, Monadjemi A. Algorithm based fault tolerant and check pointing for high performance computing systems. Journal of applied sciences. 2009;9(22):3947-56. 10.3923/jas.2009.3947.3956
21. Hamidi H, Rafebakhsh M. Analyzing factors influencing mobile social media marketing acceptance among customers. International Journal of Engineering, Transactions C: Aspects. 2022;35(6):1209-16. <https://doi.org/10.5829/ije.2022.35.06c.13>
22. Hamidi H, Seyed Lotfali S. Analysis of role of cloud computing in providing internet banking services: Case study bank melli iran. International Journal of Engineering, Transactions B: Application. 2022;35(5):1082-8. <https://doi.org/10.5829/ije.2022.35.05b.23>
23. Daliri S. Using harmony search algorithm in neural networks to improve fraud detection in banking system. Computational Intelligence and Neuroscience. 2020;2020. <https://doi.org/10.1155/2020/6503459>
24. Esmail FS, Alsheref FK, Aboutabl AE. Review of Loan Fraud Detection Process in the Banking Sector Using Data Mining Techniques. International journal of electrical and computer engineering systems. 2023;14(2):229-39. <https://doi.org/10.32985/ijeces.14.2.12>
25. Handa A, Dhawan Y, Semwal P. Hybrid analysis on credit card fraud detection using machine learning techniques. Handbook of Big Data Analytics and Forensics. 2022:223-38. https://doi.org/10.1007/978-3-030-74753-4_15
26. Hamidi H. A combined fuzzy method for evaluating criteria in enterprise resource planning implementation. Intelligent systems: Concepts, methodologies, tools, and applications: IGI Global; 2018. p. 639-70.
27. Alt R, Puschmann T. The rise of customer-oriented banking-electronic markets are paving the way for change in the financial industry. Electronic Markets. 2012;22:203-15.
28. Patil S, Nemade V, Soni PK. Predictive modelling for credit card fraud detection using data analytics. Procedia computer science. 2018;132:385-95. <https://doi.org/10.1016/j.procs.2018.05.199>
29. Hamidi H, Vafaei A, Monadjemi SA. Analysis and design of an abft and parity-checking technique in high performance computing systems. Journal of Circuits, Systems, and Computers. 2012;21(03):1250017. <https://doi.org/10.1142/S021812661250017X>
30. Rambola R, Varshney P, Vishwakarma P, editors. Data mining techniques for fraud detection in banking sector. 2018 4th International Conference on Computing Communication and Automation (ICCCA); 2018: IEEE. 10.1109/CCAA.2018.8777535

COPYRIGHTS

©2024 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

با استفاده از رسانه های اجتماعی، بانک ها می توانند راه های کارآمدتری برای جذب مشتری طراحی کنند. رسانه های اجتماعی و اپلیکیشن هایی که از طریق رسانه های تلفن همراه در دسترس مردم هستند، می توانند به بانک ها کمک کنند تا محصولات و خدمات خود را در مناطق جغرافیایی که نتوانسته اند حضور فیزیکی داشته باشند، نفوذ کنند و آن ها را معرفی کنند. بانک ها با نظارت بر اطلاعات به دست آمده از رسانه های اجتماعی، می توانند واکنش های مشتریان به محصولات ارائه شده توسط رقبا را تجزیه و تحلیل کنند و استراتژی های مناسبی برای افزایش رضایت مشتری و جذب مشتریان جدید اتخاذ کنند. این پژوهش با هدف تحلیل نقش رسانه های اجتماعی در پذیرش موبایل بانک در بانک ملی استان قزوین انجام شد. این پژوهش از نوع پیمایشی و کاربردی است. ابزار گردآوری داده ها پیمایشی و روش گردآوری داده ها مطالعه میدانی است. جامعه آماری این پژوهش کلیه مدیران و دستیاران بانک ملی استان قزوین در رده های مختلف می باشد. برای محاسبه حجم نمونه آماری به دلیل محدود بودن جامعه آماری از روش فرمول کوکران استفاده شده است. در این تحقیق ابتدا شاخص ها شناسایی و بر اساس این شاخص ها فرضیه هایی تدوین شد و در نهایت با استفاده از تکنیک های آماری، تمامی فرضیه های مطرح شده تا آنها اثبات شد و تأثیر همه شاخص ها به اثبات رسید. نتایج بدست آمده تأثیر عملکرد رسانه های اجتماعی بر مرحله شناخت فردی مصرف کنندگان بانکداری تلفن همراه را تأیید کرده است.



Scientific and Methodological Support of Sand Management During Operation of Horizontal Wells

D. S. Tananykhin*

Petroleum Department, Empress Catherine II Saint Petersburg Mining University, Saint Petersburg, Russian Federation

PAPER INFO

Paper history:

Received 01 December 2023

Received in revised form 23 December 2023

Accepted 04 January 2024

Keywords:

Sand Production

Physical Simulating

Mathematical Modeling

Shape Memory Polymers

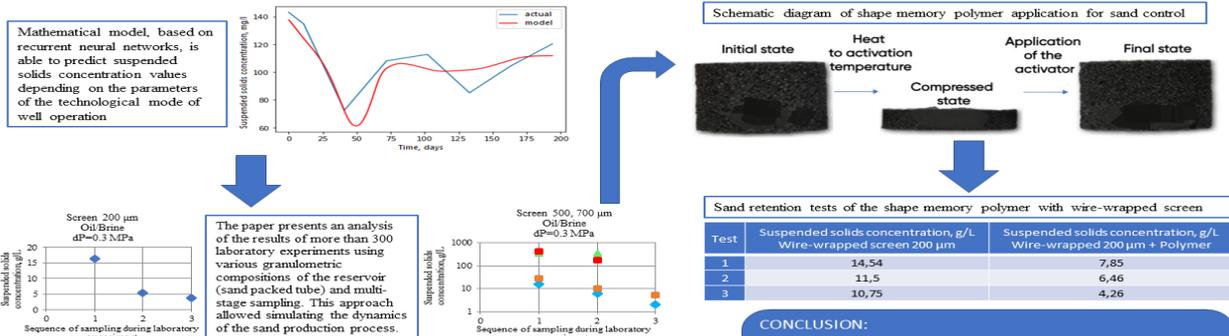
A B S T R A C T

The paper presents a new scientifically based approach for the selection of sand control technology, which is dedicated to enhance the efficiency of the development of unconsolidated reservoirs. Established laboratory and methodological complexes for physical simulating of the sand producing process were analyzed in order to obtain new knowledge and confirm the available theories. All of them have their advantages and disadvantages, but their simultaneous application revealed characteristic dependencies between the sand production and the studied parameter (grain size distribution, pressure drop, clay content, water cut, gas/oil ratio, etc.). Author proposed concepts of mathematical apparatus improvement to increase the quality of assessing the ability of formation fluids to transport particles of different grain size distribution within the formation, as well as in the inner part of tubing. The effect of each of the characterizing factors on suspended solids concentration (SSC) was studied as a result of more than 300 laboratory experiments. According to the observation, there is a sharp decrease in SSC after the first stage (sampling). Thus, the author determined that the main inflow of mechanical impurities occurs during flow stimulation and after shutdowns. In conclusion, author substantiated the method for limiting sand production using polymers with shape memory based on the results of the performed set of tests. Proposed method allows limited passage of particles with diameter less than 50 μm , which creates conditions for noncolmaticity of screen while maintaining geomechanical stability of bottom-hole formation zone.

doi: 10.5829/ije.2024.37.07a.17

Graphical Abstract

Increasing the efficiency of the development of unconsolidated reservoirs through the use of a new scientifically based approach for the selection of sand control technology



Empress Catherine II Saint Petersburg Mining University

Tananykhin D.S. Scientific and methodological support of sand management during the operation of horizontal wells. International Journal of Engineering, Transactions B: Applications, 34(2),pp. 581-586, 2021. doi:10.5829/IJE.2021.34.02B.32

*Corresponding Author Email: Tananykhin_DS@pers.spmi.ru (D. S. Tananykhin)

Please cite this article as: Tananykhin DS. Scientific and Methodological Support of Sand Management During Operation of Horizontal Wells. International Journal of Engineering, Transactions A: Basics. 2024;37(07):1395-407.

NOMENCLATURE			
u_c^*	critical velocity of the single-phase flow (ft/s)	V_{sl}	reduced particle velocity
u_0^*	friction velocity, (ft/s)	D	pipe diameter
u_t	rate of sedimentation, (m/s)	x	the energy required to maintain suspended particles per unit length of the pipe
C_V	solids concentration, (m^3/m^3)	N_{Re}	Reynolds number
ν	kinematic viscosity, (ft/s)	ρ_G	gas specific gravity
d	grain diameter	V_M	mixture velocity
$V_{L,C}$	critical velocity of the fluid, m/c	V_p	particle velocity
g	gravity	Fr	Froude coefficient
ρ_P	particle density	WC	water cut
ρ_L	fluid density	GOR	gas-oil ratio

1. INTRODUCTION

Sand production occurs in a wide range of hydrocarbon reservoirs, both composed of relatively young rocks of the Quaternary period and consolidated reservoirs at the final stages of development (1). Sand-related challenges significantly disrupt the oil production process, since it adversely affects the complex of subsurface and surface equipment and creates difficulties in the operation of wells (2, 3). The topic of horizontal wells operation with sand production is particularly relevant. Their application is justified by much higher well productivity (compared to vertical ones) and a large sweep efficiency.

Sand production is also a relevant scientific challenge. This is primarily due to the need for inherent scientific research and the development of a systematic approach, which consists in the reasonable choice of the most effective technologies (4, 5), as well as in the integrated simulating (numerical, physical and field) of processes occurring in the system «well-bottom hole zone».

Based on the results of laboratory tests and mathematical modelling (6-8), it was revealed that the greatest sand production occurs at the stage of flow stimulation due to violation of the initial geomechanical state of the bottomhole formation zone.

Sand production depends on the water/oil/gas distribution in the total flow of the fluid (9, 10). Some researchers argue that this is solely a problem of arch stability (11), whilst others argue that this is an issue of wetting and chemical reactions (12) or strain softening due to water incorporating into the rock matrix (13).

There are many technologies and technical solutions to prevent sand production, such as screens, different form of gravel packings, chemical consolidation, technological restrictions and some other forms or combinations of them. Every technology has both advantages and disadvantages. Screens (wire-wrapped, weaved, slotted liners et al.) are prone to erosion (13-15) and corrosion (16, 17). Slotted liners are usually restricted on the open to flow area (OFA), because high OFA results in decreasing strength of the whole screen (which is basically a tubing with slots (cuts) of specific pattern and opening size) (18). Another problem is screen replacement – since borehole walls can collapse onto the

screen surface, this causes an issue of lifting the whole tailpiece onto the surface.

Chemical consolidation technologies usually result in low retained permeability of the reservoir after treatment (17) and their efficiency largely depends on the clay content and water saturation of the reservoir (18). Chemical consolidation also requires treatments on the periodic manner between 0.5 and 1.5 years (19). But none of the above solves a root cause of sand production – geomechanical instability.

Gravel packs are able to solve a significant amount of complications caused by the geomechanical factor. However, in the case of horizontal completions, the use of gravel packs is limited due to the large volumes of borehole space that need to be filled with gravel. The use of unsorted (cheaper) gravel can lead to increased production of rock particles in certain packing areas under conditions of high heterogeneity. Unsorted gravel also favours conditions of “hot-spots” (zones with lack of gravel or extremely high permeability), which will result in screen losing its efficiency in terms of sand retaining and will further favour faster terms of erosion (20, 21). The use of sorted gravel is often not possible due to its high cost.

This paper presents a method of limiting sand production by utilizing shape memory polymers which have a mesh structure and can allow the passage of particles with a diameter of less than 50 μm , which leads to less "blockage" of the screen.

2. MATERIALS AND METHODS

Traditionally, the criteria for selecting sand control technologies is based either on empirical correlations or on field experience in the operation of similar objects (22-24). Nowadays, there has been a growing interest in using a laboratory infrastructure to investigate the technologies of operating wells with different types of completion. The most common methods of testing sand-related challenges are: slurry sand retention test (25-28) and prepack sand retention test (26, 29-35). A comparative analysis of current laboratory procedures is presented in Table 1.

2. 1. Mathematical Models of Critical Flow Velocity in Horizontal Wells

There are two main approaches of models for determining critical flow velocity in a horizontal well. The first approach is based on the description of the forces and processes that affect the solid particle: its picking up, transfer and deposition. This approach is reflected in the theory of Salama (36).

The second approach is based on the theory of turbulence: the balance between the energy required to weigh particles and the energy released when the turbulent vortices are stratified. Models Oroskari and Turian (37), Danielson (38) and others (39-41) are the most widespread.

However, the hydrodynamics of a multiphase system is much more complex than a single-phase one. The flow tends to separate due to the different density and viscosity

of the liquid and gas fluid (42). Various structural forms (43) can characterize the gas-liquid flow. Gas-liquid flow's structure refers to mutual arrangement or distribution of gas and liquid phases in process of their simultaneous movement through the well (44).

Different phases move at different velocity. It leads to a very important phenomenon - the "slippage" of one phase relative to others. Comparative analysis of mathematical models for calculation of critical velocity is given in Table 2 (45, 46).

2. 2. Advanced Software for Calculating Sand Production During Operation of Oil And Gas Wells

Author identifies the following software programs that allow calculating critical velocity of produced solid particles in horizontal multiphase flow:

TABLE 1. Comparative analysis of laboratory methods for testing sand-related challenges during the development of oil and gas fields

Method	Measurement parameters	Simulated conditions	Testing sample
Prepack sand retention test	1. Drop pressure during filtration	Long-term well operation, wellbore collapse	Core sample - sand packed tube /mechanical screen, gravel pack or chemical treatment
Slurry sand retention test	2. Suspended solids concentration	Flow stimulation, first inflow of the well	Not applicable/ mechanical screen, gravel pack
Linear sand control evaluation	3. Particle size distribution of produced particles	Long-term well operation, wellbore collapse	Sand packed tube / mechanical screen, gravel pack, chemical treatment
Radial sand control evaluation	4. Permeability	Long-term well operation, wellbore collapse	Sand packed tube /Full size mechanical screen, gravel pack, chemical treatment

TABLE 2. Comparative analysis of mathematical models for calculation of critical velocity in horizontal well

No.	Author	Particle diameter, μm	SSC, mg/l	Mathematical correlation
Single phase flow				
1	Thomas (44)	0.4-950	More than 20000	$u_c^* = u_0^* \left[1 + 2.8 \left(\frac{u_t}{u_0^*} \right)^{0.33} \sqrt{C_V} \right]$ $u_0^* = \left(100 u_t \left(\frac{v}{d} \right)^{2.71} \right)^{0.269}$
2	Oroskar and Turian (38)	-	-	$\frac{V_{LC}}{\sqrt{gd \left(\frac{\rho_P}{\rho_L} - 1 \right)}} = 1.85 C_V^{0.1536} (1 - C_V)^{0.3564} \left(\frac{d}{D} \right)^{-0.378} N_{Re}^{0.09} \chi^{0.30}$
Multiphase flow				
3	Salama (37)	-	-	$V = 1.3 \left(\frac{V_{sl}}{V_M} \right)^{0.53} d^{0.17} v_L^{-0.09} \left(\frac{\rho_P - \rho_L}{\rho_L} \right)^{0.55} D^{0.47}$
4	Stevenson(45)	512-1010	2 000	$\frac{V_p}{V_{sl}} = 0.95 \left(1 + \frac{V_{SG}}{V_{SL}} \right) - \left(1.38 \frac{V_{SG}}{V_{SL}} + 0.88 \sqrt{Fr} \right) \cdot \left(Re \sqrt{Fr} \left(\frac{d}{D} \right)^{1.5} \right)^{-0.18}$
5	Danielson(39)	280-550	1 000	$V_{LC} = 0.23 v^{-1/9} d^{1/9} (gD((\rho_P - \rho_L) - 1))^{5/9}$
6	Ibarra (46)	211-297	2 500 – 10 000	$\frac{V_{LC}}{\sqrt{gd \left(\frac{\rho_P}{\rho_L} - 1 \right)}} = 1.3277 \left(\frac{V_{SL}}{V_{LC}} \right)^{-0.285} (1 - C_V)^{-35.490} \left(\frac{d}{D} \right)^{-0.378} N_{Re}^{0.09}$
7	Hill (47)	211-297	2 500 – 10 000	$\frac{V_{LC}}{\sqrt{gd \left(\frac{\rho_P}{\rho_L} - 1 \right)}} = 1.85 C_V^{0.1536} (1 - C_V)^{0.3564} \left(\frac{d}{D} \right)^{0.378} N_{Re}^{0.09}$

- OLGA Schlumberger (software was used for investigations (47, 48).
- ANSYS Fluent (software was used for calculations (49).

The OLGA dynamic multiphase flow allows you to perform calculations for a complex modeling of the behavior of solid phases, calculate the change in the flow rate in general and each phase separately, and calculate the precipitation of solid particles.

Main advantages are a quick calculation time and a focus on the oil and gas industry. The main drawback is the complexity of setting up and using the program.

As for ANSYS Fluent software, the combination of DEM (discrete element method) and CFD (computational fluid dynamics) packages allows you to simulate a gas/liquid-solids system. The program allows you to take into account both the interaction between particles and a high concentration of particles. Modeling of solid particles is possible in the DEM-CFD Coupling Module, where it is possible to specify the number of suspended particles, their size and density. The advantage of CFD 3D programs is the high accuracy of calculations, taking into account different factors that affect the transport of formation particles. However, an increase in accuracy leads to an increase in calculation time. Comparative analysis of the sand simulating software is provided in Table 3.

One of the software allows us to simulate the transport of solid particle and erosion process of oil and gas equipment (50). PETRONAS software is more designed to determine the risk of equipment erosion, but it can calculate the critical flow velocity to prevent the formation of a sand plug in the well.

SYNTEF's Multiphase Transport and Flow Assurance software, which has a module for simulating the transport of solid particles in horizontal wells and pipelines, is on the global market too.

As a result of the analysis, the author registered a software program that allows calculating the size of particles that can be produced from a horizontal well by upflow.

TABLE 3. Software products related to sand production simulation

Module	Software product	Advantages	Disadvantages
Sand transport in the well	OLGA	- Quick calculation time; - Adapted for oil and gas industry;	- Complexity of setting up
	ANSYS Fluent	- High accuracy	- Not adapted for oil production; - Complexity of setting up

2. 3. Creating a Predictive Mathematical Model Using Recurrent Neural Networks

The algorithm of the mathematical model consists of using the long-term short-term memory (LSTM) model, which is a type of recurrent neural network. Each row of input data about the target parameter and characteristics has its own time - the measurement date. Therefore, the dependence of suspended solids concentration on the parameters of the process mode of the well was considered in the form of a time series (51-53).

The proposed model consists of 6 layers (Figure 1): input layer, two layers of long-term short-term memory, two layers of regularizers and output layer. The mean-square error is used as a function of calculating deviation, and the Adam algorithm (adaptive moment estimate) is used as an optimizer. The data entered into the model were pre-standardized and divided by proportion: 70% - training, 10% - validation, 20% - test.

The training process was conducted on several models to obtain reliable results. Models differ in the parameter of the lines from each other. Then best model has been selected according to the information obtained after testing.

The results of using the developed mathematical model are presented below.

3. RESULTS

This paper presents an analysis of the results of more than 300 laboratory experiments using various granulometric compositions of the reservoir (sand packed tube) and multi-stage sampling. This approach allowed simulating the dynamics of the sand production process.

The author found that the distribution of phases in the fluid flow affects the amount of suspended particles in the filtrate when using the same screens. An increase in the share of brine leads to an increase in suspended solids concentration. The SSC of any brine-oil mixture is lower than the SSC values in pure oil filtration. An increase in the gas fraction, on the contrary, leads to a decrease in SSC (Figure 2).

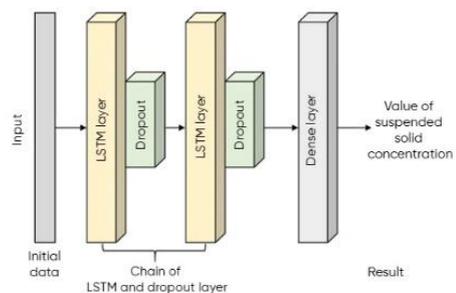


Figure 1. Schematic diagram of the proposed model

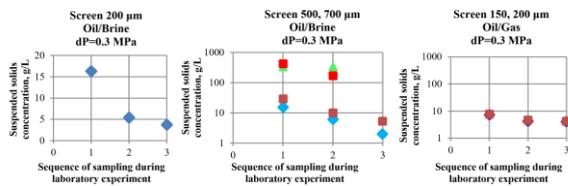


Figure 2. Amount of suspended solids concentration depending on the filterable phases

At the same time, the SSC naturally decreases, primarily due to the formation of arched systems near the holes of the screen (14, 32, 54-58). The scheme of the arch system is shown in Figure 3.

An increase in the pressure drop during filtration of various mixtures, on average, leads to an increase in SSC by 2.6 times with a range of 1.2-8 times depending on the sampling stage (Figure 4).

The results of physical experiments showed the greatest impact of the flow stimulation stage on the stability of the bottomhole formation zone and on the amount of sand produced.

3. 1. Assessment of the Ability of Fluids to Transport Formation Particles of Different Grain Size Distribution

One of the goal of the paper is to determine the critical flow velocity. Existing mathematical models and correlations on the calculation of critical flow velocity have been discussed in detail above. All correlations are based on laboratory tests performed for the flow of mechanical impurities in brine.

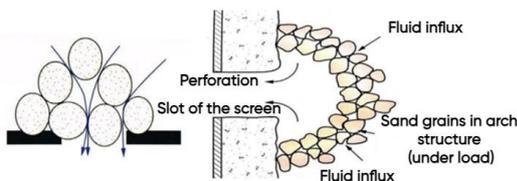


Figure 3. Arch system located near perforations/screen slots (2)

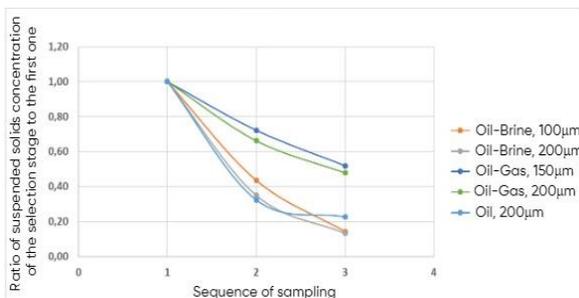


Figure 4. Stabilization of sand production during laboratory experiment

The physical characteristics of brine differ significantly from those of oil, especially high viscosity oil.

Moreover, many of the models considered are adapted only for suspensions with high concentrations of suspended particles, for example, pulp.

The author conducted a parametric study for particles of different diameters with different water cut and different gas factor. The critical velocity was determined graphoanalytically using graphs constructed as a result of mathematical modeling (Figure 5). “Black line” indicates a line showing the volume fraction of particles deposited on the bottom of the pipe; “red line” indicates the flow rate. As can be seen from the graph, when the flow rate decreases to 0.149m/s, particles with a diameter of 300 μm with a water cut of 20 % will begin to form a fixed layer.

The critical flow rate varies differently for mixtures with different proportions of water (Figure 6). The obtained results of numerical modeling are in line with the generally accepted opinion that the growth of the linear particle size leads to an increase in the critical flow rate:

1. When pure oil flows (water cut 0 %), an increase in the linear grain size leads to an increase in the critical flow rate, which is consistent with the generally accepted ideas about the transport of mechanical particles.
2. For a mixture with a low water content (water cut 5 %, 10 %, 20 %), the critical velocity does not depend on the particle diameter.
3. For a highly aqueous liquid (water cut 50 %, 70 %), the critical velocity does not change for particles in the diameter range of 100-1000 μm. At the same time, the value of the critical velocity for small particles (50 μm) is higher than for larger ones.

Thus, with a water cut of 50-70 %, a layer of water is formed in the lower part of the pipe, the low viscosity of which does not allow to "pick up" and ensure the transfer of particles with a size of less than 50 μm.

As can be seen in Figures 6 and 7, an increase in the proportion of water in the flow leads to an increase in critical velocity, since water has less ability to carry particles in a horizontal part.

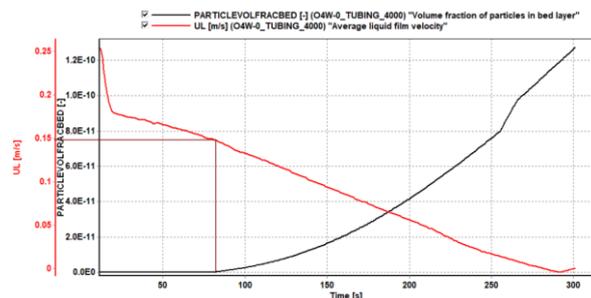


Figure 5. Critical velocity for sand grains of 600 μm with water cut 0%

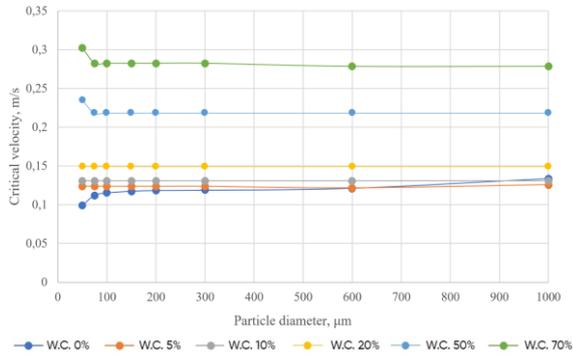


Figure 6. Critical velocity for different particle diameters and water cut (W.C.)

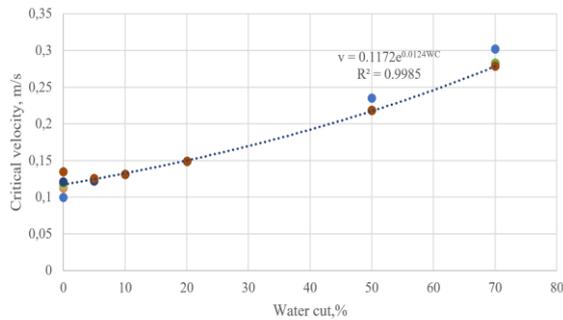


Figure 7. Dependence of critical velocity on water cut

Since critical velocity depends little on particle size, for future calculations it is proposed to use correlation 1 to calculate critical flow velocity depending on water cut (Figure 7). The coefficient of determination for the proposed expression is 0.9985.

$$v = 0,1172e^{0,0124WC} \tag{1}$$

where v – critical velocity, m/s; WC – water cut, %

Similarly, critical velocities for the gas-liquid mixture were determined (Figure 8). The increase in gas-oil ratio results in an increase in critical flow rate. At the same time, the velocity increases with an increase in the size of the solid particle.

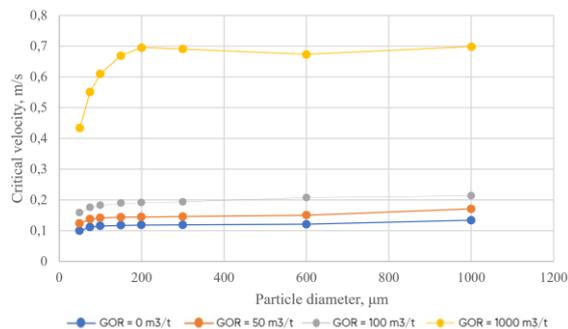


Figure 8. Critical velocity for different particle diameters and gas-oil ratio

Increasing the amount of gas reduces the viscosity and density of the oil, which reduces the drag force on the solid particle from the fluid surface. The ability of the fluid to allow the movement of sand is reduced.

3. 2. Results of Using Developed Mathematical Model Using Recurrent Neural Networks

The Mining University team initiated a program for creating a mathematical model for sand prediction, which takes into account quantitative and qualitative using recurrent neural networks based on field data on the technological modes of well operation. The model allows us to predict the trend and direction of change of the suspended solids concentration, as well as its absolute values depending on the main technological parameters. As a basis for modeling, a type of recurrent neural network (RNN) network of long short-term memory (LSTM) is used.

The proposed model consists of 6 layers: input layer, two LSTM layers, two layers of Dropout regularizers and output layer. The standard error (MSE) is used as a function for calculating deviation, and the Adam algorithm is used as an optimizer. The data submitted to the model are pre-standardized and divided by proportion: 70 % - training, 10 % - validation, 20 % - prediction. The average absolute error (MAE) acts as a measure of assessing the quality of model training. As a result, 3 out of 4 wells with a reach of more than 100 mg/l give a satisfactory result.

The applicability of the model is limited to the prediction of SSC values of 20 mg/l as a minimum and 1000 mg/l as a maximum. Values beyond these limits cannot be predicted by the model due to the peculiarities of the recurrent neural network architecture used.

Limitations in the initial data are conditioned by physical meanings of each of the indicators – there are no direct limitations of the model operation depending on the values of the initial data.

The input data fed into the model are:

1. P_{res} - minimum limitation is 1 bar, no maximum restriction.
2. $P_{at\ the\ intake}$ - minimum limit of 1 bar, no maximum restriction.
3. P_{wf} - minimum limit of 1 bar, no maximum limit, but does not exceed reservoir pressure.
4. H_D - minimum restriction of 0 m, but will not exceed well depth.
5. P_{buf} - minimum restriction of 1 bar, no maximum restriction.
6. Q_{oil} - minimum restriction of 0 m³/d, no maximum restriction.
7. Q_{liq} - minimum limit of 0 m³/d, but not less than Q_{oil} , no maximum restriction.
8. Water cut - minimum limitation of 0%, maximum limitation of 100%.

The created mathematical model, based on recurrent neural networks, is able to predict suspended solids

concentration values depending on the parameters of the technological mode of well operation. It can be used in the field to control the sand production, calculate the time of bottomhole filling, predict downhole pumping equipment failures, etc.

The results of using the developed mathematical model based on neural networks to determine the predicted indicators of the number of suspended particles in the well are presented in Figure 9. As the initial value (blue line), field data with technological modes of well operation were used.

Next step is to develop a software module to determine the scenario of filling the horizontal wellbore. Section of the horizontal well with a particle accumulation is determined by setting the initial flow rates, the well profile and grain size distribution (Figure 10).

According to the information received, it can be concluded that the greatest accumulation of mechanical particles occurs in the "toe" (the flow rate is not enough for the sand production) and in the "heel" (due to the inclination of the section) of the horizontal well.

3. 3. Justification of Sand Control Technology Using Shape Memory Polymers According to the results of analysis of more than 300 laboratory

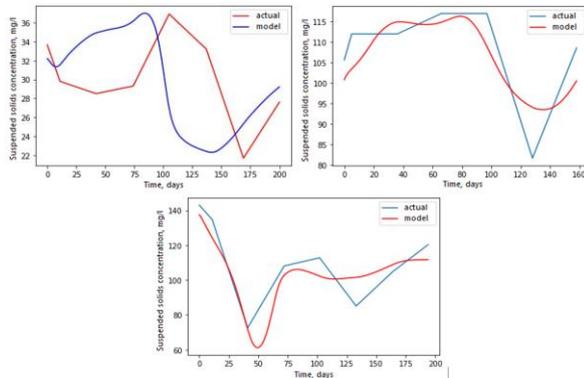


Figure 9. Results of using the developed mathematical model to predict sand production

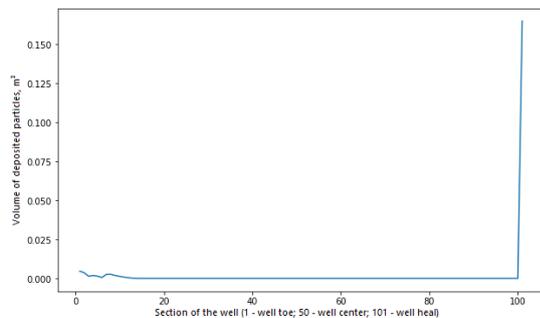


Figure 10. A number of suspended particles deposited depending on section of the horizontal well

experiments, it was established that the greatest sand production occurs at the stage of flow stimulation (i.e., the first tens to hundreds of tons of produced hydrocarbons). Therefore, it is necessary to implement sand control technologies already at the stage of flow stimulation. The use of mechanical control methods in the event of frequent well shutdowns is futile. In case of gas flow the sand production is almost not stabilized (it does not decrease during long-term filtration), which indicates the possibility of using the technology with shape memory polymer when producing gas-oil mixtures.

It is important to note that each shape memory polymer has its own activation temperature (the start point of the transition to the initial state). First, the shape memory polymer is heated to the activation temperature, then compressed to the minimum possible dimensions, after which it is cooled and the polymer in this state can be freely transported. The form recovery will begin after the polymer is reheated to its activation temperature. A screen with a polymer is installed at the bottom-hole zone (the temperature of which should be lower than the activation temperature) and activation fluids are used a catalyst that reduces the activation temperature of the polymer to a target level (Figure 11).

The major element of the shape memory polymer is the activator and its concentration in the liquid the lower the bottomhole temperature, the higher concentration of the activator is necessary to reduce the activation temperature. Almost all Russian sand-producing oil fields are confined to PK1-3 layers, which located at a depth of 1000-1200 meters, and therefore have a low bottomhole temperature (in the range of 10-30 °C). Accordingly, the use of this technology can be limited by the need to use a high concentration of activator (up to 10 % by volume) in the liquid, which reduces the economic attractiveness (high costs for polymer engineering and activator chemicals) of this method of sand control.

The sand-limiting method developed in this paper allows limited transport of particles with a diameter of less than 50 μm, which creates conditions for non-compatibility of the screen while maintaining its efficiency in terms of reducing the number of produced particles. Photo of the expandable screen element are shown in Figure 12.

The use of this polymer consists of two steps:

1) Delivery of slotted screen with pre-installed polymer pellets to bottom hole formation zone (screen

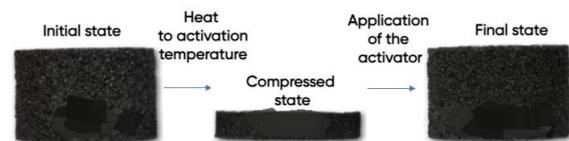


Figure 11. Schematic diagram of shape memory polymer application

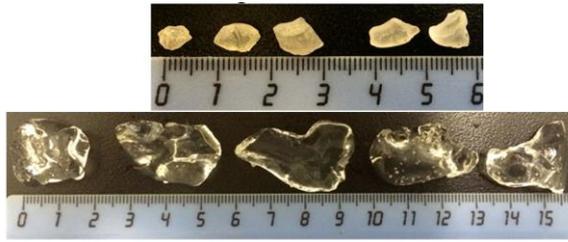


Figure 12. Photo of the expandable polymer screen in laboratory conditions

aperture does not allow polymer pellets to "fall out" during transportation).

2) Activation of the polymer in order to swell it and fill the entire volume between the lowered equipment and the wellbore by using the activator fluid.

Based on the results of laboratory studies of the developed technology using a polymer with shape memory, a decrease in the number of suspended particles by more than 46% was obtained (Table 4).

There are a slight decrease in the permeability of the polymer structure after the completion of laboratory experiments. The average permeability reduction is 6% (Table 5).

4. DISCUSSION

The selection of the technology for limiting sand production is based on the results of analysis of the following factors (59):

- Well profile: vertical, inclined or horizontal;

- Well completion: open or cased.
- Particle size analysis;
- Cost-effectiveness.

There are some tips how to make right selection of sand control technologies, for example, decision trees, based on the data presented in the scientific paper (60).

It is also necessary to understand the advantages and disadvantages of current methods. Most of them are described in Table 6.

TABLE 4. Laboratory tests of the shape memory polymer with wire-wrapped screen

Test	Suspended solids concentration, g/L	Suspended solids concentration, g/L
	Wire-wrapped screen 200 μm	Wire-wrapped 200 μm + Polymer
1	14,54	7,85
2	11,5	6,46
3	10,75	4,26

TABLE 5. Analysis of the permeability reduction of the shape memory polymer with wire-wrapped screen

Test	Initial data	Final data, % of original value
	Wire-wrapped screen 200 μm	Wire-wrapped screen 200 μm + Polymer
1	100	95%
2	100	96%
3	100	93%

TABLE 6. Advantages and disadvantages of sand control technologies

Technology	Advantages	Disadvantages
Pressure control	- Prevents formation destruction - no capital expenses	- underestimation of the well productivity
Selective perforation	- Prevents formation destruction	- underestimation of the well productivity - narrow applicability range - creation of a geomechanical reservoir model is required
Oriented perforation	- Prevents formation destruction	- underestimation of the well productivity - narrow applicability range - creation of a geomechanical reservoir model is required
Chemical consolidation	- Prevents formation destruction - cost-effective for small intervals (compared to the gravel packing) - compatible with mechanical methods of sand control	- reduction of permeability - cost-effective only for shallow thickness (up to 5 meters) - low efficiency of the repair and insulation works - service life is limited (1-2 years for epoxy resins)
Slotted screens	- Low cost - passage of medium and large size particles - easiness of installation - compatible with inflow control devices	- screen erosion - inefficient for passage small particles - not applicable in heterogeneous formations - low reliability

Wire-wrapped screens	- Compatible with inflow control devices - filtration area is larger than in slotted screens	- high cost - screen clogging - screen erosion - risk of screen damage during installation - high skin factor
Wire-mesh screens	- Compatible with inflow control devices - filtration area is larger than in slotted screens	- high cost - screen clogging - screen erosion - risk of screen damage during installation - high skin factor
Expandable screen	- Supports the wellbore; - low skin factor	- high cost - risk of screen damage during installation
Ceramic screens	- Less prone to erosion than metal screens	- high cost - screen clogging - risk of screen damage during installation
Wire-wrapped screens with gravel packed	- Thinnest passage zone - transport of different sizes particles can be achieved by controlling the packaging	- high cost - screen clogging - screen erosion - not applicable formations with a clay content
Gravel Pack in the open hole	- Applicable in long sections (up to 150 m) - filtration of small and medium-sized particles - durability - high reliability	- sharp decrease in productivity - complexity of screen creation process - high cost

Currently, single sand screens (slot or wire-wrapped) are the most widespread within the oil and gas fields in Russia.

Regardless of sand control technology that limits sand production, some of the destructed formation will be produced from the reservoir into the well. Produced sand will be accumulated in the well if the flow of liquid or gas does not transport solids to the wellhead. The accumulation of mechanical particles in the well will lead to a gradual filling of perforations and a decrease in production rate. In extreme cases, it is possible to form a sand plug and completely stop the well.

It is necessary to determine the nature of the movement of the sand-liquid mixture to predict accident-free operation of producing wells.

The main measurement parameters of the developed procedure for conducting laboratory experiments on sand packed tubes to establish the effectiveness of the sand control technology are: the number of suspended particles in the produced mixture, the particle size distribution of mechanical particles in the produced mixture during multi-stage sampling, and the relative permeability. These parameters make it possible to comprehensively assess the dynamics of the sand production process when using any sand control technology.

As a result of a series of laboratory experiments, it was established:

1. The number of suspended particles in the produced mixture depends on the volume distribution of phases in the fluid flow - filtration of pure oil leads to the highest suspended solids concentration (SSC). An increase in the proportion of brine leads to an increase in SSC. Filtration of gas-oil mixtures leads to the lowest SSC.

2. The number of suspended particles naturally decreases over time, due to the formation of arched systems and blocking of pore channels. The process of reducing SSC over time has been called "stabilization." The rate of "stabilization" depends on the volume distribution of phases in the fluid flow:

Water-oil mixtures have a high initial SSC, but over time the SSC value decreases to 15 – 25 % of the initial one;

• Gas-oil mixtures have a low initial SSC (compared to water-oil mixtures), but over time the SSC decreases only to 45 – 55 % of the initial one.

This phenomenon is associated with turbulization of the gas phase in the fluid flow high velocities of the gas phase lead to turbulent flow near the walls of the pore channel, due to which the separation of particles is more likely.

3. The increase in pressure drop leads to an increase

in the number of suspended particles in the produced mixture. It was experimentally established that with a quadruple increase in pressure drop (depression), on average, SSC increases 2.6 times. In this case, depending on the volume distribution of phases in the fluid flow and the "stage" of sand production, the SSC can increase by 1.2 to 8 times.

4. A decrease in grain size distribution does not result in an increase in the number of produced particles when using wire screens of the same aperture and design. This is due to an increase in the strength of the sand packed tubes due to cohesive interaction between particles. In addition, the decrease in grain size distribution indicates a high content of clay in its composition, which also plays the role of a consolidating material in the formation.

5. Oil flow leads to the transport of the ever-increasing diameter of rock particles due to the high flow force – over time, the flow of the oil phase entrains and carries out increasingly large particles. In the fine to medium part of the grain-size distribution (D25 and D50 of the original composition), the volume distribution of phases in the fluid flow does not significantly affect the transport capacity of particles. Water-oil fluids tend to increase the diameter of the produced mechanical impurities over time in the area of large grain-size distribution.

6. Application of mechanical screens leads to a decrease in the permeability of the system «screen – sand packed tube». However, the screen aperture and filtration pressure drop do not significantly affect the permeability. In most cases, decrease in the permeability does not depend on the suspended solids concentrations in the produced mixture and the particle size distribution of the produced particles, which suggests that the permeability of the system «screen – sand packed tube» changes mainly in the remote zone of the formation.

4. 1. Assessment of the Ability of Formation Fluids to Transport Particles of Different Grain Size Distribution

The complete production of sand particles can only be achieved by increasing the flow rate of fluid. To prevent the formation of a sand plug in the production casing, with a water cut of 4 %, the critical liquid flow rate should exceed 0.123 m/s, which corresponds to a production rate of 214 m³/day.

The increase in water cut will have a beneficial effect on the transport of sand in the horizontal well due to the increase in liquid production rate.

The growth of the gas factor, on the contrary, will negatively affect the mode of sand movement.

The amount of suspended particles in the studied range (20 – 2000 mg/l) does not significantly affect the transport of sand in the well. Areas with fixed sand are observed in the same areas. However, an increase in suspended solids concentration will lead to faster sand accumulation and early formation of a sand plug.

Critical velocities for streams with 50 μm and 200 μm particle diameters are substantially the same. For this reason, it is not necessary to install well screens with a small gap size. It is recommended to install wire-wrapped screens with an aperture of 200 μm instead of 150 μm, which will reduce the pressure drop on the screen and increase the well productivity.

4. 2. Development of a Combined Method for Limiting Sand Production Using Swellable Screen Elements

The concept of the method is to use already existing mechanical screens with the addition of a gravel-like chemicals. Expandable polymer is required to prevent coarse sand particles and the mechanical screen in turn will trap fine particles. The screen can also be holes inside the base pipe (production string), so that this element is not an obligatory part of the structure. The approximate structure of the technology is shown in Figure 13.

There are three main mechanisms for initiating the process of sand production. Two of them consist in violation of petroleum geomechanics due to exceeding the compressive and tensile strength by shear and tensile stresses, respectively. The dynamics of sand production due to tensile stresses, as a rule, is a short-term and local nature and does not lead to significant consequences during well operation. The third mechanism is associated with the volumetric destruction of pore space and is currently poorly studied due to the complexity of physical processes and the impossibility of clearly formalizing the challenge due to too many influencing factors. In addition, the presence of such a factor as the migration of microparticles of rocks during the fluid production in a porous medium should be noted.

According to laboratory experiments, a high pressure drop leads to selective production of sand particles with diameters of 1 – 20 μm and 40 – 200 μm. Literature analysis (61, 62) has shown that particles with a diameter of less than 50 μm do not create sand plugs, since they are carried out of the wellbore even at low rates, but have a significant effect on skin factor of the screen.

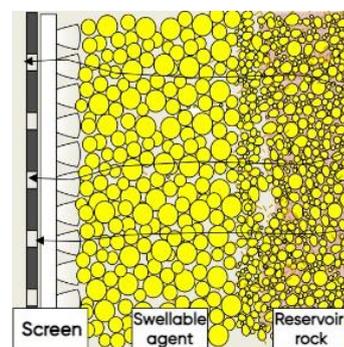


Figure 13. Scheme of reservoir with wire-wrapped screen with polymer

5. CONCLUSION

The paper presents a new concept of using shape memory polymers with wire-wrapped screens. Results of integrated modeling, including physical (laboratory), hydrodynamic and geomechanical, as well as analysis of the experience of using mechanical and chemical technologies for sand control in field conditions were revealed. Shape memory polymer technology will significantly reduce the suspended solids concentration during the well operation and will allow to avoid changes in the structure of the pore space near the bottom hole formation zone.

Results of the investigation might be presented in the following aspects:

1. A physical modeling method was developed to study the sand production process depending on the simulated operating conditions of the well (water cut, free gas content, pressure drop, etc.).

2. Author improved a set of mathematical models describing the process of sand production from the stage of initiation of violation of the stable geomechanical state and until the separation of rock particles from the walls of the wellbore and their further transportation along the horizontal well.

3. A mathematical model was created using recurrent neural networks to predict sand production in horizontal wells. The developed model makes it possible to determine the relationship between the number of suspended particles concentration and the main process parameters (oil flow rate, liquid flow rate, water cut, drop pressure) and use them to predict sand production.

4. A software module has been developed to determine the sections of the most probable accumulation of sand sediment with the calculation of the percentage of overlap of the inner diameter along the length of wellbore.

5. Proposed and substantiated is technology of sand control using shape memory polymer with wire-wrapped screen, which will allow limited passage of particles with diameter less than 50 μm , which creates conditions for non-colmatibility of screen while maintaining permeability.

Directions for future research:

1. The creation of a database of well-founded and confirmed experiments will make it possible to evaluate the proposed sand control technologies at the expert level.

2. The application of the developed technology will improve the efficiency of wells operation that complicated by sand occurrence and significantly reduce the workover operations.

3. The use of the proposed methodological approach will allow for a better choice of technology for sand control and will also allow to maintain a geomechanically stable of the bottomhole formation zone.

6. REFERENCES

- Peretomode E, Oluyemi G, Faisal NH. Sand production due to chemical-rock interaction. A Review. *Engineering failure analysis*. 2022;142:106745. <https://doi.org/10.1016/j.engfailanal.2022.106745>
- Raupov I, Burkhanov R, Lutfullin A, Maksyutin A, Lebedev A, Safiullina E. Experience in the application of hydrocarbon optical studies in oil field development. *Energies*. 2022;15(10):3626. <https://doi.org/10.3390/en15103626>
- Aleksander G P, Yifan T, Fuming Z. Predicting Service Life of Polyethylene Pipes under Crack Expansion using. *International Journal of Engineering, Transactions C: Aspects*. 2023;36(12):2243-52. 10.5829/ije.2023.36.12c.14
- 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. <https://doi.org/10.31897/PMI.2022.100>
- Mahmoudi M, Roostaei M, Ghalambor A, editors. Sand screen design and optimization for horizontal wells using reservoir grain size distribution mapping. *SPE International Conference and Exhibition on Formation Damage Control*; 2016: SPE.
- Garolera D, Carol I, Papanastasiou P. Micromechanical analysis of sand production. *International Journal for Numerical and Analytical Methods in Geomechanics*. 2019;43(6):1207-29. 10.1002/nag.2892
- Mahmud HB, Leong VH, Lestariyono Y. Sand production: A smart control framework for risk mitigation. *Petroleum*. 2020;6(1):1-13. <https://doi.org/10.1016/j.petlm.2019.04.002>
- Eshiet KI-I, Yang D, Sheng Y. Computational study of reservoir sand production mechanisms. *Geotechnical Research*. 2019;6(3):177-204. <https://doi.org/10.1680/jgere.18.00026>
- Han G, Dusseault M, editors. Quantitative analysis of mechanisms for water-related sand production. *SPE International Conference and Exhibition on Formation Damage Control*; 2002: SPE.
- Vaziri H, Allam R, Kidd G, Bennett C, Grose T, Robinson P, et al. Sanding: a rigorous examination of the interplay between drawdown, depletion, startup frequency, and water cut. *SPE Production & Operations*. 2006;21(04):430-40. <https://doi.org/10.2118/89895-MS>
- Zhang R, Shi X, Zhu R, Zhang C, Fang M, Bo K, et al. Critical drawdown pressure of sanding onset for offshore depleted and water cut gas reservoirs: modeling and application. *Journal of Natural Gas Science and Engineering*. 2016;34:159-69. <https://doi.org/10.1016/j.jngse.2016.06.057>
- Raupov I, Milic J, editors. Improvement of operational efficiency of high water-cut oil wells. *IOP Conference Series: Earth and Environmental Science*; 2022: IOP Publishing.
- Zamani MAM, Knez D. A new mechanical-hydrodynamic safety factor index for sand production prediction. *Energies*. 2021;14(11):3130. <https://doi.org/10.3390/en14113130>
- Bianco L, Halleck P, editors. Mechanisms of arch instability and sand production in two-phase saturated poorly consolidated sandstones. *SPE European Formation Damage Conference and Exhibition*; 2001: SPE.
- Zhang R, Wang J, Liu S, Zhang Z, Ma L, Meng W, et al. Modeling and Durability Behavior of Erosion-Corrosion of Sand Control Screens in Deepwater Gas Wells. *ACS omega*. 2021;6(37):23943-51. 10.1021/acsomega.1c02960
- Abduljabbar A, Mohyaldinn ME, Younis O, Alghurabi A, Alakbari FS. Erosion of sand screens by solid particles: a review of experimental investigations. *Journal of Petroleum Exploration*

- and Production Technology. 2022;12(8):2329-45. <https://doi.org/10.1007/s13202-022-01467-4>
17. Tananykhin D, Grigorev M, Korolev M, Solovyev T, Mikhailov N, Nesterov M. Experimental evaluation of the multiphase flow effect on sand production process: Prepack sand retention testing results. *Energies*. 2022;15(13):4657. <https://doi.org/10.3390/en15134657>
 18. Tananykhin D, Grigorev M, Simonova E, Korolev M, Stecyuk I, Farrakhov L. Effect of wire design (Profile) on sand retention Parameters of wire-wrapped screens for conventional production: Prepack sand retention testing results. *Energies*. 2023;16(5):2438. <https://doi.org/10.3390/en16052438>
 19. Cameron J, Zaki K, Jones C, Lazo A, editors. Enhanced flux management for sand control completions. SPE Annual Technical Conference and Exhibition?; 2018: SPE.
 20. Alakbari FS, Mohyaldinn ME, Muhsan AS, Hasan N, Ganat T. Chemical sand consolidation: From polymers to nanoparticles. *Polymers*. 2020;12(5):1069. <https://doi.org/10.3390/polym12051069>
 21. Sveta PE, Gustavsen Ø, Solli HS, Brandal Ø, editors. Identifying point of failure and repairing damaged sand screens in gravel packed wells. A Case History From The Heidrun Field. SPE European Formation Damage Conference and Exhibition; 2009: SPE.
 22. Haftani M, Kotb O, Nguyen PH, Wang C, Salimi M, Nouri A. A Novel sand control testing facility to evaluate the impact of radial flow regime on screen performance and its verification. *Journal of Petroleum Science and Engineering*. 2020;195:107903. <https://doi.org/10.1016/j.petrol.2020.107903>
 23. Raupov I, Rogachev M, Sytnik J. Design of a polymer composition for the conformance control in heterogeneous reservoirs. *Energies*. 2023;16(1):515. <https://doi.org/10.3390/en16010515>
 24. Shishlyannikov D, Zvonarev I, Rybin A, Zverev V, Ivanchenko A. Assessment of Changes in the Abrasiveness of Solid Particles in Hydraulic Mixtures Pumped with ESPs. *Applied Sciences*. 2023;13(3):1885. <https://doi.org/10.3390/app13031885>
 25. Ballard T, Beare S, editors. An investigation of sand retention testing with a view to developing better guidelines for screen selection. SPE International Conference and Exhibition on Formation Damage Control; 2012: SPE.
 26. Chanpura RA, Hodge RM, Andrews JS, Toffanin EP, Moen T, Parlar M. A review of screen selection for standalone applications and a new methodology. *SPE Drilling & Completion*. 2011;26(01):84-95. <https://doi.org/10.2118/127931-PA>
 27. Constien VG, Skidmore V, editors. Standalone screen selection using performance mastercurves. SPE International Conference and Exhibition on Formation Damage Control; 2006: SPE. 10.2118/98363-MS
 28. Underdown D, Dickerson R, Vaughan W, editors. The nominal sand control screen: a critical evaluation of screen performance. SPE Annual Technical Conference and Exhibition?; 1999: SPE.
 29. Anderson M, editor SAGD sand control: large scale testing results. SPE Canada Heavy Oil Conference; 2017: SPE.
 30. Ballard T, Beare S, editors. Media sizing for premium sand screens: Dutch twill weaves. SPE European Formation Damage Conference and Exhibition; 2003: SPE.
 31. Ballard T, Beare SP, editors. Sand retention testing: the more you do, the worse it gets. SPE International Conference and Exhibition on Formation Damage Control; 2006: SPE.
 32. Devere-Bennett N, editor Using prepack sand-retention tests (SRT's) to narrow down liner/screen sizing in SAGD wells. SPE Thermal Integrity and Design Symposium; 2015: SPE.
 33. Montero J, Chissonde S, Kotb O, Wang C, Roostaei M, Nouri A, et al., editors. A critical review of sand control evaluation testing for SAGD applications. SPE Canada Heavy Oil Conference; 2018: SPE.
 34. Wang Y, Wu B, editors. Borehole collapse and sand production evaluation: Experimental testing, analytical solutions and field implications. ARMA US Rock Mechanics/Geomechanics Symposium; 2001: ARMA.
 35. Tabatabaee Moradi S, Nikolaev N. Study of bonding strength at salt-cement interface during cementation of salt layers. *International Journal of Engineering, Transactions B: Applications*. 2021;34(2):581-6. 10.5829/IJE.2021.34.02B.32
 36. Salama MM. Sand production management. *J Energy Resour Technol*. 2000;122(1):29-33. 10.1115/1.483158
 37. Oroskar AR, Turian RM. The critical velocity in pipeline flow of slurries. *AIChE Journal*. 1980;26(4):550-8. 10.1002/AIC.690260405
 38. Danielson TJ, editor Sand transport modeling in multiphase pipelines. Offshore Technology Conference; 2007: OTC. 10.4043/18691-MS
 39. Petrakov D, Loseva A, Alikhanov N, Jafarpour H. Standards for selection of surfactant compositions used in completion and stimulation fluids. *International Journal of Engineering, Transactions C: Aspects*. 2023;36(9):1605-10. 10.5829/ije.2023.36.09c.03
 40. Rajaoalison H, Zlotkowski A, Rambolamanana G. Mechanical properties of sandstone using non-destructive method. *Записки Горного института*. 2020;241:113-7. <https://doi.org/10.31897/pmi.2020.1.113>
 41. Podoprigora D, Byazrov R, Sytnik J. The comprehensive overview of large-volume surfactant slugs injection for enhancing oil recovery: Status and the Outlook. *Energies*. 2022;15(21):8300. <https://doi.org/10.3390/en15218300>
 42. Dorfman MB, Sentemov AA, Belozero IP. The study of displacing ability of lignosulfonate aqueous solutions on sand packed tubes. *Записки Горного института*. 2023(264 (eng)):865-73.
 43. Thomas DG. Transport characteristics of suspensions: Part IX. Representation of periodic phenomena on a flow regime diagram for dilute suspension transport. *AIChE Journal*. 1964;10(3):303-8. 10.1002/aic.690100307
 44. Stevenson P, Thorpe RB. Energy dissipation at the slug nose and the modeling of solids transport in intermittent flow. *The Canadian Journal of Chemical Engineering*. 2003;81(2):271-8. 10.1002/CJCE.5450810213
 45. Ibarra R, Mohan RS, Shoham O, editors. Critical sand deposition velocity in horizontal stratified flow. SPE International Conference and Exhibition on Formation Damage Control; 2014: SPE. 10.2118/168209-MS
 46. Hill AL. Determining the Critical Flow Rates For Low-concentration Sand Transport in Two-phase Pipe Flow by Experimentation and Modeling: University of Tulsa; 2011.
 47. Leporini M, Terenzi A, Marchetti B, Corvaro F, Polonara F. On the numerical simulation of sand transport in liquid and multiphase pipelines. *Journal of Petroleum Science and Engineering*. 2019;175:519-35. 10.1016/J.PETROL.2018.12.057
 48. Mendoza L, Marin M, Nascimento C, Peraza R, editors. Sand transport modeling in heavy oil gathering network in Orinoco oil belt, Venezuela. SPE Canada Heavy Oil Conference; 2017: SPE.
 49. Dabirian R, Arabnejad Khanouki H, Mohan RS, Shoham O. Numerical simulation and modeling of critical sand-deposition velocity for solid/liquid flow. *SPE Production & Operations*. 2018;33(04):866-78. 10.2118/187049 PA
 50. Baghdadi AF, Gupta A, Kamat D, Singh R, Borhan A, Jadid M, editors. Development of Sand Erosion and Transport Software for Efficient Sand Management. Offshore Technology Conference Asia; 2016: OTC. 10.4043/26448-ms

51. Shahsavari MH, Khamehchi E, Fattahpour V, Molladavoodi H. Investigation of sand production prediction shortcomings in terms of numerical uncertainties and experimental simplifications. *Journal of Petroleum Science and Engineering*. 2021;207:109147. <https://doi.org/10.1016/j.petrol.2021.109147>
52. Linh N, Gabov V, Lykov Y, Urazbakhtin R. Evaluating the Efficiency of Coal Loading Process by Simulating the Process of Loading onto the Face Conveyor with a Shearer with an Additional Share. *International Journal of Engineering, Transactions A: Basics*. 2021;34(7):1804-9. 10.5829/IJE.2021.34.07A.25
53. Rogachev MK, Aleksandrov AN. Justification of a comprehensive technology for preventing the formation of asphalt-resin-paraffin deposits during the production of highly paraffinic oil by electric submersible pumps from multiformation deposits. *Записки Горного института*. 2021;250:596-605. <https://doi.org/10.31897/PMI.2021.4.13>
54. Ahad NA, Jami M, Tyson S. A review of experimental studies on sand screen selection for unconsolidated sandstone reservoirs. *Journal of petroleum exploration and production technology*. 2020;10(4):1675-88. <https://doi.org/10.1007/s13202-019-00826-y>
55. Poplygin VV. Study on Application of Arps Decline Curves for Gas Production Forecasting in Senegal. *International Journal of Engineering, Transactions C: Aspects*. 2023;36(12):2207-13. 10.5829/ije.2023.36.12c.10
56. Kuncoro B, Ulumuddin B, Palar S, editors. Sand control for unconsolidated reservoirs. *National Symposium, Jakarta, Indonesia*; 2001.
57. Коробов ГЮ, Воронцов АА. STUDY OF CONDITIONS FOR GAS HYDRATE AND ASPHALTENE-RESIN-PARAFFIN DEPOSITS FORMATION IN MECHANIZED OIL PRODUCTION. *Bulletin of the Tomsk Polytechnic University Geo Assets Engineering*. 2023;334(10):61-75. <https://doi.org/10.18799/24131830/2023/10/4181>
58. Leone JA, Scott EM. Characterization and control of formation damage during waterflooding of a high-clay-content reservoir. *SPE reservoir engineering*. 1988;3(04):1279-86. <https://doi.org/10.2118/16234-PA>
59. Wang X, Osunjaye G, editors. Advancement in openhole sand control applications using shape memory polymer. *SPE Annual Technical Conference and Exhibition?*; 2016: SPE.
60. Araujo-Guerrero E, Alzate-Espinosa G, Cross-Arroyave Y, Vega-Niño YP, Cartagena-Pérez DF, Naranjo-Agudelo AJ, editors. A new methodology for selecting sand control or sand management as strategy in wells with sand production potential. *ISRM International Symposium Geomechanics*; 2021: ISRM.
61. Vatani M, Domiri-Ganji D. Experimental examination of gas-liquid two-phase flow patterns in an inclined rectangular channel with 90 bend for various vertical lengths. *International Journal of Engineering*. 2022;35(4):685-91. 10.5829/IJE.2022.35.04A.07
62. Mahmoudi M, Roostaei M, Fattahpour V, Sutton C, Fermaniuk B, Zhu D, et al., editors. Standalone sand control failure: review of slotted liner, wire wrap screen, and premium mesh screen failure mechanism. *SPE Annual Technical Conference and Exhibition?*; 2018: SPE.

COPYRIGHTS

©2024 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 رویکرد مبتنی بر علمی جدید را برای انتخاب فناوری کنترل شن و ماسه ارائه می‌کند که به افزایش کارایی توسعه مخازن تجمع نشده اختصاص دارد. مجتمع‌های آزمایشگاهی و روش‌شناسی ایجاد شده برای شبیه‌سازی فیزیکی فرآیند تولید شن و ماسه به منظور کسب دانش جدید و تایید نظریه‌های موجود مورد تجزیه و تحلیل قرار گرفت. همه آنها مزایا و معایب خود را دارند، اما کاربرد همزمان آنها وابستگی‌های مشخصه‌ای بین تولید ماسه و پارامتر مورد مطالعه (توزیع اندازه دانه، افت فشار، محتوای رس، قطع آب، نسبت گاز به نفت و غیره) را نشان داد. نویسنده مفاهیم بهبود دستگاه ریاضی را برای افزایش کیفیت ارزیابی توانایی سیالات سازند برای انتقال ذرات با توزیع اندازه دانه‌های مختلف در سازند و همچنین در قسمت داخلی لوله پیشنهاد کرد. اثر هر یک از عوامل مشخص‌کننده بر هماهنگی جامدات معلق (SSC) در نتیجه بیش از ۳۰۰ آزمایش آزمایشگاهی مورد مطالعه قرار گرفت. با توجه به مشاهده، کاهش شدید SSC پس از مرحله اول (نمونه برداری) وجود دارد. بنابراین، نویسنده تعیین کرد که جریان اصلی ناخالصی‌های مکانیکی در طول تحریک جریان و پس از خاموش شدن رخ می‌دهد. در نتیجه، نویسنده روش محدود کردن تولید شن و ماسه با استفاده از پلیمرهای دارای حافظه شکل را بر اساس نتایج مجموعه آزمایش‌های انجام شده اثبات کرد. روش پیشنهادی اجازه عبور محدود ذرات با قطر کمتر از ۵۰ میکرومتر را می‌دهد که شرایطی را برای غیرکلماتیک بودن صفحه و در عین حال حفظ ثبات ژئومکانیکی ناحیه تشکیل سوراخ پایین ایجاد می‌کند.



Bi-objective Economic Production Quantity with Partial Backordering under Uncertainty

M. Najafi^a, A. Ghodratnama^{*a}, S. H. R. Pasandideh^a, R. Tavakkoli-Moghaddam^b

^a Department of Industrial Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran

^b School of Industrial Engineering, College of Engineering, University of Tehran, Tehran, Iran

PAPER INFO

Paper history:

Received 27 August 2023

Received in revised form 22 November 2023

Accepted 27 December 2023

Keywords:

Bi-objective Mathematical Model

Economic Production Quantity

Rework and Shortage

Meta-heuristics

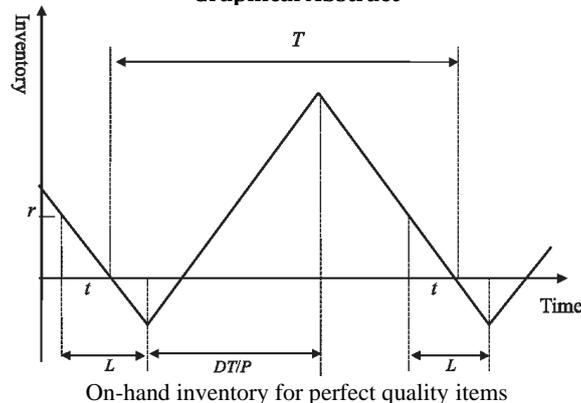
Uncertainty

ABSTRACT

The economic production quantity (EPQ) model considers the production rate, demand rate, setup costs, holding costs, and shortage costs to find the production quantity that minimizes the sum of these costs. The goal is to balance the costs associated with production, holding inventory, and potential shortages. In this paper, two objectives include the costs of production and ordering and others in a separate objective function. In the objectives of the other costs, The cost of storage space as a supply is defined to be minimized. This study considers scrap and reworks in the EPQ model. This inventory model accounts for many items on a single machine. The production capacity is reduced, and there are shortages when only one machine exists. By determining the quantities of the products produced by the manufacturing facility, the storage space for each product, cycle time, and product scarcity, we can reduce both the overall cost and the supply cost of warehouse space due to non-linearity and the inability to solve commercial software in large dimensions, a multi-objective meta-heuristic algorithm, namely the non-dominated sorting genetic algorithm (NSGA-II), is used. The findings are further validated using the non-dominated ranking genetic algorithm (NRGA). Also, the obtained Pareto front is studied with several indicators. To perform these two algorithms at the best condition, we employed the Taguchi approach and related orthogonal arrays and performed algorithms for each array considering several factors. Also, to validate the mathematical model, we used the augmented epsilon-constraint method executed in the GAMS environment. It is clear that GAMS commercial software yields better results; however, these two algorithms are justifiable when the problem becomes bigger. Finally, by performing a sensitivity analysis for these indicators and the objective functions, the behavior of the proposed algorithms is compared and examined in detail. Also, the superior algorithm is chosen using the TOPSIS as a multi-criteria decision-making method. Numerical examples show how the presented model and the proposed algorithms may be used efficiently. A surveying literature review clarifies that the related objective functions, constraints, and solution approaches have not been investigated until now.

doi: 10.5829/ije.2024.37.07a.18

Graphical Abstract



*Corresponding Author's Email: ghodratnama@khu.ac.ir (A. Ghodratnama)

NOMENCLATURE			
d_i	Demand for the i -th product	A_i	Setup expense for producing a batch of the i -th product
p_i	Production rate of the i -th product	ts_i	Duration of machine setup to manufacture the i -th product
u_i	Pauly products of the i -th product	o_i	Supply cost per unit of storage space
v_i	Product backorders as a percentage of its shortfall	f_i	Space occupied by each unit of the i -th product
c_i	Manufacturing cost per unit of the i -th product	q_i^s	For each cycle ($i = 1, 2, \dots, n$), the size of the manufacturing lot for the i -th product
k_i	Reworking the costs per item of the i -th product	$T \geq 0$	Cycle's duration
h_i	Price per item and holding time for the i -th product	$N \geq 0$	Cycles per year, number
b_i	Backorder price for the i -th product	$X_i \geq 0$	Continuous random variable represents the product's storage space.,
sc_i	Scrap percentage of the i -th product	$s_i \geq 0$	Overall shortfall amount for the i -th product of the cycle
H_i	After normal manufacturing ends, the product's maximum amount of inventory	H_i^{\max}	Final amount of available inventory for the product

1. INTRODUCTION

Inventory control is crucial for maintaining inventory levels and minimizing system costs. It ensures that an organization's existing items are available for production, distribution, sales, and engineering operations management departments, considering factors like time, location, quantity, quality, and cost. Inventory control includes raw materials and products stored in the warehouse. Operations management involves designing and managing products, processes, services, and supply chains. It includes strategic, tactical, and operational levels.

Improving competitive power and comprehensiveness in supply chains is essential for efficient systems. Although the economic production quantity model is often used in inventory management and manufacturing, it is crucial to investigate if damaged goods are included in inventory models. This study investigates the economic production quantity (EPQ) model during production periods for different time components, shortages, and lost sales, focusing on production costs, ordering, and storage space.

However, the EPQ model is specifically designed for situations where items are produced or manufactured rather than simply ordered. It is often applied to scenarios where production rates are finite and may vary. The EPQ model takes into account factors such as production rate constraints, setup costs, holding costs, and demand for the product. The key components of the EPQ model include:

- Demand: The rate at which customers are requesting the product.
- Setup (or production) cost: The cost associated with setting up the production process, including the cost of preparing the machinery, changing tools, etc.
- Holding (or carrying) cost: The cost of holding or storing inventory, including expenses related to warehousing, insurance, and potential obsolescence.
- Production rate: The rate at which units are produced.

The goal of the EPQ model is to find the production quantity that minimizes the total cost, taking into consideration the trade-off between setup costs and holding costs. The formula for the Economic Production Quantity is derived based on mathematical optimization techniques, and it helps businesses determine the most cost-effective production quantity to meet demand.

This paper's remaining sections are organized as follows.. Section two scrutinizes the related literature review meticulously. The mathematical formulation of the issue is covered in section 3. The methods for the solutions are in section 4. Section 5 deals with the solution and comparison of numerical instances. The conclusion and some recommendations for more research are included in section 6. References are also included in section 7.

2. LITERATURE REVIEW

Cunha et al. (1) examined the economic production quantity model with partial backordering and a discount for batches of subpar quality. Shah and Vaghela (2) created and refined a flawed production inventory model for time- and effort-dependent demand under inflation and optimum dependability. Taleizadeh et al. (3) developed sustainable economic output quantity models for shortage inventory systems. Al-Salamah (4) investigated how much economic production might be produced in a manufacturing process that included faults and configurable synchronous and asynchronous rework rates. The economic order quantity (EOQ) and EPQ inventory models with two backorder charges were developed by Lin (5) using analytical geometry and algebra. Marchi et al. (6) examined the economic production quantity model. It includes learning production, quality reliability, and energy efficiency.

The EOQ and EPQ inventory models with partial backorder issues were studied by Thinakaran et al. (7). For the economic output quantity and the joint economic lot size, Zavanella et al. (8) considered energy. To resolve

an EPQ model for an inefficient manufacturing process, De et al. (9) employed a game-based approach and a neutrosophic fuzzy method. Ganesan and Uthayakumar (10) created EPQ models for a flawed manufacturing system that considers warm-up production runs, shortages during hybrid maintenance periods, and partial backordering.

Guha and Bose (11) presented the EPQ in Batch Manufacturing with imperfect quality and non-destructive acceptance sampling. Insights from an EPQ model were investigated by Hauck et al. (12) on the impact of early inspection on the functionality of production systems. According to Kalantari and Taleizadeh (13), mathematical modeling may be used to find the best replacement for failed items in an EPQ model with several shipments. In 2020, that was formulated by Nobil et al. (14). An economic production quantity inventory model with discrete delivery orders, common production standards, and budgetary constraints for several items produced on a single machine. The Development and Solvency were reported by Rahaman et al. (15).

The production model in terms of amount under arbitrary commands both with and without deterioration. Artificial bee colony optimization was the basis for Rahman et al.'s (16) synergetic analysis of the fractional-order economic production quantity model. Economic output quantity models that are sensitive to forecasting maintenance and modified later on. An economic production quantity model for three tiers of work was designed by using the Weibull distribution degradation and shortage. Bose and Guha (17) looked at the economic production lot size under the conditions of low quality, online inspection, and inspection errors. A comment on the cost comparison method used to address the EOQ and EPQ concerns.

Using fuzzy geometric programming (GP), different fuzzification and defuzzification approaches, and an unconstrained multi-item model, Kalaiarasi et al. (18) claim that this model was optimized. Moghdani et al. (19) considered a multi-item fuzzy economic production quantity model with multiple deliveries. Shekhawat et al. (20) looked at the EPQ model for deteriorating items with a Weibulleterioration rate throughout the finite time horizon. To solve a multi-product, single-machine EPQ inventory model utilizing GP mode, Kalaiarasi et al. (21) employed Python.

A quantitative model of economic production with a fluctuating energy price has been constructed. Nobil et al. (22) considered a setup time/cost function for a multi-product imperfect manufacturing system and an economic production quantity inventory model. Priyan et al. (23) examined a cleaner EPQ inventory model with synchronous and asynchronous rework procedures and investments in green technologies. Edalatpour et al. (24) integrated sustainability concerns with pricing and inventory decisions for degrading items. Also, some

researchers studied other aspects of EPQ issue (25-27). Figure 1 illustrates the number of papers published in this regard.

According to the considerable literature on the economic production quantity model, it is evident that interest in the topic of an incomplete production system, a problem that affects real-world manufacturing, is still expanding. Since the model created in this work is challenging to solve analytically, Pareto fronts are discovered using a non-dominated sorting genetic algorithm (NSGA-II) and the non-dominated ranking genetic algorithm (NRGA).

3. PROPOSED MATHEMATICAL MODEL

In this section, the assumptions of the proposed model are explained first, and then the parameters and variables are defined; in the following, the problem's mathematical model and the constraints' definition are discussed in detail.

3. 1. Assumptions The main assumptions of the proposed model are as follows:

- Considering a manufacturing system with flawed production procedures.
- At a rate of u_i ; $i=1,2,\dots,n$ every cycle, incomplete objects of n various sorts are produced, and among these goods, the SC_i part is considered to be junk, while the other portion may be reworked.
- In each cycle, most parameters are seen as unknown.
- The parameters are produced randomly using uniform distributions in the respective ranges for various issues.
- One machine is used to make all of the goods.
- All things are believed to have a certain cycle length, $T_1=T_2=\dots=T_n=T$.
- Assume that the number of goods produced corresponds to a corresponding demand in every cycle, with a production rate of π per cycle.

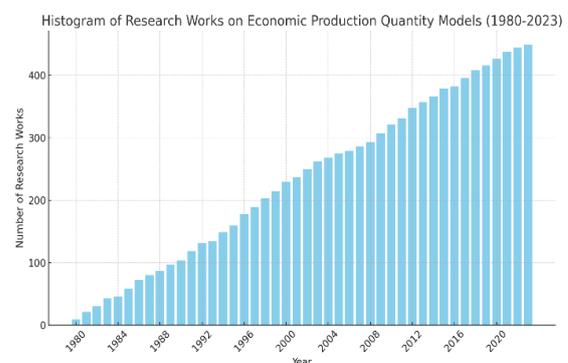


Figure 1. Histogram representing the number of papers published in this regard

- It is assumed that the number of goods produced, which corresponds to a corresponding demand, d_i , in every cycle, has a production rate of p_i per cycle for the i -th item.
- At the end of the rework period, we expect that all reworkable objects will be reworked, and one part will be left behind as scrap
- We expect all of the reworkable objects to be reworked, and a m_i part will be left behind as scrap after the rework period.
- Producers use the same resource for production and rework simultaneously.
- The budget and capacity for the standard manufacturing system are limited, and some shortfalls are being backordered. The fundamental concept of the EPQ inventory model with the rework process is that the production rate of fewer defectives must always be more than or equal to the demand (28). The production cycle length is determined as the average of the good and incomplete item production up times (t_i^1 and t_i^5 , respectively), the reworking time (t_i^2), and the good and incomplete item production downtimes (t_i^3 and t_i^4 , respectively).

$$T = \sum_{j=1}^5 t_i^j \tag{1}$$

Figure 2 depicts the cycle time for each product since all products are produced on a single machine with a limited capacity. This has led to the following equations:

$$t_i^1 = \frac{q_i^s}{p_i} - \frac{v_i s_i}{(1-u_i-sc_i)p_i-d_i} \tag{2}$$

$$t_i^1 = u_i \frac{q_i^s}{p_i} \tag{3}$$

$$t_i^3 = \frac{H_i^{\max}}{d_i} = \left(\frac{(1-sc_i-m_i u_i)}{d_i} - \frac{(1+u_i)}{p_i} \right) q_i^s - \frac{v_i s_i}{d_i} \tag{4}$$

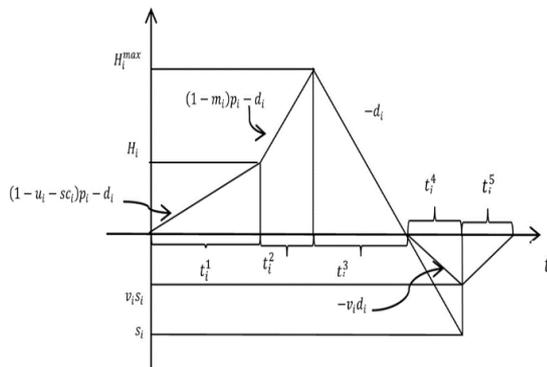


Figure 2. On-hand inventory for perfect quality items

$$t_i^4 = \frac{s_i}{d_i} \tag{5}$$

$$t_i^5 = \frac{v_i s_i}{(1-u_i-sc_i)p_i-d_i} \tag{6}$$

Thus, based on Equation 1, the length of the cycle for a single product is as follows:

$$H_i = \left((1-u_i-sc_i)p_i-d_i \right) \frac{q_i^s}{p_i} - v_i s_i \tag{7}$$

$$H_i^{\max} = H_i + u_i \left((1-m_i)p_i-d_i \right) \frac{q_i^s}{p_i} \tag{8}$$

$$H_i^{\max} = \left((1-sc_i-m_i u_i)p_i - (1+u_i)d_i \right) \frac{q_i^s}{p_i} - v_i s_i \tag{9}$$

$$T = \frac{(1-v_i)s_i + (1-sc_i-m_i u_i)q_i^s}{d_i} \tag{9}$$

$$q_i^s = \frac{Td_i - (1-v_i)s_i}{(1-sc_i-m_i u_i)} \tag{10}$$

3. 2. Function of the Total Cost The following describes the model's total cost function:

$$TC = CA + CP + CR + CH + CB + CL \tag{11}$$

3. 2. 1. Setup Cost A setup costs and occurs N times annually.

As a result, the yearly setup cost is as follows:

$$CA = \sum_{i=1}^n NA_i \tag{12}$$

$$N = \frac{1}{T} \tag{13}$$

3. 2. 2. Production Cost The sum of the total production cost is the sum of the production costs per unit and quantity per period for all i -th commodities, respectively of manufacturing annually:

$$CP = \frac{1}{T} \sum_{i=1}^n c_i q_i^s \tag{14}$$

3. 2. 3. Rework Cost k_i reflects how much of the i -th product has to be changed. The sum of the rework cost per unit of the i -th product is what is referred to as the annual rework cost. The year of work may be determined by multiplying the total cost by N . The cost of this shared insurance is as follows:

$$CR = \frac{1}{T} \sum_{i=1}^n k_i u_i q_i^s \quad (15)$$

3. 2. 4. Holding Cost

The holding costs of the inventory system for independent and collaborative production strategies expressed in Equation 16 as shown in Figure 1,

$$CH = \frac{1}{T} \sum_{i=1}^n h_i \left[\frac{H_i}{2} (t_i^1) + \frac{H_i + H_i^{\max}}{2} (t_i^2) + \frac{H_i^{\max}}{2} (t_i^3) \right] \quad (16)$$

3. 2. 5. Backorder Cost

Related to Figure 1, expressions 17 and 18 indicate the back-ordered and lost selling expenditures for each cycle.

$$CB = \frac{1}{2T} \sum_{i=1}^n b_i v_i s_i (t_i^4 + t_i^5) \quad (17)$$

$$CL = \frac{1}{2T} \sum_{i=1}^n l_i (1 - v_i) s_i \quad (18)$$

3. 2. 6. Lost Sale Cost

As a consequence, the model's objective function is as follows:

$$\begin{aligned} TC &= CA + CP + CR + CH + CB + CL \\ &= \frac{1}{T} \sum_{i=1}^n A_i + \frac{1}{T} \sum_{i=1}^n c_i q_i^s + \frac{1}{T} \sum_{i=1}^n k_i u_i q_i^s \\ &\quad + \frac{1}{T} \sum_{i=1}^n h_i \left[\frac{H_i}{2} (t_i^1) + \frac{H_i + H_i^{\max}}{2} (t_i^2) + \frac{H_i^{\max}}{2} (t_i^3) \right] \\ &\quad + \frac{1}{2T} \sum_{i=1}^n b_i v_i s_i (t_i^4 + t_i^5) + \frac{1}{2T} \sum_{i=1}^n l_i (1 - v_i) s_i \end{aligned} \quad (19)$$

3. 5. Cost of Storage Space as a Supply

The supply cost of a warehouse is determined as the product of the supply cost per storage space and the continuous random variables X_i representing the storage area of a particular product, respectively.

$$G = \sum_{i=1}^n o_i X_i \quad (20)$$

3. 6. Constraints

The related constraints are as follows:

3. 6. 1. Capacity Constraint

In collaborative production systems that include rework, the combined production, rework, and setup times need to be less than the cycle time. In our issue, T must be less than or equal to $\sum_{i=1}^n (t_i^1 + t_i^2 + t_i^5) + \sum_{i=1}^n t s_i$. As a result, this is the capacity-constrained model:

$$\sum_{i=1}^n (t_i^1 + t_i^2 + t_i^5) + \sum_{i=1}^n t s_i \leq T \quad (21)$$

Equations 2, 3, and 6 give the capacity constraint model as follows:

$$\sum_{i=1}^n (1 + u_i) \frac{T d_i - (1 - v_i) s_i}{(1 - s c_i - m_i u_i) p_i} + \sum_{i=1}^n t s_i \leq T \quad (22)$$

3. 6. 2. Budget Constraint

Given that the entire budget is W , the manufacturing quantity is represented, and the i -th is reworked $u_i q_i^s$. The current budget constraint is as follows:

$$\sum_{i=1}^n (c_i q_i^s + k_i u_i q_i^s) \leq W \quad (23)$$

3. 6. 3. Service Level Constraint

For the service level constraint, the i -th product's annual demand, safety margin for allowable shortage, period-by-period shortfall amount, and several periods are, in that order: S_i , d_i , SL , and N . The current service level limitation is as follows:

$$\sum_{i=1}^n \frac{S_i}{T d_i} \leq SL \quad (24)$$

Constraints 22-24.

3. 6. 4. Warehouse-Space Constraint

There is definite room in the warehouse to keep the goods.

$$f_i H_i^{\max} \leq X_i \quad (25)$$

3. 7. Final Model

$$\begin{aligned} Min Z &= \frac{1}{T} \sum_{i=1}^n A_i + \frac{1}{T} \sum_{i=1}^n C_i \frac{T d_i - (1 - v_i) s_i}{(1 - s c_i - m_i u_i)} + \\ &\quad \frac{1}{T} \sum_{i=1}^n k_i u_i \frac{T d_i - (1 - v_i) s_i}{(1 - s c_i - m_i u_i)} + \frac{1}{T} \sum_{i=1}^n h_i \\ &\quad \left[\frac{1}{2} \left(\left((1 - u_i - s c_i) p_i - d_i \right) \left(\frac{T d_i - (1 - v_i) s_i}{(1 - s c_i - m_i u_i) p_i} \right) - v_i s_i \right) \right. \\ &\quad \left. \left(\frac{T d_i - (1 - v_i) s_i}{(1 - s c_i - m_i u_i) p_i} - \frac{v_i s_i}{(1 - u_i - s c_i) p_i - d_i} \right) + \right. \\ &\quad \left. \left(\left((1 - 0.5 u_i - s c_i - 0.5 m_i u_i) - (1 + 0.5 u_i) \frac{d_i}{p_i} \right) \right) - v_i s_i \right. \\ &\quad \left. \left(\frac{T d_i - (1 - v_i) s_i}{(1 - s c_i - m_i u_i) p_i} \right) \right. \\ &\quad \left. \left(\frac{T d_i - (1 - v_i) s_i}{u_i (1 - s c_i - m_i u_i) p_i} \right) \right. \\ &\quad \left. \frac{1}{2 d_i} \left(\left((1 - s c_i - m_i u_i) p_i - (1 + u_i) d_i \right) \left(\frac{T d_i - (1 - v_i) s_i}{(1 - s c_i - m_i u_i) p_i} \right) - v_i s_i \right)^2 \right] \\ &\quad + \frac{1}{2T} \sum_{i=1}^n b_i v_i s_i \left(\frac{s_i}{d_i} + \frac{v_i s_i}{(1 - u_i - s c_i) p_i - d_i} \right) + \frac{1}{2T} \sum_{i=1}^n l_i (1 - v_i) s_i \end{aligned} \quad (26)$$

$$\text{Min } G = \sum_{i=1}^n o_i X_i$$

s.t.

$$\sum_{i=1}^n (c_i + k_i u_i) \left(\frac{Td_i - (1-v_i)s_i}{(1-sc_i - m_i u_i)} \right) \leq W \quad (27)$$

$$f_i \left(\left((1-sc_i - m_i u_i) p_i - ((1+u_i)d_i) \right) \left(\frac{Td_i - (1-v_i)s_i}{(1-sc_i - m_i u_i) p_i} \right) - v_i s_i \right) \leq X_i \quad (28)$$

4. SOLUTION APPROACHES

In terms of approaches to solutions For reactive berth allocation and scheduling at maritime container ports in reaction to disturbances, Dulebenets (29) took into consideration a dispersed memetic optimizer. An adaptable polyloid memetic algorithm was also considered by Dulebenets (30) for truck scheduling at a cross-docking terminal. In multi-objective settings, Pasha et al. (31) used a factory-in-a-box to look at precise and metaheuristic algorithms for the vehicle routing issue. Singh and Pillay (32) considered analyzing ant-based pheromone spaces for generating perturbative meta-heuristics. The development of precise and heuristic optimization techniques for safety enhancement projects at level crossings under competing goals was considered by Singh et al. (33). Chen and Tan (34) provide a quick, self-adaptive, efficient fireworks approach for large-scale optimization. An effective multi-objective metaheuristic algorithm for the sustainable harvest planning issue was considered by Fathollahi-Fard et al. (35).

4.1. Augmented Epsilon-Constraint Method The augmented epsilon-constraint method is a technique used in multi-objective optimization to handle constraints in the optimization process. Multi-objective optimization involves optimizing multiple conflicting objectives simultaneously, and constraints are conditions that must be satisfied for a solution to be considered feasible.

In the augmented epsilon-constraint method, the idea is to transform the constrained multi-objective optimization problem into an unconstrained one by introducing additional variables and constraints. The method is particularly useful when dealing with problems where finding feasible solutions is challenging.

Here's a basic overview of the augmented epsilon-constraint method:

- **Original Problem:** Let's say you have a multi-objective optimization problem with objectives $f_1(x), f_2(x), \dots, f_m(x)$, and constraint functions, $g_1(x), g_2(x), \dots, g_p(x)$ where x is the vector of decision variables.

- **Introduce Slack Variables:** Introduce slack variables $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)$ to represent the violation of each constraint. These slack variables are non-negative and measure how much a solution violates a particular constraint.

- **Transform Constraints:** Transform the original constraints into equality constraints using the slack variables. The transformed constraints may look like this:

$$g_i(x) + \varepsilon_i = 0 \quad i=1,2,\dots,p \quad \text{where } \varepsilon_i \geq 0$$

- **Augmented Objective Function:** Modify the objective functions to penalize violations of constraints. The augmented objective function may include a penalty term that depends on the slack variables:

$$F(x) = (f_1(x), f_2(x), \dots, f_m(x)), \lambda_1 \varepsilon_1, \lambda_2 \varepsilon_2, \dots, \lambda_p \varepsilon_p$$

Here, $\lambda_1, \lambda_2, \dots, \lambda_p$ are penalty coefficients.

- **Optimization:** Solve the augmented unconstrained problem using a multi-objective optimization algorithm. The algorithm seeks to optimize the augmented objective function, and the penalty terms encourage the optimization process to find solutions that minimize violations of constraints.

- **Post-Processing:** After obtaining solutions from the augmented problem, analyze the trade-offs between conflicting objectives and check the values of slack variables to ensure constraint satisfaction.

The augmented epsilon-constraint method helps convert a constrained multi-objective optimization problem into a form that can be addressed by standard multi-objective optimization algorithms. This approach allows for a more flexible and efficient exploration of the solution space in the presence of constraints (36).

Hybrid multiobjective optimization problems involve single-objective optimization using multicriteria decision-making methods and single-objective evolutionary algorithms (SOEA) like NSGA-II, NPGA, and MOPSO for finding Pareto optimal fronts in a single simulation run.

4.2. NSGA-II Deb et al. (37) developed the NSGA-II, a GA-based multi-objective optimization technique. They created a random parent population of size ($nPop$), assessed objective values, and sorted using the nondomination method. They selected two individuals and formed a new offspring population with $nPop + n$ sizes. The NSGA-II implementation produced non-dominated Pareto-optimum solutions.

4.3. NPGA The NPGA works identically to NSGA-II, except for selecting and reproducing the parents in the mating pool. Before using the Pareto-based population-ranking approach, one of the fronts is chosen using the ranked-based roulette wheel (RBRW) selection operator,

initially proposed by Al Jadaan et al. (38, 39). The Pseudo-code of the NSGA-II stated in Figure 3.

The next step is to choose one option from the candidate front using the same method. Consequently, the chances of picking a solution inside the best non-dominated set of the initial front are greatest. In contrast, solutions inside a set of the second front have lower probability, and so on.

4. 4. Algorithms' Characteristics

4. 4. 1. Chromosome Structure Chromosomes are gene collections arranged in a specific order. Designing an appropriate chromosomal structure is crucial for algorithm execution. The study's chromosomal solution is represented as a $2n$ matrix, displaying storage space, low supply, shortfall, and storage locations for each commodity. Figure 4 shows this graphically.

4. 5. Algorithms' Mechanism This part explains the crossover, mutation, evaluation function, and stop criteria of the two algorithms' four key characteristics.

4. 5. 1. Crossover Operator A crossover process involves DNA switching between parent chromosomes, creating better chromosomes and favorable genes for offspring. This study generates offspring using linear chromosome vectors and arithmetic crossover operators with a random weighting factor

$$offspring_1 = \alpha \times parent_1 + (1 - \alpha) \times parent_2 \quad (29)$$

$$offspring_2 = (1 - \alpha) \times parent_1 + \alpha \times parent_2 \quad (30)$$

4. 5. 2. Mutation Operator Mutation preserves

-
1. Create N_{pop} random solutions (Initialization)
 2. Determine the values of the objective function for the first solutions.
 3. Determine rankings for the solutions using Goldberg's ranking method
 4. Determine the crowding distance.
 5. While stop requirements are not met
 - a. Create the mating pool and add individuals to it using binary tournament selection.
 - b. To the mating pool, apply the crossover and mutation operators.
 - c. To the new solutions' objective function values.
 - d. Combine the existing populace with the newly developed solutions.
 - e. Determine rank using the Goldberg's ranking method.
 - f. Determine the crowding distance.
 - g. Group people and choose superior options.
-
- Stop while

Figure 3. Pseudo-code of the NSGA-II

$$\begin{bmatrix} X_i [X_1 X_2 \dots X_n] \\ S_i [S_1 S_2 \dots S_n] \end{bmatrix}$$

Figure 4. Structure of a chromosome

genetic diversity in populations, protecting against data loss. In this regard, firstly, we chose a normal random variable, then by substituting in Expression 32 regarding γ between zero and one, we obtain β after that if β is positive Expression 34 is regarded else Expression 36 is taken into account. Note that Upperbound is one and lowerbound is zero; they are updated during algorithm processing.

$$z \sim Normal(0,1) \quad (31)$$

$$\beta = \tanh(\gamma \times z) \quad (32)$$

$$\text{if } \beta \geq 0 \quad (33)$$

$$offspring = parent + \beta \times (Upperbound - parent) \quad (34)$$

$$\text{else} \quad (35)$$

$$offspring = parent + \beta \times (parent - Lowerbound) \quad (36)$$

4. 5. 3. Controlled Elitism In the NSGA-II, elitism is regulated to balance maintaining high-quality solutions and preserving diversity in the population. This is crucial because excessive elitism can lead to premature convergence (where the algorithm converges to sub-optimal solutions early). In contrast, insufficient elitism can lead to a loss of good solutions.

4. 5. 4. Evaluation The fitness of chromosomes in each generation is assessed using the optimization model's objective function. However, the inventory model has four limitations, making synthetic chromosomes unlikely. A penalty function is used to increase the likelihood of constraint violations, defining the penalty and fitness function.

$$Penalty\ function = \begin{cases} \sum_{i=1}^4 \frac{E_i}{4} & \text{If the chromosome} \in \text{inf feasible region} \\ 0 & \text{If the chromosome} \in \text{feasible region} \end{cases} \quad (37)$$

$$\text{Fitness function} = \text{Penalty function} + \text{Objective function}$$

4. 5. 5. Stopping Criterion A stopping condition in the MOEA leads to Pareto-optimal solutions, eliminating the need for mutation and crossover operators. After a certain number of generations, algorithms are believed to be finished. Finding an early perfect solution requires both the NSGA-II and NPGA algorithms in MATLAB 8.20 and statistical techniques.

5. NUMERICAL RESULTS

In this step, twenty test problems are regarded, ranging from the size of the two products to twenty products. Also, related parameters are generated randomly following uniform distribution between associated bounds. Table 1 displays the ranges of numerical data. Firstly, we report solutions obtained by GAMS commercial software for all twenty test problems as the best-quality solutions. For the augmented epsilon-constraint method, this approach outperforms two other solution approaches regarding the quality metric index. Thus, we did not consider it. The number of Pareto front solutions oscillates between 400 and 500. However, for other criteria, no specific trend exists. However, the Pareto front curve is the best compared to other methods, and convexity is outward.

To perform in the best condition regarding two metaheuristic approaches. We chose six factors for each of the three levels. Table 3 shows the factors and related levels. Then, using Minitab software, we generated orthogonal arrays, and for specific levels, we obtained the metrics mentioned above. We obtained weighted values

after normalizing the solution and determining positive or negative indexes. At last, the distance from the maximum value of weighted sum values is computed. The distance from the maximum as the best solution should be minimized. Tables 4 to 5 show the numerical results, and Figures 5 to 6 show the optimum level of regarded factors (less response value is better).

We used the Topsis approach, a multi-attribute decision-making approach, to decide which algorithm was better. Results show that the NPGA outperforms the NSGA-II regarding related indexes. For further information refer to literature [32].

TABLE 1. Parameter range

$A \sim$ uniform [500 1900]	$m \sim$ uniform [0.02 0.04]	$b \sim$ uniform [5 33]
$c \sim$ uniform [6 34]	$u \sim$ uniform [0.05 0.25]	$l \sim$ uniform [1 29]
$d \sim$ uniform [150 1000]	$k \sim$ uniform [1 15]	$ts \sim$ uniform [0.0003 0.0007]
$v \sim$ uniform [0.5 0.7]	$h \sim$ uniform [2 30]	$f \sim$ uniform [2 5]
$sc \sim$ uniform [0.045 0.065]	$p \sim$ uniform [5000 12000]	$o \sim$ uniform [3 10]
W=20000000 SL=0.99999		

TABLE 2. Problems vs. obtained indexes related to the augmented epsilon-constraints method executed by GAMS commercial software

Problem	Pareto front metric		Objective function mean values		Computational time
	No. of non-dominated solutions	Spacing metric	Mean z_1	Mean z_2	
1	429	6.0161	33020.6	11022.1	711
2	491	6.2039	68276.68	7271.221	291
3	329	0.08952	109109	13461.52	229
4	492	5.7082	126659.1	17500.58	270
5	490	6.0684	156536.3	20625.78	256
6	389	7.2765	190277.9	34970.78	271
7	490	3.927	223838.8	32008.53	309
8	492	4.1316	225652.9	25609.38	345
9	491	3.1453	206231.8	30771.12	318
10	2	0	151504.9	145595.4	37
11	1	0	176127	1119.57	20
12	1	0	205488.2	1190.5	20
13	1	0	177252.2	1432.92	20
14			Infeasible		
15	492	2.1907	351029.9	49065.39	338
16	490	1.856	401399.8	54128.24	435
17	489	2.0345	449593.3	59183.4	539
18	491	1.8456	424752	77587.54	533
19	489	1.563	521761.7	62907.14	417
20	490	1.6378	470140	62391.54	332

TABLE 3. Considered structural parameters and levels of both algorithms

Factors	Levels		
	1	2	3
nPop	30	50	70
MaxIt	30	50	70
pCrossover	0.3	0.5	0.7
pMutation	0.3	0.5	0.7
Gamma	0.3	0.5	0.7
ControlledElitismParameter	0.3	0.5	0.7

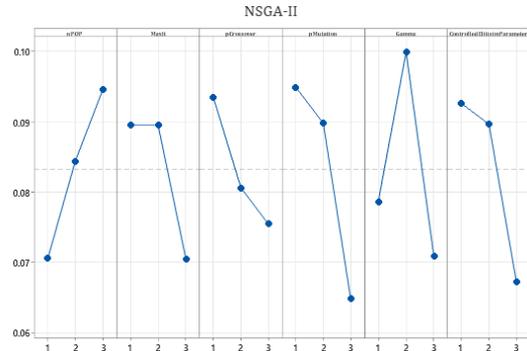


Figure 5. Levels vs. response values related to the NSGA-II algorithm (less response value is better)

TABLE 4. Orthogonal arrays vs. obtained indexes related to the NSGA-II

Array	Pareto front metric		QM	Objective function mean values		Computational time
	Number of non-dominated solutions	Spacing metric		Mean z ₁	Mean z ₂	
1	18	2.18	1	5.37E+06	1.73E+05	2.385609
2	6	1.92	0	5.75E+06	1.51E+05	2.827144
3	6	1.35	1	4.59E+06	1.56E+05	2.868053
4	7	1.25	0	3.65E+06	1.27E+05	5.840853
5	7	1.74	0	3.39E+06	1.22E+05	6.133792
6	4	0.44	1	2.00E+06	7.32E+04	6.862424
7	4	0.94	1	2.19E+06	7.37E+04	12.0195
8	2	0	0	1.21E+06	3.96E+04	12.77704
9	1	0	1	6.85E+05	2.38E+04	13.75613
10	15	0.57	0	3.51E+06	1.27E+05	10.614
11	1	0	1	2.37E+06	8.86E+04	11.75421
12	12	1.226	0	4.49E+06	1.47E+05	13.51858
13	19	2.38	0	4.20E+06	1.52E+05	13.30175
14	7	1.59	1	3.28E+06	1.03E+05	14.68792
15	6	0.93	0	2.80E+06	8.92E+04	15.39808
16	4	0.63	1	2.54E+06	8.74E+04	18.23107
17	5	1.75	0	3.07E+06	1.09E+05	17.88389
18	1	0	0	1.48E+06	4.86E+04	21.27136
19	7	1.47	1	3.87E+06	1.37E+05	18.79754
20	2	0	0.22	3.17E+06	1.11E+05	19.69584
21	4	1.076	0.2	3.00E+06	1.01E+05	19.88637
22	1	0	1	2.31E+06	7.88E+04	29.65523
23	4	1.21	0	2.21E+06	7.79E+04	29.67642
24	2	0	0	1.42E+06	5.02E+04	32.4939
25	11	1.13	0	3.13E+06	1.05E+05	32.45348
26	4	0.066	0	2.38E+06	8.18E+04	34.08444
27	2	0	1	9.92E+05	3.22E+04	42.43041

TABLE 5. Orthogonal arrays vs. indexes related to the NRGa

Array	Pareto front metric		Objective function mean values			Computational time
	Number of non-dominated solutions	Spacing metric	QM	Mean z_1	Mean z_2	
1	21	2.14	0	7.01E+05	1.88E+05	2.628553
2	12	1.39	1	4.66E+05	1.40E+05	2.805269
3	8	0.88	0	6.22E+05	1.92E+05	2.751699
4	16	0.913	1	2.71E+05	9.74E+04	6.375257
5	2	0	1	3.38E+05	1.02E+05	7.816881
6	8	2.39	0	4.64E+05	1.44E+05	6.523603
7	3	1.7	0	3.10E+05	9.70E+04	14.80842
8	2	0	1	1.06E+05	3.51E+04	13.41777
9	1	0	0	7.08E+04	2.39E+04	15.72957
10	5	1.11	1	2.93E+05	1.04E+05	12.59192
11	3	0.006	0	4.04E+05	1.32E+05	15.59124
12	4	0.88	1	2.96E+05	1.03E+05	13.4974
13	12	1.53	1	4.04E+05	1.26E+05	17.41597
14	11	1.43	0	3.94E+05	1.24E+05	15.31773
15	1	0	1	2.31E+05	7.77E+04	15.9429
16	2	0	0	2.65E+05	9.03E+04	18.27887
17	1	0	1	1.79E+05	6.76E+04	23.95663
18	2	0	1	7.97E+04	2.63E+04	32.82435
19	22	0.71	0	4.53E+05	1.54E+05	22.75148
20	8	1.42	0.77	2.95E+05	1.14E+05	24.62209
21	4	1.92	0.8	3.01E+05	9.90E+04	21.58282
22	4	0.94	0	2.67E+05	9.48E+04	34.73219
23	1	0	1	1.10E+05	3.49E+04	42.13635
24	1	0	1	9.14E+04	2.80E+04	40.36455
25	8	0.55	1	2.41E+05	8.30E+04	37.56112
26	2	0	1	1.13E+05	3.77E+04	39.72398
27	2	0	0	1.19E+05	3.82E+04	53.15664

TABLE 6. Problems vs. obtained indexes related to the NSGA-II

Problem	Pareto front metric		Objective function mean values			Computational time
	Number of non-dominated solutions	Spacing metric	QM	Mean z_1	Mean z_2	
1	3	1.5089	1	1.72E+05	6.97E+03	11.99048
2	10	0.75766	0.21622	9.80E+05	1.27E+04	11.09406
3	11	0.88645	0.44	4.25E+06	1.28E+04	13.00945
4	13	0.80206	0.20588	3.90E+06	2.34E+04	12.14553
5	12	1.0162	0.083333	4.89E+06	2.37E+04	13.08017
6	9	0.87584	0.29167	6.70E+06	3.27E+04	11.87017
7	11	0.95535	0.19355	9.09E+06	3.91E+04	11.83453
8	11	0.84226	0.91667	6.78E+06	31495.19	12.17597
9	14	1.1263	0.34286	7.07E+06	46441.94	12.5615

10	10	0.69828	0.34783	9.34E+06	55771.99	12.04777
11	11	1.0172	1	1.11E+07	59469.65	12.2284
12	11	1.1176	1	1.19E+07	59824.82	12.26202
13	10	0.79429	1	1.37E+07	67954.2	12.0311
14	12	0.71965	1	2.00E+07	73020.04	12.36023
15	9	1.364	0.29412	2.52E+07	90003.51	11.13272
16	12	1.8151	0.90909	2.67E+07	103629.3	11.90267
17	12	0.7592	0	2.27E+07	106882.5	11.68452
18	10	0.61745	1	1.89E+07	115649.1	12.14915
19	10	1.2073	0	3.96E+07	112046.6	11.08706
20	9	0.37443	0.28	3.46E+07	126621.2	11.36034

TABLE 7. Problems vs. obtained indexes related to the NRGGA

Problem	Pareto front metric		Objective function mean values			Computational time
	Number of non-dominated solutions	Spacing metric	QM	Mean z_1	Mean z_2	
1	6	1.4405	0	17220.22	6971.906	34.33418
2	36	2.2765	0.78378	97960.15	12584.02	20.96036
3	36	0.93509	0.56	415115.9	13117.35	20.67863
4	35	1.4365	0.79412	394099.5	22696.6	21.08452
5	35	1.4198	0.91667	485820.9	23265.65	21.11902
6	34	0.83329	0.70833	678311.7	31882.03	21.18444
7	35	0.95884	0.80645	882258.1	39061.81	21.10709
8	33	1.5107	0.083333	671786.3	32377.67	21.0122
9	35	0.86176	0.65714	680924.6	47891.94	20.27082
10	35	1.0508	0.65217	950517.3	56985.32	20.27982
11	34	1.4061	0	1135262	66045.51	20.00956
12	33	1.1839	0	1231447	64292.26	20.86337
13	36	1.9886	0	1462616	77532.3	19.67027
14	29	2.0254	0	2152912	91681.01	19.55916
15	33	2.5172	0.70588	2604750	89019.46	19.9761
16	34	2.4407	0.090909	2675543	109897.8	19.36939
17	7	2.3533	1	1949359	88772.28	20.74752
18	15	1.9437	0	2098233	115609.2	19.69793
19	3	1.1994	1	3200771	104068.1	20.23629
20	32	2.1377	0.72	2996309	134638.6	18.99066

TABLE 8. Topsis final score decision matrix

Alternatives	Indexes					
	No. of non-dominated solutions	Spacing metric	QM	Mean z_1	Mean z_2	Computational time
NSGA-II	0.35	0.52	0.74	0.72	0.7	0.495
NRGA	0.93	0.86	0.67	0.69	0.72	0.86

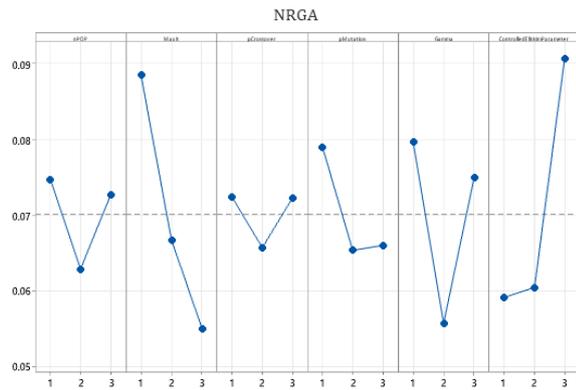


Figure 6. Levels vs. response values related to the NRGA algorithm (less response value is better)

Figure 7 shows the obtained Pareto front by three approaches. Solving problem shows that the best quality is related to the GAMS commercial software (the augmented epsilon-constraint). The NSGA-II yields more Pareto solutions; however, NRGA solutions are closest to the ideal point (origin). As a whole, the conflict between objective functions is strongly evident.

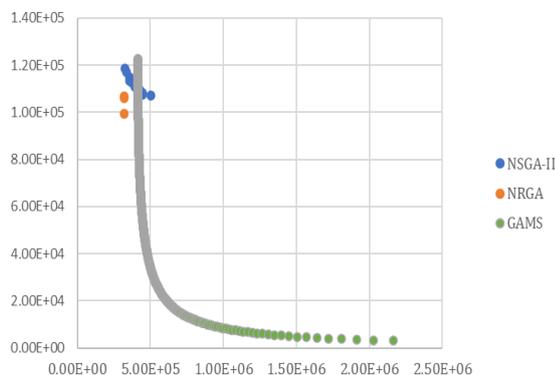


Figure 7. Obtained Pareto front regarding two meta-heuristic algorithms (NSGA-II and NRGA) and GAMS (augmented epsilon-constraint method)

6. CONCLUDING REMARKS AND SUGGESTIONS

This study utilized a partial back ordering, rework, and garbage EPQ model to minimize manufacturing facility costs and warehouse space supply costs. We investigate two separate objective functions regarding operational constraints. Regarding the research literature gap, we contributed our novelty (developing the mathematical model and using solution approaches). First, the best solutions obtained by GAMS software were presented using the augmented epsilon-constraint method. Due to the complexity of the related model, The model was solved using the NSGA-II and NRGA approaches.

Various Pareto front indexes were considered, and we tuned the structural parameters of these two algorithms using the Taguchi method. Numerous test problems were investigated and reported all of them meticulously. The obtained Pareto front confirms the conflict between objective functions. Topsis method was used to specify the best approaches. As a whole, GAMS yields the best quality solutions. However, in a large size, the two aforementioned algorithms are justifiable.

Future research should consider using various meta-heuristics, comparing performance metrics, accounting for uncertain parameters, and considering multi-product systems with multiple stages and product limits.

7. REFERENCES

- Cunha LRA, Delfino APS, dos Reis KA, Leiras A. Economic production quantity (EPQ) model with partial backordering and a discount for imperfect quality batches. *International Journal of Production Research*. 2018;56(18):6279-93. <https://doi.org/10.1080/00207543.2018.1445878>
- Shah NH, Vaghela CR. Imperfect production inventory model for time and effort dependent demand under inflation and maximum reliability. *International Journal of Systems Science: Operations & Logistics*. 2018;5(1):60-8. <https://doi.org/10.1080/23302674.2016.1229076>
- Taleizadeh AA, Soleymanfar VR, Govindan K. Sustainable economic production quantity models for inventory systems with shortage. *Journal of cleaner production*. 2018;174:1011-20. <https://doi.org/10.1016/j.jclepro.2017.10.222>
- Al-Salamah M. Economic production quantity in an imperfect manufacturing process with synchronous and asynchronous flexible rework rates. *Operations research perspectives*. 2019;6:100103. <https://doi.org/10.1016/j.orp.2019.100103>
- Lin SS-C. Note on "The derivation of EOQ/EPQ inventory models with two backorders costs using analytic geometry and algebra". *Applied Mathematical Modelling*. 2019;73:378-86. <https://doi.org/10.1016/j.apm.2019.04.025>
- Marchi B, Zanoni S, Jaber MY. Economic production quantity model with learning in production, quality, reliability and energy efficiency. *Computers & Industrial Engineering*. 2019;129:502-11. <https://doi.org/10.1016/j.cie.2019.02.009>
- Thinakaran N, Jayaprakas J, Elanchezian C. Survey on inventory model of EOQ & EPQ with partial backorder problems. *Materials Today: Proceedings*. 2019;16:629-35. <https://doi.org/10.1016/j.matpr.2019.05.138>
- Zavanella LE, Marchi B, Zanoni S, Ferretti I. Energy considerations for the economic production quantity and the joint economic lot sizing. *Journal of business economics*. 2019;89(7):845-65. <https://doi.org/10.1007/S11573-019-00933-6>
- De SK, Nayak PK, Khan A, Bhattacharya K, Smarandache F. Solution of an EPQ model for imperfect production process under game and neutrosophic fuzzy approach. *Applied Soft Computing*. 2020;93:106397. <https://doi.org/10.1016/j.asoc.2020.106397>
- Ganesan S, Uthayakumar R. EPQ models for an imperfect manufacturing system considering warm-up production run, shortages during hybrid maintenance period and partial backordering. *Advances in Industrial and Manufacturing Engineering*. 2020;1:100005. <https://doi.org/10.1016/j.aime.2020.100005>

11. Guha A, Bose D. A note on “Economic production quantity in batch manufacturing with imperfect quality, imperfect inspection, and destructive and non-destructive acceptance sampling in a two-tier market”. *Computers & Industrial Engineering*. 2020;146:106609. <https://doi.org/10.1016/j.cie.2020.106609>
12. Hauck Z, Rabta B, Reiner G. Impact of early inspection on the performance of production systems—insights from an EPQ model. *Applied Mathematical Modelling*. 2022;107:670-87. <https://doi.org/10.1016/j.apm.2022.03.003>
13. Kalantari SS, Taleizadeh AA. Mathematical modelling for determining the replenishment policy for deteriorating items in an EPQ model with multiple shipments. *International Journal of Systems Science: Operations & Logistics*. 2020;7(2):164-71. <https://doi.org/10.1080/23302674.2018.1542753>
14. Nobil AH, Sedigh AHA, Cárdenas-Barrón LE. A multiproduct single machine economic production quantity (EPQ) inventory model with discrete delivery order, joint production policy and budget constraints. *Annals of Operations Research*. 2020;286:265-301. <https://doi.org/10.1007/S10479-017-2650-9>
15. Rahaman M, Mondal SP, Shaikh AA, Ahmadian A, Senu N, Salahshour S. Arbitrary-order economic production quantity model with and without deterioration: generalized point of view. *Advances in Difference Equations*. 2020;2020:1-30. <https://doi.org/10.1186/S13662-019-2465-X/FULLTEXT.HTML>
16. Rahaman M, Mondal SP, Shaikh AA, Pramanik P, Roy S, Maiti MK, et al. Artificial bee colony optimization-inspired synergetic study of fractional-order economic production quantity model. *Soft Computing*. 2020;24:15341-59. <https://doi.org/10.1007/S00500-020-04867-Y>
17. Bose D, Guha A. Economic production lot sizing under imperfect quality, on-line inspection, and inspection errors: Full vs. sampling inspection. *Computers & Industrial Engineering*. 2021;160:107565. <https://doi.org/10.1016/j.cie.2021.107565>
18. Kalaiarasi K, Daisy S, Sumathi M. WITHDRAWN: Solving a multi product single machine EPQ inventory model with GP mode: By using python. Elsevier; 2021.
19. Moghdani R, Sana SS, Shahbandarzadeh H. Multi-item fuzzy economic production quantity model with multiple deliveries. *Soft Computing*. 2020;24(14):10363-87. <https://doi.org/10.1007/S00500-019-04539-6>
20. Shekhawat S, Rathore H, Sharma K. Economic production quantity model for deteriorating items with weibull deterioration rate over the finite time horizon. *International Journal of Applied and Computational Mathematics*. 2021;7:1-21. <https://doi.org/10.1007/S40819-021-00972-0>
21. Kalaiarasi K, Begum MS, Sumathi M. WITHDRAWN: Optimization of unconstrained multi-item (EPQ) model using fuzzy geometric programming with varying fuzzification and defuzzification methods by applying python. Elsevier; 2021.
22. Nobil AH, Niaki STA, Niaki SAA, Cárdenas-Barrón LE. An economic production quantity inventory model for multi-product imperfect production system with setup time/cost function. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales Serie A Matemáticas*. 2022;116:1-23. <https://doi.org/10.1007/S13398-021-01182-5>
23. Priyan S, Mala P, Palanivel M. A cleaner EPQ inventory model involving synchronous and asynchronous rework process with green technology investment. *Cleaner Logistics and Supply Chain*. 2022;4:100056. <https://doi.org/10.1016/j.clscn.2022.100056>
24. Edalatpour MA, Mirzapour Al-e-Hashem SMJ, Fathollahi-Fard AM. Combination of pricing and inventory policies for deteriorating products with sustainability considerations. *Environment, Development and Sustainability*. 2023;1-41. <https://doi.org/10.1007/S10668-023-02988-6>
25. Poursoltan L, Seyedhosseini S, Jabbarzadeh A. An extension to the economic production quantity problem with deteriorating products considering random machine breakdown and stochastic repair time. *International Journal of Engineering, Transactions B: Applications*. 2020;33(8):1567-78. 10.5829/IJE.2020.33.08B.15
26. Deiranlou M, Dehghanian F, Pirayesh MA. The simultaneous effect of holding safety stock and purchasing policies on the economic production quantity model subject to random machine breakdown. *International Journal of Engineering, Transactions B: Applications*. 2019;32(11):1643-55. <https://doi.org/10.5829/ije.2019.32.11b.16>
27. Jain M, Sharma G, Rathore S. Economic production quantity models with shortage, price and stock-dependent demand for deteriorating items. *International Journal of Engineering, Transactions A: Basics*. 2007;20(2):159-68.
28. Pattnaik S, Nayak MM, Abbate S, Centobelli P. Recent trends in sustainable inventory models: A literature review. *Sustainability*. 2021;13(21):11756. <https://doi.org/10.3390/SU132111756>
29. Dulebenets MA. A Diffused Memetic Optimizer for reactive berth allocation and scheduling at marine container terminals in response to disruptions. *Swarm and Evolutionary Computation*. 2023;80:101334. <https://doi.org/10.1016/J.SWEVO.2023.101334>
30. Dulebenets MA. An Adaptive Polyploid Memetic Algorithm for scheduling trucks at a cross-docking terminal. *Information Sciences*. 2021;565:390-421. <https://doi.org/10.1016/J.AEI.2022.101623>
31. Pasha J, Nwodu AL, Fathollahi-Fard AM, Tian G, Li Z, Wang H, et al. Exact and metaheuristic algorithms for the vehicle routing problem with a factory-in-a-box in multi-objective settings. *Advanced Engineering Informatics*. 2022;52:101623. <https://doi.org/10.1145/3583131.3590367>
32. Singh E, Pillay N, editors. A study of ant-based pheromone spaces for generation perturbative hyper-heuristics. *Proceedings of the Genetic and Evolutionary Computation Conference; 2023*.
33. Singh P, Pasha J, Moses R, Sobanjo J, Ozguven EE, Dulebenets MA. Development of exact and heuristic optimization methods for safety improvement projects at level crossings under conflicting objectives. *Reliability Engineering & System Safety*. 2022;220:108296. <https://doi.org/10.1016/J.SWEVO.2023.101314>
34. Chen M, Tan Y. SF-FWA: A Self-Adaptive Fast Fireworks Algorithm for effective large-scale optimization. *Swarm and Evolutionary Computation*. 2023;80:101314. <https://doi.org/10.1016/J.COR.2023.106304>
35. Fathollahi-Fard AM, Tian G, Ke H, Fu Y, Wong KY. Efficient multi-objective metaheuristic algorithm for sustainable harvest planning problem. *Computers & Operations Research*. 2023;158:106304.
36. Greco S, Figueira J, Ehrgott M. *Multiple criteria decision analysis*: Springer; 2016.
37. Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation*. 2002;6(2):182-97. 10.1109/4235.996017
38. Al Jadaan O, Rao CR, Rajamani L, editors. Solving constrained multi-objective optimization problems using non-dominated ranked genetic algorithm. 2009 Third Asia International Conference on Modelling & Simulation; 2009: IEEE. 10.1109/AMS.2009.38
39. Al Jadaan O, Rao C, Rajamani L, editors. Parametric study to enhance genetic algorithm performance, using ranked based roulette wheel selection method. *International conference on multidisciplinary information sciences and technology (InSciT2006)*; 2006.

COPYRIGHTS

©2024 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**

مدل مقدار تولید اقتصادی (EPQ) نرخ تولید، نرخ تقاضا، هزینه‌های راه‌اندازی، هزینه‌های نگهداری و هزینه‌های کمبود را در نظر می‌گیرد تا مقدار تولیدی را بیابد که مجموع این هزینه‌ها را به حداقل برساند. هدف متعادل کردن هزینه‌های مرتبط با تولید، نگهداری موجودی و کمبودهای احتمالی است. در این مقاله، دو هدف شامل هزینه‌های تولید و سفارش مورد نظر است در هدف دوم هزینه فضای ذخیره‌سازی قرار است تا به حداقل برسد. این مطالعه ضایعات و کار مجدد را در مدل EPQ در نظر می‌گیرد. مدل موجودی برای اقلام در یک ماشین مورد نظر است. ظرفیت تولید کاهش می‌یابد و زمانی که فقط یک دستگاه وجود داشته باشد کمبودهایی وجود دارد. با تعیین مقدار محصولات تولید شده توسط واحد تولیدی، فضای ذخیره‌سازی برای هر محصول، زمان چرخه و کمیاب محصول، می‌توان هم هزینه کلی و هم هزینه تامین فضای انبار و کاهش داد. برای حل از نرم افزار تجاری و در ابعاد بزرگ، از یک الگوریتم فراابتکاری چندهدفه، یعنی الگوریتم ژنتیک مرتب سازی غیر غالب (NSGA-II) استفاده شده است. یافته‌ها با استفاده از الگوریتم ژنتیک رتبه‌بندی غیرمسلط (NRGA) اعتبار بیشتری دارند. همچنین جبهه پارتو به دست آمده با چند شاخص مورد مطالعه قرار می‌گیرد. برای اجرای این دو الگوریتم در بهترین شرایط، از رویکرد تاگوچی و آرایه‌های متعامد مربوطه استفاده شد. همچنین برای اعتبارسنجی مدل ریاضی، از روش محدودیت اسپیلون ادغام شده، اجرا شده در محیط GAMS استفاده کردیم. واضح است که نرم افزار تجاری نتایج بهتری به همراه دارد، اما استفاده از این دو الگوریتم زمانی قابل توجیه هستند که ابعاد بزرگتر شود. در نهایت با انجام تحلیل حساسیت این شاخص‌ها و توابع هدف، رفتار الگوریتم‌های پیشنهادی با جزئیات مقایسه و بررسی می‌شود. همچنین الگوریتم برتر با استفاده از روش Topsis به عنوان یک روش تصمیم‌گیری چند معیاره انتخاب شده است. مثال‌های عددی نشان می‌دهد که چگونه مدل ارائه شده و الگوریتم‌های پیشنهادی ممکن است به طور موثر مورد استفاده قرار گیرند. بررسی ادبیات روشن می‌کند که توابع هدف مرتبط، محدودیت‌ها و رویکردهای راه‌حل تاکنون مورد بررسی قرار نگرفته‌اند.



Fuzzy Dynamic Modeling for Export Consortia in Small and Medium-Sized Enterprises

F. Shamshiri, P. Shahnazari-Shahrezaei*, M. Fallah, H. Kazemipour

Department of Industrial Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran

PAPER INFO

Paper history:

Received 10 October 2023

Received in revised form 15 December 2023

Accepted 23 December 2023

Keywords:

Export Consortia

Small and Medium-sized Enterprises

System Dynamics

Fuzzy Inference System

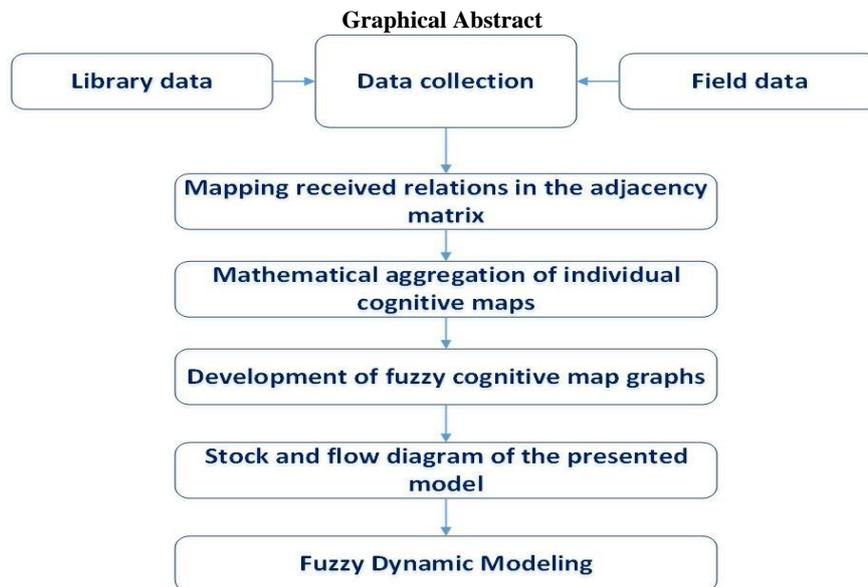
Fuzzy Cognitive Map

Grounded Theory

ABSTRACT

The absence of active export consortia and the lack of a technical, serious, and codified plan for their development are among the most important reasons for Iran's small and medium-sized enterprises (EMSs) remaining in the country's export coordinates. In this study, the data are collected and analyzed with a mixed (qualitative-quantitative) approach, which is a critical paradigm. The data are collected using library research and field methods. In the field section, structured, exploratory, and collaborative interviews are used in the qualitative phase, and the researcher-made questionnaires are used in the quantitative phase. The data are analyzed using grounded theory, brainstorming sessions, fuzzy cognitive map (FCM), fuzzy inference system (FIS), and system dynamics modeling (SDM). According to the results, "features of consortium members", "export operational plan", "consortium strengthening factor", "recognition of export support", "transnational factors", "government factors", and "product features" are the seven main success factors of private sector export consortia in Iranian industries. Furthermore, identifying a suitable promoter, identifying potential members, conducting the desired study and contacting interested companies, appointing representatives, holding meetings between potential members, conducting a feasibility study and preparing a business plan, officially forming a consortium, and following up on consortium affairs are eight steps for establishing private sector export consortiums in Iranian industries.

doi: 10.5829/ije.2024.37.07a.19



*Corresponding Author Email: parisa_shahnazari@yahoo.com (P. Shahnazari-Shahrezaei)

1. INTRODUCTION

Small and Medium Enterprises (SME) are one of the most important factors of economic development and their importance in regional economic growth (1). National development in today's world cannot be imagined without considering foreign trade. According to the review of the economic and financial developments of the world after the industrial revolution, the developed countries managed to enhance their political and economic power by using the important and influential part of export. Exports contribute significantly to the development of a country's economy in the present era, and all governments are aware of its economic impact (2). Nowadays, small and medium-sized enterprises (SMEs) have to deal dynamically and innovatively with the challenges caused by the liberalization of trade worldwide and globalization. SMEs have a low export share and low competitiveness to be present in foreign markets, especially in the face of large market powers, due to their small sizes (3, 4).

There are usually problems faced by SMEs in exporting to foreign markets. SMEs may lack the necessary knowledge and financing, fail to meet foreign regulatory requirements, and produce products of inadequate quantity or quality for foreign buyers among many other potential problems (5). Accordingly, they must take measures to survive in this fast-paced movement. Cooperation between SMEs and the formation of export consortia are some of the solutions to these problems (6). SMEs can significantly increase their export potential and reduce the costs and risks of penetrating international markets by combining knowledge, financial resources, and connections in the form of an export consortium. Export consortia provide a wide range of services from basic administrative services to those that improve the production structure in the form of medium-term to long-term strategic cooperation to provide specialized services to facilitate access to foreign markets (7). Lack of proper organizational support, weak judicial and legal systems, non-adherence to written and approved regulations, and increasing costs caused by uncertainty and lack of credibility can be solved by forming consortia (8-11). Service-oriented approaches should prevail, current and emerging markets should be developed, a new organizational system should be adopted, and business networks in the country should be investigated, created, analyzed, and developed due to recent economic developments and to take advantage of their benefits (12). Thus, the export capability of SMEs can be improved and the basis for developing and expanding their activities, generating income, improving the

quantity and quality of products, increasing exports, and reducing unemployment can be provided by a better understanding of the main success factors of SME export consortia and having a conceptual model for their development.

If consortium members pool their resources to acquire equipment, systems, and services (marketing, logistics, training, technical consulting, etc.), economies of scale may emerge. So, as a group, they increase the bargaining power, enabling access to products and services under better conditions. Besides, a common definition should be provided, and "regulations of production" should be developed (to achieve quality certificates representing significant achievements in terms of added value) (13). Continuous information exchange between SMEs (for example, HRM practices) directly contributes to their competitiveness. The results of conventional SME upgrading programs can be enhanced through export consortia because cost reductions, economies of scale, and multiplier effects can significantly increase the number of SMEs benefiting from modernization initiatives (14). Export consortia act as powerful tools to strengthen the relationship between the value chains of member companies and increase competition. Moreover, in an export consortium, strengthening competitive advantages in export markets mainly leads to the expansion of local market share¹. Many countries have promoted consortia and other forms of networks among SMEs due to the importance of networks for the growth and competitiveness of SMEs. The initial network and development of trust between members can be facilitated through public agencies (15). As facilitators, government trade promotion agencies act consciously in helping SMEs, contribute to their goals, and follow a legal nature accordingly. However, the commitment and active participation of the consortium members cannot be replaced by external incentives to establish a consortium. One of the necessary conditions for the emergence and development of effective networks is a positive understanding of the results of network establishment (6).

SMEs form the basis of Iran's economy. Due to their limited activities, SMEs naturally cannot take advantage of the cost-effectiveness of production. Private sector agencies are suitable and reliable platforms for establishing different consortia, both for production and export. However, they have advantages and disadvantages that should be recognized. Although the field research indicates that there are inconsistent examples of export consortia in the leather, furniture, and carpet industries in Iran. These consortia have taken no successful action. Mass production, securing foreign

¹ [www.unido.org\(2005\)](http://www.unido.org(2005))

exchange income, increasing employment, improving quality, and reducing the price of manufactured products are some of the benefits of strengthening export consortia. Appropriate models should be used to take advantage of export consortia in achieving economic growth and development as much as possible. Identifying the success factors of export consortia and how to establish them and designing a scenario-based model in this field to deal with unforeseen changes are some strategies to strengthen export consortia .

Ortega et al. (16) used fuzzy logic and system dynamics (SD) to deal with uncertainties and ambiguities in epidemic problems such as ambiguity in risk factors, contact patterns, and infection conditions and risks. They used Mamdani's max-min inference method for the multiple input multiple output (MIMO) model and the center of area (COA) defuzzification method to calculate the deterministic output (16). In a study by Chang et al. (17), the applications of fuzzy computing in SDM were investigated and the results were evaluated for the recruitment-customer-producer model. In this model, triangular membership functions were used for the variables "order quantity receipt rate" and "labor productivity". However, these fuzzy variables did not interact with each other in the model, and the combination of fuzzy variables was not reflected in the study (17).

In their study, Ghazanfari et al. (18) investigated the applications of fuzzy logic in the dynamics of systems and provided causal loop diagrams (CLDs) with fuzzy relations. In a study by Xu & Li (19), a conceptual model was proposed for initial values and flow and stock variables using SD and fuzzy optimization. FCMs and qualitative relations were used in SD models in a study by Carvalho & Tome (20). In a study by Herrera et al. (21), a combined approach of fuzzy logic and dynamic system modeling was proposed. Reviewing the literature on the use of fuzzy theory in dynamic systems modeling suggests that although dynamic systems modeling tries to better draw uncertain parameters, it sometimes causes problems (22). For example, the soft variables that are mainly based on the subjective opinions of the modeler are shown using lookup functions. So, many applications of fuzzy theory in SD are focused on modeling imprecise fuzzy parameters and variables. However, they lack a comprehensive approach to use fuzzy inference systems (FISs) as decision-making tools in export consortium dynamics. In addition, previous studies have provided no evidence of implementing the FIS output representation in an SD model. The use of SD models in past studies is mainly limited to drawing mathematical formulas. Thus, the principle of SD is not fully covered .

This study can be considered a model for the success of export consortia to help SMEs to enter international markets. Studies can investigate the application of the SD model based on fuzzy logic in private sector export consortia in Iran in subjective, perceptual, and external

aspects, as well as the necessary factors for the success of such consortia and their performance evaluation. In addition, correct decisions can be made in ambiguous fuzzy environments and a comprehensive view can be presented by simultaneously considering the role of all members with a system approach based on fuzzy logic to export consortia in Iran.

The rest of the study is organized as follows: In section 2, the theoretical foundations and literature review are provided. Section 3 describes the methodology and data analysis techniques. Section 4 presents data analysis and results, and section 5 provides conclusions and recommendations.

2. LITERATURE REVIEW

Researchers have conducted various studies on export consortia over the past years, each of which has made a significant contribution to the development of export consortia literature, developed different concepts and used specific methodologies. The summary of past domestic and foreign studies is given below.

According to investigation presented by Mousavi Neqabi and Anoushe (23) on South Pars Phase 11 agreement was concluded with a consortium of France's Total, China's CNPC, and Iran's Petropars. The first part of the contract includes the construction of two production platforms and two pipelines, and the second part includes the construction of pressure-boosting platforms. According to the reviews, this is a service and buy-back agreement. In this type of agreement, the motivation of investors is increased and the interests of both parties are aligned as a result of making changes compared to conventional buy-back agreements.

Also Mousavi Neqabi and Anoushe (24) showed that the development of SMEs is one of the key requirements of balanced regional development, considering their role in job creation and fair distribution of wealth. Cluster development is one of the strategies used in regional economic development programs of many countries for the development of SMEs. Although SMEs have many advantages over large enterprises, there are limitations on the development of SMEs in the areas of finance, marketing, production, R & D, etc., due to limited resources.

Yan et al. (25) found SMEs make an important contribution to economic-social development through job creation, the use of people's small capital in productive work, fair distribution of wealth, balanced development of regions, exploitation of regional advantages, etc. However, SMEs face financial, marketing, knowledge, information, and human resource limitations due to their small size. Policymakers have put various strategies on their agenda, such as supporting financing, developing domestic and export markets, promoting technology, empowering human resources,

and networking for the development of SMEs according to their limitations. In addition, Yan et al. (25) investigated on the strategies of entering foreign markets, foreign investment performance, and the effect of international trade and product compatibility on these factors were investigated in pharmaceutical companies affiliated to the syndicate using the principles of marketing management and international trade during six-months from March 2017 to September 2018. According to the results of the descriptive statistics of the variables, the strategies of entering foreign markets have the highest average, product compatibility has the smallest average and range of changes, and foreign investment performance has the lowest standard deviation, indicating the unanimity of opinion among the respondents. The results also indicated that there is a significant positive relationship between the strategies of entering foreign markets and foreign investment performance, between the strategies of entering foreign markets and international trade, and between the strategies of entering foreign markets and product compatibility. It was also found that international trade plays a mediating role in the relationship between the strategies of entering foreign markets and foreign investment performance.

Mousavi Niarki et al. (26), conducted a study on identification and prioritization reflecting on export of products of knowledge-based companies. The data needed to answer the questions were collected using questionnaires with open questions that were given to experts, namely 5 university professors and 5 marketing managers of knowledge-based companies located in Tehran Science and Technology Park. The experts were interviewed to identify the factors affecting the commercialization of products. In this way, 17 indicators in 4 sub-indicators were identified. Finally, 5 sub-indicators were removed from the main sub-indicators due to their lack of importance, and 12 final indicators remained. All the data were analyzed using the ANP method. According to the results, the rate of job creation by technology in the case of commercialization and mass production with a weight of 0.99, the presence of detailed regulations, instructions, and rules with a weight of 0.97, and the technical complexities of technology with a weight of 0.96 had the highest weights among the sub-indicators and were ranked first to third. The other weights obtained received other ranks. Ahmadi et al. (27) developed a comprehensive, quantitative hybrid design research method for knowledge-based companies.

Du et al. (28) summarized the main problems and obstacles in the protection of wildlife resources and their use in China. Based on Analytical Hierarchy Process (AHP), the main technological factors affecting on conservation and use of wildlife resources in China were identified. Finally, the direction of future wildlife conservation research development based on critical factors is discussed. It can provide guidance for the

development of wildlife resource conservation and utilization for a sustainable ecosystem in China.

Nozari et al. (29) argued by identifying and prioritizing these factors plays an important role in establishing a balance between the domestic economy and the global economy and determining the position of the food industry in the volume of trade and export products of the country. To this end, researcher-made questionnaires including 4 main measures and 15 main indicators were designed and distributed among 31 food exporting companies in Alborz Province. The validity of the questionnaires was confirmed through interviews with experts, and their reliability was confirmed using Cronbach's alpha. The weights of the measures were then determined using the AHP method, and the best indicators were prioritized to determine the factors determining export performance. The results suggested that the objective features of the company had the most weight and that the export commitment had the least weight among the measures. Besides, the first to third priorities respectively belonged to "export drivers", "direct export channel", and "duration of export experience".

Moosavi et al. (30), analyzed supply chain disruptions during the Coronavirus era and presented management strategies in order to create a stable and resilient supply chain.

Gillespie & Riddle (31) conducted a study on export promotion organization emergence and development. In most of the studies, the export promotion organization's performance is evaluated by evaluating the profits of the companies, awareness, understanding, and use of the organization's services. A small number of studies describe how the export promotion organization makes decisions about providing services. This study was conducted to provide a better understanding of how and why to determine export promotion organizations and provide and modify proposals for their services. It first reviewed the literature on the export promotion organization and suggested that both macro and micro approaches derived from the literature can be adapted to more effectively explore the effect of export promotion organization emergence and changes in its services.

Antoldi et al. (32) analyzed the performance of export consortia as cooperation, facilitating the achievement of innovation, and the export performance of Brazilian members through research consortia with different companies and economic sectors. According to the results of the survey analysis using descriptive statistics and Pearson's correlation, the enterprises sought to comply with international technological standards.

Bekerman et al. (33) suggested that the internationalization of SMEs is now a major challenge for the development of Argentina's economy due to the high concentration of Argentina's exports in the hands of some large companies. In this study, the experience of export consortia in Argentina was examined as a tool that

can play an essential role in the process of internationalization of SMEs.

Forte & Oliveira (4) argued that exporting is an essential strategy for the growth and survival of enterprises, especially SMEs. However, SMEs generally cannot overcome internal and external export barriers, and a consortium can help them to internationalize. Accordingly, this study was conducted mainly to identify the main effects of a consortium on the member enterprises as well as the gap in the literature. So, it reviewed the available empirical studies, which are mainly focused on case studies of export consortia. Export consortia can be effective in different areas of enterprises such as finance, market, and credit and help them overcome some export obstacles. Identifying the weaknesses of the literature suggests that future studies should focus on empirical studies with larger samples and consider other possible factors affecting exports to conduct a more detailed empirical analysis of the effects of consortia on enterprises' exports.

The reviewed domestic and foreign studies are listed in Table 1 based on their objectives and methodologies. Investigations indicate that the studies have neglected to propose a comprehensive approach that analyzes the state of the export consortium of SMEs and provides an accurate prediction of their future performance. So, the most important innovations and contributions of the study are as follows:

- Analyzing and recognizing the factors affecting the export of SMEs;
- Bringing the study closer to reality using a fuzzy inference engine;
- Developing an SD model to check and predict the model's future performance for making decisions about the export of SMEs.

3. METHODOLOGY

The methodology is referred to as a set of activities performed to determine where, how, and with what tools data is collected and analyzed to obtain the necessary results. This is a basic and applied-developmental study based on the objectives. Basic research is a theoretical one that is conducted solely to develop knowledge without a specific applied objective. Basic research is conducted for one of the objectives of science, which is a kind of science for science. This kind of research attracts a lot of attention because they pave the way for other studies (34). Applied research is used to meet human needs and improve and optimize tools, objects, and patterns in the development of welfare and comfort and the improvement of human living standards by using the cognitive background and information provided through basic research (34). Developmental studies are conducted to develop a more comprehensive indicator of the results of studies and expand their scope. Thus, developmental studies are those that are carried out as a continuation of previous studies to investigate additional dimensions and issues. These studies have a mixed approach. The mixed method uses a combination of quantitative and qualitative methods simultaneously. As a combination of quantitative and qualitative methods, mixed methods provide an opportunity for researchers to bring together the methodologies, designs, models, and methods used in a single study (34).

This is a basic and applied-developmental study based on objectives or results. It is also a descriptive exploratory study in terms of method and a survey one in terms of the data collection method. The data is collected and analyzed using a mixed (qualitative-quantitative)

TABLE 1. Literature review

Study type	Author	Year	Objective	Methodology	Other factors
Domestic	Ortega	(16)	Examining technical, economic, and legal dimensions of the agreement	Two-part analysis	-
	Mousavi Neghabi & Anousheh	(23)	Examining the role of the cluster in the development of sustainability of small enterprises	Cluster analysis	Sustainability
	Mousavi Neghabi	(24)	Examining solutions for sustainable development of small enterprises	Networking	Sustainability
	Yan et al.	(25)	Examining the effect of international trade on entering foreign markets	Statistical analysis	-
	Mousavi Niarki et al.	(26)	Prioritizing factors affecting the export of knowledge-based companies	ANP	-
	Ahmadi	(27)	Identifying and prioritizing factors affecting food exports	AHP	-
Foreign	Antoldi et al.	(32)	Examining export performance to obtain innovation	Statistical analysis	-
	Bekerman et al.	(33)	Examining the export productivity of small industries	Statistical analysis	-
	Forte & Oliveira	(4)	Reviewing small business exports	Statistical analysis	-

method. Moreover, the study is part of the critical paradigm. The statistical population is selected because the Iran Trade Promotion Organization is responsible for strengthening and developing export capacities at the country level, including export consortia. The sample consisting of 11 experts and consultants of export consortia in the Iran Trade Promotion Organization willing to cooperate in the study is selected by purposive expert sampling. The experts have at least 10 years of experience and study about export and export consortia. The demographic features of the population are given in Table 2.

In this study, the data are collected using library research and field methods. The data on the subject description and literature review are collected using the library research method and referring to domestic and foreign books, papers, and theses (via the Internet). The data required by the field section are collected in qualitative and quantitative phases.

In the qualitative phase, in-depth, exploratory, and collaborative interviews began by asking the following questions: "What are the main success factors of private sector export consortiums in Iranian industries?" and "What steps should be taken to establish private sector export consortia in Iranian industries?". Eleven focus group meetings are then organized with the experts. In each session, an expert is selected as a facilitator to explain the main success factors of export consortia and the steps to establish an export consortium based on his/her knowledge and experience. In the quantitative phase, FCM technique questionnaires are designed according to the main factors identified in the previous phase and are distributed among the experts. The experts' opinions concerning the questionnaires are then obtained through frequent contact with them. A brainstorming session is held with the experts to determine if-then rules for the FIS.

TABLE 2. Demographic features of the population

Expert	Gender	Age (year)	Education level	Work experience (year)
1	Male	50	Master's degree	20
2	Male	52	PhD.	26
3	Male	44	PhD.	11
4	Male	48	Master's degree	19
5	Male	45	Master's degree	18
6	Male	43	Master's degree	14
7	Male	49	PhD.	19
8	Female	40	Master's degree	16
9	Male	56	PhD.	28
10	Female	45	PhD.	18
11	Male	43	Master's degree	15

The questionnaires concern the FCM technique. They are the matrixes of the power of communication between the main and secondary factors, where the experts express their opinions about the relationship and the degree of effects between the two factors.

The reliability of the measurement tool should be ensured before performing any data analysis and statistical inference (28). Reliability indicates whether the same results will be obtained if the study is repeated with the same test. In this study, different data collection tools are used because it consists of qualitative and quantitative phases.

The FCM questionnaires are used in the quantitative phase. Content validity is evaluated to ensure the validity of the questionnaires. In the FCM technique, the matrix for determining relationships and the degree of effects of the factors are the same questionnaires whose validity is confirmed by the experts. The techniques used to analyze the data and the reasons for their use can be seen in Table 3.

Various problem-solving techniques have been proposed, each of which has specific features and conditions of application. In this study, decisions are made using FCM, SD, and FIS, which are briefly described below.

TABLE 3. The techniques used to analyze the data and the reasons for their use

Objectives	Techniques	Reasons
The first objective: identifying the main success factors of export consortia in Iranian industries	Grounded theory based on qualitative coding	The qualitative coding method is used to analyze the qualitative data in the grounded theory approach (35)
The second objective: determining the steps to establish export consortia in Iranian industries		
The third objective: designing a fuzzy roadmap of the main success factors of export consortia in Iranian industries	FCM	The FCM technique is used to analyze causal maps and improve the ability of decision-makers to understand the dynamic behavior of causal cognitive maps. The technique represents the causal relationship between concepts and analyzes inference patterns.
The fourth objective: Fuzzy dynamic modeling for private sector export consortia in Iranian industries	Fuzzy logic-based SDM	SDM does not exclusively deal with future quantitative prediction but seeks to gain extensive knowledge about the mutual dynamic relationship between the system components [39]

3. 1. Fuzzy Cognitive Mapping (FCM) Technique

Kosko (36) defines an FCM as a directed graph to show the causal relationships between factors, which is specified by a number in the range [-1,1]. In this model, the SD is shown by simulating the interactions between the concepts and factors of the system. The FCM technique is used to show both quantitative and qualitative data. The inputs obtained from the experience and knowledge of experts in the subject matter are required for the development of an FCM. So, in this technique, the accumulated experiences of people are integrated with the existing knowledge in the field for which the model is drawn, and the causal relationships between the factors constituting the system are formed accordingly (37). An FCM model is made up of a set of nodes or concepts represented with abbreviations in Equation 1,

$$C_i, \quad i = 1, 2, \dots, n \tag{1}$$

In this model, each node represents a key factor in the modeled system and is valued by Equation 2.

$$i = 1, 2, \dots, n, \quad A_i \in [0,1] \tag{2}$$

The model nodes are internally connected by weighted arcs. If a fuzzy cognitive map with n nodes C_j is given, the value of each node in each iteration can be calculated as Equation 3.

$$A_i^{t+1} = f(A_i^t + \sum_{j=1}^n A_j^t W_{ij}) \tag{3}$$

where A_i^{t+1} is the quantity of concept C_j at the time of $t + 1$, and A_j^t is the quantity of concept C_j at the time t . In addition, W_{ij} corresponds to the fuzzy weight between two groups, and f is a threshold function that converts the result of the coefficient into a number in the interval [0,1]. The logistic function, which is given in Equation 3, is the most prevalent kind of function f .

$$f(x) = \frac{1}{1+e^{-\lambda x}} \tag{4}$$

where $\lambda > 0$ determines the slope of the continuous function f .

$W_{ij} > 0$: indicates a positive cause and effect relationship

$W_{ij} < 0$: indicates a negative cause and effect relationship

$W_{ij} = 0$: no relationship

The methodology used in this research contains 6 phases as follows (38, 39,37, 40):

Step 1: Forming the relationship matrix or determining the relationships (effectiveness and impressionability) between the factors: Research professionals identify the relationships between the required parameters in this stage. A factor's relationships with its linked sub-factors are also established. Based on professional consensus, the determined relationships have been established.

Step 2: Determining the intensity of relationships based on linguistic variables: The linguistic variables are

initially defined in this stage. Then, depending on linguistic factors, each research specialist evaluates the strength of the associations under consideration (very little, little, medium, much, very much).

The number of linguistic variables multiplied by the relevant fuzzy number yields a result that represents the intensity of relationships as a fuzzy number. At this point, each relationship's fuzzy number has to be determined using the following formula. We therefore have:

J : set of linguistic variables ($j \in J$)

j : Linguistic variables ($j = VL, L, M, H, VH$)

K : The set of triangular fuzzy number values (upper bound, most probable, lower bound) ($k \in K$)

k : Triangular fuzzy number values (upper bound, most probable, lower bound) ($k = l, m, u$)

x_{jk} : The value of linguistic variable j with triangular fuzzy number bound k

I : The set of intensity of the factors' effectiveness on each other($i \in I$)

i :The intensity of the factors' effectiveness on each other($I=VL,L,M,H,VH$)

$f_i^{(A \rightarrow B)}$: The frequency of experts' opinions regarding the effectiveness of factor A on factor B with the effectiveness intensity i

$\alpha_k^{(A \rightarrow B)}$:The value of the triangular fuzzy number (upper limit, most probable, lower limit): k regarding the effectiveness of factor A on factor B . It is shown in the equation below.

$$\alpha_k^{(A \rightarrow B)} = \sum_{i=VL}^I \sum_{j=VL}^J f_i^{(A \rightarrow B)} x_{jk} \quad \forall k \in K \tag{5}$$

Step 3: De-fuzzification: The fuzzy mean method in Equations 3-6 were used for de-fuzzification.

$$Score_{(Crisp)}^{(A \rightarrow B)} = \frac{\alpha_l^{(A \rightarrow B)} + 4\alpha_m^{(A \rightarrow B)} + \alpha_u^{(A \rightarrow B)}}{6} \tag{6}$$

Step 4: effectiveness power and impressionability capacity: In this stage, we try to figure out how much effectiveness each component has and how much impressionability it can have. In this sense, each factor's effectiveness and impressionability are combined concerning other factors. That is, the total effectiveness power and the total impressionability capacity of each factor are obtained separately.

Step 5: Determining the axial index: The total of each factor's effectiveness and impressionability is added together to determine the axial index in this stage. The axial index demonstrates the significance of the variables. Indeed, a factor is referred to the superior factor or axial the more entire effectiveness power and impressionability ability it possesses.

Step 6: Designing fuzzy cognitive map: The identified relationships, effectiveness power, impressionability capacity, and axial index of each element are used to create the fuzzy cognitive map in this stage.

3. 2. System Dynamics The model given in this research continues its path based on the following steps, which are five in the system dynamics method (38).

Step 1: problem articulation: Modelers should clearly define their modeling objectives for themselves and completely comprehend the true problem, not just its side effects. A model, then, is a condensed portrayal of the real world.

Step 2: Formulating dynamic hypotheses: The boundary of the model is established at this point by identifying the major variables influencing the problem under investigation. Using the information already known about the problem, causal circles are then formed and gradually filled in to produce a final simplified representation of the problem's actual occurrence.

Step 3: Formulating a simulation model: The model is put into action when dynamic hypotheses and system boundaries are generated. The dynamic interactions between the elements in this mental model are frequently so complex that they are difficult to comprehend. This necessitates simulating and applying the model in a virtual environment. There are other programs available for this purpose, however, Vensim software is employed in this study.

Step 4: Model testing and its validation: Comparison of the model's simulated behavior with real-world behavior is essential. This comparison demonstrates that the model is ready to be utilized as a foundation for policy.

Step 5: Policy-making: It entails resolving tactics, frameworks, and new decision-making policies to enhance the system. To attain a more successful outcome in this direction, many integrated policies and scenarios are typically employed.

3. 3. Fuzzy Interference System There are two significant types of fuzzy inference system: Mamdani and Sugeno. The most popular inference system in

literature, Mamdani's fuzzy inference system, was one of the first fuzzy control systems. Fuzzy sets must be used as both input and output for Mamdani. After the aggregation operation, the output becomes defuzzified. In a Mamdani fuzzy inference system, the input variables are combined in such a way that if x_1, x_2 and x_n are equal to $A_1, A_2,$ and A_n , then y is equal to B , where x_i ($i = 1, 2, 3, \dots, m$) are input variables and y is an output variable. A_1, A_2, \dots, A_n and B represent the linguistic terms. The inputs in Mamdani fuzzy inference system may be clear or unclear (39).

4. RESULTS ANALYSIS

The grounded theory technique was used in this study to identify the key success criteria of private-sector export consortia in the Iranian industry. Following the interviews, the information was categorized using three steps of qualitative coding (open, axial and selective) (40).

"Open coding" is the initial stage of data analysis. In the current study, the interviews were initially open-coded into ideas pertinent to the research issue. Axial coding is connected to the second phase. Open codes are now grouped into many categories (subcategories) according to their shared content. In other words, depending on shared axes, the concepts were compared to one another and classified into subclasses. "Selective coding" is the name of the third level. Generally, 26 axial codes, 7 selective codes, and 142 open codes were found. The results of qualitative coding using the grounded theory approach are displayed in Table 4 (41).

The fuzzy cognitive map approach (FCM) was employed to create a fuzzy cognitive strategic map for the primary success determinants of private sector export consortia in Iranian industries. Six phases were used in the fuzzy cognitive mapping approach. Accordingly, the matrix of the strength of communication and the power

TABLE 4. Qualitative coding to identify the main success factors of private sector export consortia in Iranian industries

Main categories code	Main categories (selective codes)	Sub-categories code	Sub-categories (axial codes)
FCMs	Features of consortium members	FCMs1	Consortium members' experiences
		FCMs2	The export share of consortium members
		FCMs3	Power of production/supply of consortium members
EOP	Export assistance programs	EOP1	Identifying potential export markets and creating a database related to their features
		EOP2	Export consulting inside and outside the country
		EOP3	Participating in foreign international exhibitions
		EOP4	Establishing branches in target markets
		EOP5	Online advertising of export goods (online exports)
		EOP6	Determining the export strategy (direct, indirect and cooperative)

CB	Consortium strengthening factor	CB1	Establishing a teamwork culture instead of a solitary entrepreneurial spirit
		CB2	Co-branding and its strengthening
		CB3	Quality assimilation and its improvement
		CB4	Value-creation through after-sale services
		CB5	Competitive pricing
ES	Recognizing export support	ES1	Familiarity and using the support of government organizations (such as receiving subsidies, services provided by the trade development organization, facilitating export procedures, etc.)
		ES2	Awareness and use of support by incubation centers and export accelerators
		ES3	Knowledge and use of private and foreign government networks
TFs	Transnational factors	TFs1	Using foreign investments
		TFs2	Certificate of international standards
		TFs3	Forming the global trade promotion team
		TFs4	Employing foreign expert human resources
GFs	Governmental factors	GFs1	Efforts to establish cooperation between business institutions of trading countries
		GFs2	Financing through export-oriented securities
PFs	Product features	PFs1	Improving the quality of existing products
		PFs2	Developing new products
		PFs3	Attractive packaging

of effectiveness, impressionability and axial index about the primary success factors and associated sub-factors was obtained. In this regard, expert opinions were gathered to determine the matrix of communication between the factors and the strength of the relationship between them. The visual representation of the fuzzy

cognitive map is created following Figure 1 based on the data acquired in the earlier phases, the creation of the relationship matrix between the secondary variables, and the relationship matrix of the final relationship between the key elements.

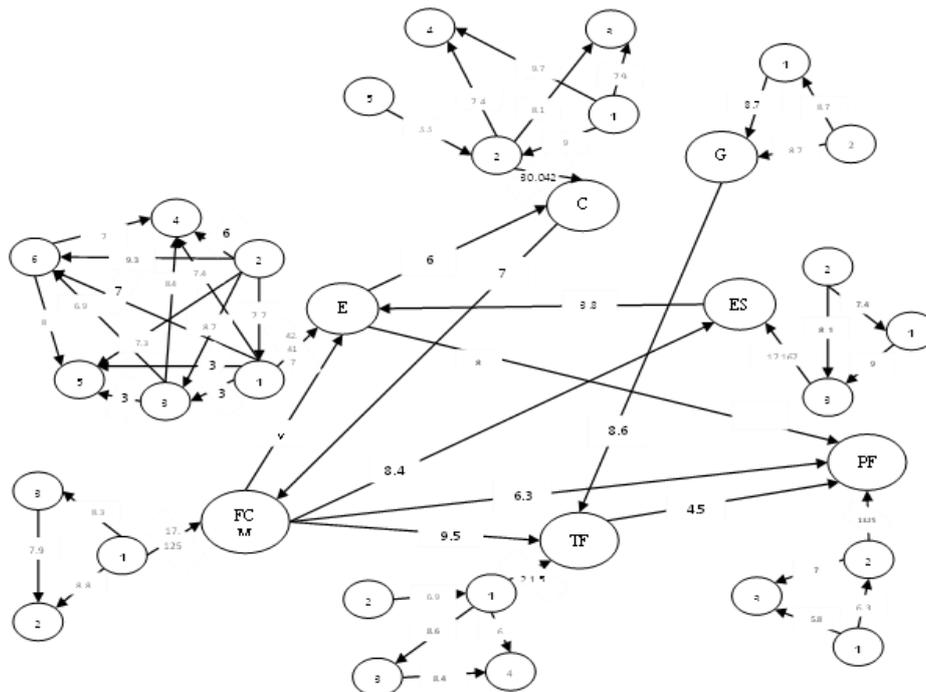


Figure 1. Fuzzy cognitive map for the main success factors of private sector export consortia in industries

consortia. The initial value of "performance of export consortia" is taken in this study to be 0.2. The integral of the sum of the flow or rate variables, which are the primary success determinants for export consortia from 2011 to 2031, is used to determine the performance of export consortia (stock variable), which has been shown in Equations 7 and 8.

$$\int (\text{consortium member features} + \text{operational program exports} + \text{consortium strengthening agent} + \text{export recognition support} + \text{transnational agents} + \text{governmental agents} + \text{product features}) \quad (7)$$

$$\int (\text{FCM} + \text{EOP} + \text{CB} + \text{ES} + \text{TF} + \text{GF} + \text{PF}) \quad (8)$$

4. 2. Flow Variables (Rate Variable) The primary export consortium success factors in this study's flow variables are listed below. Table 5 generally provides instructions on how to determine the variables of the study model.

Features of consortium members: The features of consortium members are determined by a variety of factors, including "consortium members' experiences," "consortium members' export share," and "consortium members' production/supply power". When analyzing the connections between these variables, DELAYII and RAMP functions were utilized to take into account their significance and effectiveness.

Export assistance program: This variable includes a series of different activities such as "familiarity with potential export markets and creating a database related to their features", "national and international export consultation", "participation in international exhibitions abroad", "creating branches in the target markets", "online advertising of export goods" and "determining the export strategy", all of which and the relationships between them constitute the export assistance program. The SMOOTH, STEP, DELAYII, and RAMP functions have been employed in the interaction between these elements following their meaning and effectiveness on one another. The variables "features of consortium members" and "recognition of export support" is also incorporated in the export assistance program based on the cause and effect relationships of the key success criteria.

Consortium strengthening factor: This variable consists of a series of different activities such as "creating a teamwork culture instead of the individual spirit of entrepreneurship", "co-branding and its strengthening", "standardization of quality and its improvement", "value creation through after-sales services" and "competitive pricing" which together with the relationships between them constitute the strengthening factor of the

consortium. In the relationships between these components, under their significance and mutual effect, DELAYII and RAMP functions have been applied. Additionally, the variable "export assistance program" is added to the consortium strengthening factor based on the cause-and-effect correlations of the primary success variables.

Identifying export supports: This variable includes a variety of activities, including "awareness and use of support from government organizations," "knowledge and understanding and use of support from incubation centers and export accelerators," and "recognition and use of private and government networks abroad." These activities individually and in combination together make up the recognition of export support. The RAMP function has been applied to the relationship between these variables while taking into account their significance and effectiveness. Additionally, the variable "features of the consortium members" is included in the recognition of export support based on the causal links of the key success criteria.

Transnational factors: The activities included in this variable, along with their relationships, constitute transnational factors. These activities include "Using Foreign Investments," "Certificate of International Standards," "Establishing a Global Trade Promotion Team," and "Employing Foreign Expert Human Resources." The SMOOTH, DELAYII, STEP, and RAMP functions have been employed in the relationships between these components based on their concepts and effectiveness on one another. Furthermore, the variables "features of consortium members" and "government factors" are added to the transnational factors based on the cause-and-effect correlations of the primary success factors.

Governmental factors: A variety of activities, including "efforts to establish cooperation between commercial entities of trading countries" and "financing through export-oriented securities," are included in this variable. These activities, along with the interconnection between them, are what are known as government agents. STEP and RAMP functions have been employed in the interaction between these components based on their concept and effectiveness.

Product features: The actions included in this variable, which together with their relationships create the qualities of the product, are "increasing the quality of existing products," "creating new products," and "attractive packaging." The relationship between these parameters has been examined using the DELAYII and RAMP functions, taking into account their significance and effectiveness. Additionally, the variable "features of the consortium members" is added to the product features based on the causal relationships between the primary success drivers.

TABLE 5. Method of calculating (formula) variables of the research model

Research model variables	Variables' code	Calculation method (formula)
export consortia performance	ECP	$0.2 + \int_{1390}^{1410} (FCMs + EOP + CB + ES + TFs + GFs + PFs)$
Features of consortia members	FCMs	FIS * DELAY1 ((RAMP (((FCMs1 + FCMs2 + FCMs3)/3) + (ECP)/2), 1395, 1405)), 1395)
The export share of consortium members	FCMs2	DELAY1I (((FCMs1 + FCMs3)/2), 1395, 0)
Power of production/supply of consortium members	FCMs3	RAMP (((FCMs1 + FCMs2)/2), 1395, 1405)
Export assistance programs	EOP	FIS * ((RAMP (((EOP1 + EOP2 + EOP3 + EOP4 + EOP5 + EOP6)/6), 1395, 1405) + ((ECP + ES + FCMs)/3)/2)
Identifying potential export markets and creating a database related to their features	EOP1	SMOOTH (((EOP2 + EOP3 + EOP6)/3), 1395)
Participation in international exhibitions	EOP3	DELAY1I (((EOP2 + EOP4 + EOP5)/3), 1395, 0)
Establishing branches in target markets	EOP4	STEP (((EOP1 + EOP2 + EOP6)/3), 1395)
Online advertising of export goods (online exports)	EOP5	RAMP (((EOP1 + EOP2 + EOP3 + EOP6)/4), 1395, 1405)
Determining the export strategy (direct, indirect and cooperative)	EOP6	STEP (((EOP2 + EOP3)/2), 1395)
Consortium strengthening factor	CB	FIS * ((RAMP (((CB1 + CB2 + CB3 + CB4 + CB5)/5), 1395, 1405) + ((EOP + FCMs)/2)/2)
Co-branding and its strengthening	CB2	DELAY1I ((RAMP (((CB1 + CB5)/2), 1395, 1405)), 1395, 0)
Quality standardization and its enhancement	CB3	RAMP (((CB1 + CB2)/2), 1395, 1405)
Value creation through after-sales service	CB4	RAMP (((CB1 + CB2)/2), 1395, 1405)
Identifying export support	ES	FIS * ((RAMP (((ES1 + ES2 + ES3)/3), 1395, 1405) + (FCMs)/2)
Familiarity with and using the support of government organizations (such as receiving subsidies, services supplied by the trade development organization, facilitating export procedures, etc.)	ES1	RAMP (ES2, 1395, 1405)
Identifying and utilizing private and international government networks	ES3	RAMP (((ES1 + ES2)/2), 1395, 1405)
Transnational factors	TFs	FIS * (SMOOTH (((TFs1 + TFs2 + TFs3 + TFs4)/4), 1395) + ((FCMs + GFs)/2))
Using foreign investments	TFs1	DELAY1I (((TFs2 + TFs2)/2), 1395, 0)
Forming the global trade promotion team	TFs3	RAMP (TFs4, 1395, 1405)
Employing international expert human resources	TFs4	DELAY1I (STEP ((TFs1, 1397)), 1395, 0)
Governmental agents	GFs	FIS * ((RAMP (((GFs1 + GFs2)/2), 1395, 1405) + (ECP)/2)
Financing through export-oriented securities	GFs2	DELAY1I (GFs1, 1395, 1405)), 1395, 0)
Product features	PFs	FIS * ((RAMP (((PFs1 + PFs2 + PFs3)/3), 1395, 1405) + ((ECP + FCMs)/2)/2)
Developing new products	PFs2	DELAY1I (PFs1, 1395, 1405)), 1395, 0)
Attractive packaging	PFs3	RAMP (((PFs1 + PFs2)/2), 1395, 1405)

4. 3. Simulation and Strategy Selection After the model and formulation were developed, a simulation was carried out with Vensim software. In this part, several policy and performance scenarios for export consortiums are explored utilizing the feedback mechanisms developed for the model. After the model and formulation were developed, a simulation was carried out with Vensim software. In this part, several

policy and performance scenarios for export consortiums are explored utilizing the feedback mechanisms developed for the model. The simulation time unit "year" has been chosen, and the simulation's duration is 21 years, taking into account that reports are prepared and results are reviewed yearly. The findings indicate that the performance of the export consortium has improved over time as a consequence of improvements made to the key

success determinants that have been identified. Finally, ideas were put out and several scenarios were examined based on the key success elements that had been found.

The model's implementation findings demonstrate that export consortiums may provide outstanding outcomes in the upcoming years if they appropriately apply the key success elements that have been discovered. Overall, the findings show that the successful execution of the export consortium's key success determinants raises the performance level of export consortia (Table 6).

Hypothesis/Scenario 1: Considering the consortium members' features improves the performance of export consortiums.

The number of the consortium members' features is effectively determined in this part by two separate currents, which are as follows:

1. Current one: To generate the minimal level of performance that is acceptable for the features of the consortium members, the consortium attempts to modify all the sub-factors that are connected to those qualities (Figure 3). Using a fuzzy inference system that establishes the weight of the variables in the relationships of the system dynamics approach, the choice of a combination of these sub-factors has been made. Through if-then rules, this decision support system also serves as an audit for the upper and lower bounds of the attributes of consortium members.
2. Current two: The consortium is attempting to quantify the impact of the key success variables on the features of the consortium members in this current.

According to the findings (Figure 3), the "features of consortium members" factor is influenced by the pertinent sub-factors and "export consortium performance."

TABLE 6. The performance results of consortia during 2011-2031

Year	Export consortia performance	Year	Export consortia performance
2011	0.2	2022	25.2407
2012	0.336312	2023	42.2763
2013	0.512666	2024	68.596
2014	0.740848	2025	108.04
2015	1.03612	2026	165.707
2016	1.41822	2027	248.327
2017	1.91272	2028	355.282
2018	2.86319	2029	493.735
2019	4.74533	2030	672.959
2020	8.30188	2031	904.956
2021	14.6215	---	---

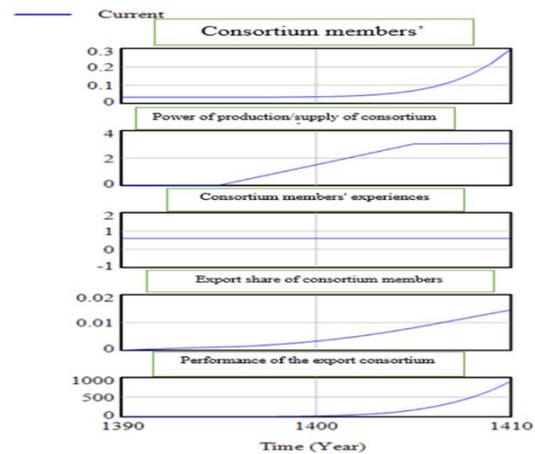


Figure 3. The results of the features of the consortium members and their influencing factors

Hypothesis/Scenario 2: Considering the export operational plan improves the performance of export consortiums.

The quantity of the export operational plan in this section is determined by two separate currents, which are as follows:

1. Current one: The consortium strives to achieve a minimal level of acceptable performance for the export operational plan during this step by adjusting all the connected sub-factors that have an impact on it (Figure 4). Using a fuzzy inference system that establishes the weight of the variables in the relationships of the system dynamics approach, the choice of a combination of these sub-factors has been made. Through if-then logic, this decision support system also serves as an audit for the export operating plan's upper and lower limitations.
2. Current two: The consortium is now attempting to quantify the influence of the key success elements on the export operating plan (Figure 4).

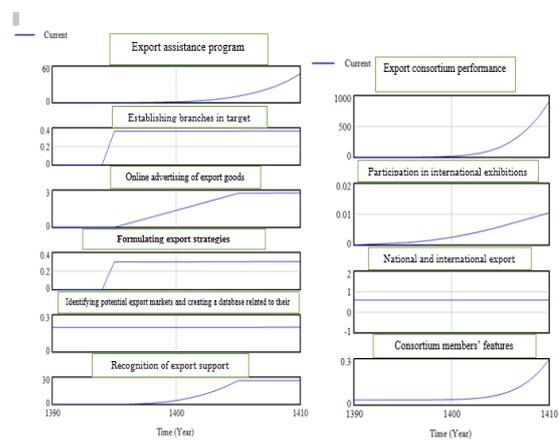


Figure 4. The results of the export assistance program and its affecting factors

According to the findings (Figure 4), the export assistance program is influenced by the pertinent sub-variables as well as the factors "features of consortium members," "awareness of export support," and "performance of export consortia."

Hypothesis/Scenario 3: Considering the consortia strengthening factor improves the performance of export consortia.

The quantity of the consortium strengthening factor in this section may be calculated using two distinct currents, which are as follows:

1. Current one: The consortium works to generate the lowest acceptable performance level for the consortium amplifier throughout this current by adjusting all the sub-factors that are connected to and have an impact on the consortia amplifier (Figure 5).

Using a fuzzy inference system that establishes the weight of the variables in the relationships of the system dynamics approach, the choice of a combination of these sub-factors has been made. Through if-then rules, this decision support system also serves as an audit for the maximum and lower bounds of the consortium enhancing factor.

2. Current two: The consortium attempts to evaluate the impact of the key success elements that are applied to the factor that strengthens the collaboration in this current (Figure 5).

According to the findings (Figure 5), the consortia strengthening factor is influenced by the pertinent sub-factors as well as the factors "features of consortium members" and "export assistance programs."

Hypothesis/Scenario 4: Considering the recognition of export support improves the performance of export consortia.

The amount of recognition of export support is determined in this section by two separate currents, which are as follows:

1. Current one: The consortium attempts to alter each sub-factor involved in the recognition of export support

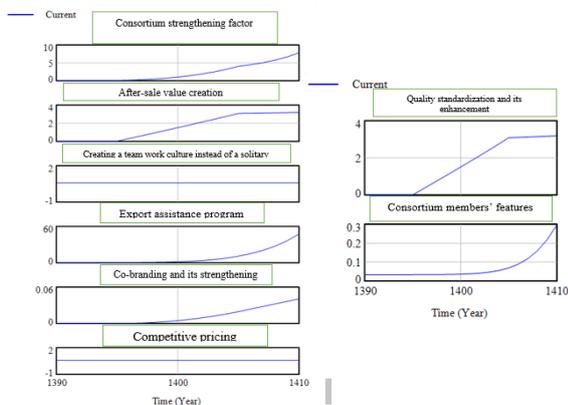


Figure 5. The results of the consortium strengthening factor and its affecting factors

during this process to define the minimum level of acceptable performance for its recognition (Figure 6). Using a fuzzy inference system that establishes the weight of the variables in the relationships of the system dynamics approach, the choice of a combination of these sub-factors has been made. Through if-then rules, this decision support system also serves as an audit for the top and lower limits of the recognition of export assistance.

2. Current two: The collaboration is now attempting to quantify the impact of the key success features on the acceptance of export assistance (Figure 6).

According to the findings (Figure 6), the factor "features of consortium members" and its related sub-factors have an impact on identifying export support.

Hypothesis/Scenario 5: Considering the transnational factors improves the performance of export consortia.

The level of focus on transnational elements in this part is effectively influenced by two separate currents, which are as follows:

1. Current one: The consortium attempts to modify each sub-factor connected to the transnational factors in this current such that the transnational factors perform at a minimal level that is deemed acceptable (Figure 7).

Using a fuzzy inference system that establishes the weight of the variables in the relationships of the system dynamics approach, the choice of a combination of these sub-factors has been made. Through if-then rules, this decision support system also serves as an audit for the top and lower bounds of multinational factors.

2. Current two: The collaboration is attempting to quantify the influence of the key success variables imposed on international elements in this current (Figure 7).

According to the findings (Figure 7), the transnational variables are influenced by the pertinent sub-factors as well as the "features of consortium members" and "government factors."

Hypothesis/Scenario 6: considering governmental factors improves the performance of export consortia.

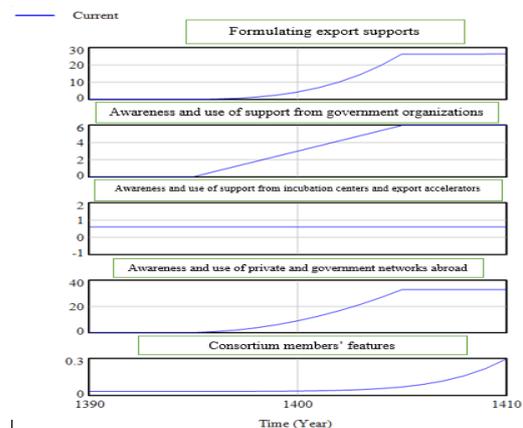


Figure 6. The results of identifying export support and its affecting factors

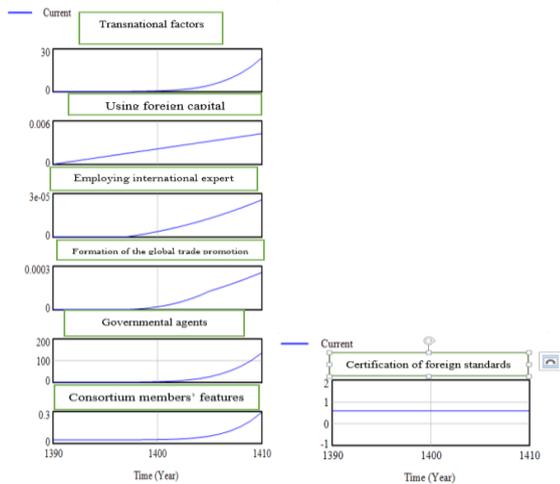


Figure 7. The results of transnational factors and their affecting factors

The degree of consideration given to governmental factors in this part may be effectively determined by two separate currents, which are as follows:

1. Current one: The consortium works to develop the minimal level of acceptable performance for the governmental factors by adjusting all the sub-factors that are relevant to them during this current (Figure 8).

Using a fuzzy inference system that establishes the weight of the variables in the relationships of the system dynamics approach, the choice of a combination of these sub-factors has been made. Through if-then rules, this decision support system also serves as an audit for the upper and lower bounds of government agents.

2. Current two: The consortium attempts to gauge how the key success variables affect government employees during this current (Figure 8).

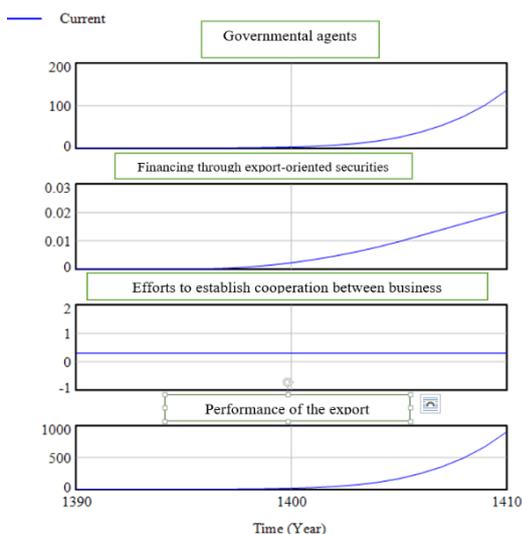


Figure 8. The results of governmental agents and their affecting factors

The findings are depicted in Figure 8, which demonstrates how the relevant sub-factors and "export consortium performance" affect the governmental agents.

Hypothesis/Scenario 7: Considering product features improves the performance of export consortiums.

The following two currents are useful in determining how much attention is paid to product features in this section:

1. Current one: To determine the minimum acceptable performance level for the product features, the consortium adjusts all the associated sub-factors to the product features during this procedure (Figure 9).

Using a fuzzy inference system that establishes the weight of the variables in the relationships of the system dynamics approach, the choice of a combination of these sub-factors has been made. Through if-then rules, this decision support system also serves as an audit for the top and lower bounds of product features.

2. Current two: The consortium is now attempting to quantify the impact of the key success factors on the product features (Figure 9).

4. 4. Model Validation

A developed model must be valid to the individuals engaged in the pertinent events for it to be effective in policy analysis. After passing the validity tests, a model may be used for this purpose to examine how various policies affect the system. Various tests have been applied in this study to determine the model's validity as follows:

Boundary Adequacy test: This test is carried out to see if the model's structure is appropriate for its intended use. In other terms, the following two issues are brought up: 1. Are the primary influencing factors visible in the model as endogenous variables?

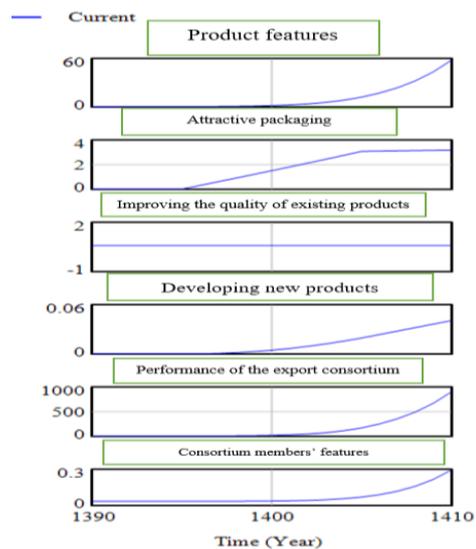


Figure 9. The results of the product features and its affecting factors

2. Is the period taken into proper consideration? All of the factors that affected the export consortium's performance in this study were regarded as endogenous factors. All of the variables included in the suggested model were, in other words, created as endogenous factors. It is important to note that research professionals confirmed the desired variables in the suggested model. It is required to run the model in short- and medium-term periods and check the outcome by expert review or comparison with historical data in the original model to confirm the proper time frame. This led to the implementation of the study model in two short-term (two years from the beginning of 2011 to the beginning of 2013) and medium-term (five years from the beginning of 2011 to the beginning of 2016) seasonal time frames with a time step of 0.25, the results of which are given in Table 7. As shown, it can be concluded that the suggested model may provide a decent approximation of the real behavior in light of the lack of change in the performance of export consortia throughout the short- and medium-term time frames compared to the considered time frame, as well as the validation by the experts.

Structure Assessment: The primary concern is whether the structure of the proposed model is consistent with the system's current rules and decision-making procedures. Accordingly, a test was carried out to see if the structure of the research model and the organization of export consortia in practice conflicted were compatible. The study professionals acknowledged that the behavior of the variables to improve the export consortium's performance and accomplish the targeted aims is entirely compatible with reality.

TABLE 7. The results of the research model in the short-term and medium-term periods to validate the adequacy of the boundaries test.

Time frame	Year	The performance level of export consortiums in different seasons			
		First three months	Second three months	Third three months	Fourth three months
Short term	2011	0.2	0.234078	0.270659	0.309926
	2012	0.352078	0.397327	0.4459	0.498043
	2013	0.554018	---	---	---
Medium-term	2011	0.2	0.234078	0.270659	0.309926
	2012	0.352078	0.397327	0.4459	0.498043
	2013	0.554018	0.614106	0.678612	0.74786
	2014	0.822198	0.902002	0.987674	1.07965
	2015	1.17838	1.28438	1.39817	1.52033
	2016	0.65147	---	---	---

Extreme Condition test: Here, it should be assessed whether the model exhibits proper behavior and reactivity in the event of unexpected changes in particular parameters. The question of whether a rapid change in the model is apparent over a long time may also be used to study this problem. Accordingly, the sensitivity of the model to these changes was tested by changing the parameter values to very high and very low states. The outcomes demonstrate the model's logical behavior in these circumstances. Furthermore, just the slope or range of the graphs varies with the change in the export consortium's performance, and the behavior of the model is not sensitive to this change. Besides, the research model was used for 50 years before any abrupt behavior or alterations were noticed. This demonstrates that, in terms of the end conditions, the research model is valid.

Partial tests of model components: After breaking the model down into smaller, simpler sub-models for this test, each of the sub-models underwent the boundary behavior test before being added to the main model.

Surprise Behavior: Is the study model capable of producing unusual, weird, and undetected behavior over the chosen period? The findings show that the study model exhibits no surprise behavior.

Parameter Assessment: Here, the topic of whether the model's parameters and their numerical values are consistent with the system's data is raised. Experts validated the conformity of the model's parameters with reality based on this test. The model parameters include "business of consortium members", "certificate of foreign standards", "efforts to establish cooperation between business institutions of trading countries", "awareness and use of support from incubation centers and export accelerators", "national and international export consultation", "competitive pricing", "creating a team work culture instead of the spirit of individual entrepreneurship" and "improving the quality of existing products". Moreover, these factors' values, which reflect the situation of export consortia at the moment, in reality, were established by specialists.

Parameter Sensitivity Analysis: If there are substantial changes in numerical values, behavior, and consequent policies with changes in parameters, boundaries, and periods, it should be identified in the sensitivity analysis. Thus, three scenarios are suggested to reinforce the research model's parameters and enhance export consortia' performance (Table 8). Because the study specialists think the reinforcements can be executed within 2 years and the outcomes may be employed, these scenarios attempt to strengthen 10% of all model parameters within that time. The first scenario is to demonstrate what happens when all model parameters are strengthened by 10% in 2013. The second scenario aims to demonstrate what happens when all model parameters are strengthened by 10% in 2023 and 2025. The final scenario aims to demonstrate what happens

when all model parameters are strengthened by 10% in 2023, 2025, and 2027. The outcomes of the model parameter augmentation policies based on the three suggested scenarios are displayed in Figure 10 and Table 9, respectively. According to the outcomes of the

scenarios (i.e., 10% strengthening of all model parameters throughout three time periods in 1402, 1404, and 1406), the performance of export consortiums in 2031 under scenarios 1 to 3 can improve by 9.5%, 20.6%, and 27.5%, respectively, over the baseline model.

TABLE 8. The way of applying 10% strengthening policy of all model parameters in 2023, 2025 and 2027

Research model parameters	Parameters 'code	Initial value	Scenario 1 (10% increase in 2023)	Scenario 2 (10% increase in 2025)	Scenario 3 (10% increase in 2027)
Consortium members' experiences	FCMs1	0.6	0.6 + RAMP (0.06, 1402, 1404)	0.6 + RAMP (0.06, 1402, 1404) + RAMP (0.12, 1404, 1406)	0.6 + RAMP (0.06, 1402, 1404) + RAMP (0.12, 1404, 1406) + RAMP (0.18, 1406, 1408)
National and international export consulting	EOP2	0.6	0.6 + RAMP (0.06, 1402, 1404)	0.6 + RAMP (0.06, 1402, 1404) + RAMP (0.12, 1404, 1406)	0.6 + RAMP (0.06, 1402, 1404) + RAMP (0.12, 1404, 1406) + RAMP (0.18, 1406, 1408)
fostering a culture of teamwork rather than a single entrepreneurial mentality	CB1	0.6	0.6 + RAMP (0.06, 1402, 1404)	0.6 + RAMP (0.06, 1402, 1404) + RAMP (0.12, 1404, 1406)	0.6 + RAMP (0.06, 1402, 1404) + RAMP (0.12, 1404, 1406) + RAMP (0.18, 1406, 1408)
Competitive pricing	CB5	0.6	0.6 + STEP (0.06, 1402)	0.6 + STEP (0.06, 1402) + STEP (0.12, 1404)	0.6 + STEP (0.06, 1402) + STEP (0.12, 1404) + STEP (0.18, 1406)
Understanding and using incubation centers' and export accelerators' help	ES2	0.6	0.6 + RAMP (0.06, 1402, 1404)	0.6 + RAMP (0.06, 1402, 1404) + RAMP (0.12, 1404, 1406)	0.6 + RAMP (0.06, 1402, 1404) + RAMP (0.12, 1404, 1406) + RAMP (0.18, 1406, 1408)
Certificate of international standards	TFs2	0.6	0.6 + STEP (0.06, 1402)	0.6 + STEP (0.06, 1402) + STEP (0.12, 1404)	0.6 + STEP (0.06, 1402) + STEP (0.12, 1404) + STEP (0.18, 1406)
Initiatives to foster collaboration between trade countries' business entities	GFs1	0.3	0.3 + RAMP (0.03, 1402, 1404)	0.3 + RAMP (0.03, 1402, 1404) + RAMP (0.06, 1404, 1406)	0.3 + RAMP (0.03, 1402, 1404) + RAMP (0.06, 1404, 1406) + RAMP (0.09, 1406, 1408)
Improving the quality of existing products	PFs1	0.6	0.6 + STEP (0.06, 1402)	0.6 + STEP (0.06, 1402) + STEP (0.12, 1404)	0.6 + STEP (0.06, 1402) + STEP (0.12, 1404) + STEP (0.18, 1406)

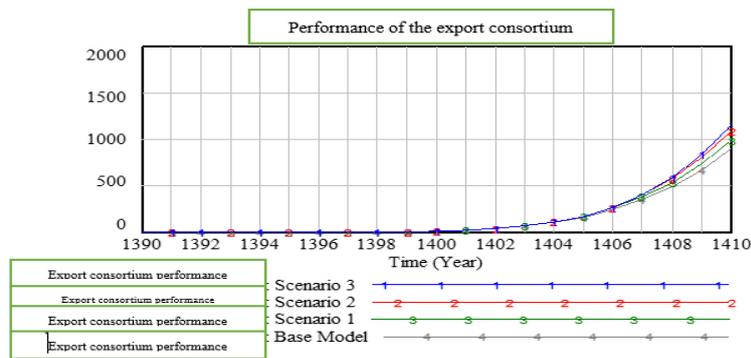


Figure 10. Performance results of export consortia based on three proposed scenarios

TABLE 9. Performance results of export consortia based on three proposed scenarios

Year	Main model	Scenario 1 (10% increase in 2023)	Scenario 2 (10% increase in 2025)	Scenario 3 (10% increase in 2025)
2011	0.2	0.2	0.2	0.2
2012	0.336312	0.336312	0.336312	0.336312
2013	0.512666	0.512666	0.512666	0.512666
2014	0.740848	0.740848	0.740848	0.740848
2015	1.03612	1.03612	1.03612	1.03612

2016	1.41822	1.41822	41822	41822
2017	1.91272	1.91272	1.91272	1.91272
2018	2.86319	86319	86319	86319
2019	4.74533	4.74533	4.74533	4.74533
2020	8.30188	8.30188	8.30188	8.30188
2021	14.6215	14.6215	14.6215	14.6215
2022	25.2407	25.2407	25.2407	25.2407
2023	42.2763	42.2763	42.2763	42.2763
2024	68.596	68.596	68.596	68.596
2025	108.04	110.0629	110.0629	110.0629
2026	165.707	173.2904	173.5811	173.5811
2027	248.327	264.6616	271.7304	271.7304
2028	355.282	382.945	405.131	405.653
2029	493.735	536.0629	577.8181	588.533
2030	672.959	734.2695	801.3565	834.8707
2031	904.956	990.8378	1090.715	1153.537

5. DISCUSSION

Among the various benefits of consortia, the literature review shows increased presence in foreign and international markets, reflecting increased exports, increased business contacts, increased bargaining power, and acquisition of knowledge about markets. A review of studies focusing on consortium case studies confirms the identified benefits. In particular, it can be seen that the internal competencies of the member companies related to internationalization, such as the number of customers abroad and their position in the international market, have improved significantly (43). Of course, it should be noted that many changes are always and the conditions including economic, political, cultural and environmental conditions are also influential in improving or not improving these conditions. For example, the Corona epidemic has been one of the changes and conditions that have had many effects on consortia and other business sectors (44). During the last few years, the effects of this big change have been the basis of many researches. Some researchers investigated the effects of various types of risks and trust in post-pandemic industries (45). Many researches have also been done on supply chains and the effects of intelligence on the supply chain of small and medium businesses after the corona virus pandemic. In this regard, some researchers proposed specific management strategies to manage supply chains after the pandemic (46). Therefore, it can be seen that the Corona pandemic has many effects on the export and import of products in small and medium-sized companies. For this reason, in this research, an effort has been made to investigate the role of consortia in the export of small and

medium companies by presenting a fuzzy dynamic model. Of course, many other approaches have been used in the literature. For example, some researchers have used meta-heuristic methods and evolutionary optimization methods such as particle swarm to solve models. These researchers have used these innovative and meta-heuristic methodologies for their allocation and placement models. But by reviewing the literature, we can see that one of the best methods is the dynamic modeling method, which is discussed in this research.

According to the findings, "features of consortium members," "export operational plan," "consortium strengthening factor," "recognition of export support," "transnational factors," "governmental factors," and "product features" are the seven main success factors of private sector export consortia in Iran's industries. For the establishment of private sector export consortiums in Iranian industries, the following eight procedures were established: 1. choosing an appropriate promoter, 2. Finding potential participants, 3. Carrying out interest research and contacting possible participants. 4. Selecting representatives. 5. Organizing meetings with prospective participants. 6. Carrying out a feasibility study and creating a business plan 7. Formally establishing a collaboration. 8. Monitoring consortium activities.

The results show that a hazy strategic map was created for the major elements that contributed to private sector export consortia's performance in Iranian industries, and "features of the consortium members" was identified as the most important and determining element. In the fuzzy cognitive map, the terms "export operational program," "transnational factors," and

"product features" each have the maximum axial indices. The findings indicated that the fuzzy dynamic model for private sector export consortia in Iranian industries comprises seven rate variables (current) that are crucial to export consortia success and one state variable (inventory) titled "performance of export consortia."

6. CONCLUSION

Researchers can draw inspiration from the subject of this study and other statistical and decision-making techniques and tools to conduct studies such as compiling guidelines for export consortia based on the key success factors identified in various industries, designing consortia maturity models, and offering an export excellence model to help export consortia perform better.

7. REFERENCES

- Ghayoomi M, Abooei M, Vahdatzad M, Ebrahimi A. Designing a model for creation of export consortia business cluster. *International Journal of Engineering, Transactions C: Aspects*, . 2020;33(3):459-67. <https://doi.org/10.5829/ije.2020.33.03c.10>
- Rahman R, Shahriar S, Kea S. Determinants of exports: A gravity model analysis of the Bangladeshi textile and clothing industries. *FIIB Business Review*. 2019;8(3):229-44. <https://doi.org/10.1177/2319714519872643>
- Sohrabi R, Pouri K, Sabk Ara M, Davoodi SM, Afzoon E, Pourghader Chobar A. Applying sustainable development to economic challenges of small and medium enterprises after implementation of targeted subsidies in Iran. *Mathematical Problems in Engineering*. 2021;2021:1-9. <https://doi.org/10.1155/2021/2270983>
- Forte R, Oliveira T. The role of export consortia in the internationalization of small and medium enterprises: a review. *International Entrepreneurship Review*. 2019;5(4):7. <https://doi.org/10.15678/IER.2019.0504.01>
- O'Reilly S, Haines M, Arfini F. Food SME networks: Process and governance-The case of Parma ham. *Journal on Chain and Network Science*. 2003;3(1):21-32. <https://doi.org/10.3920/JCNS2003.x027>
- Vaseei M, Agha MNJ, Abolghasemian M, Chobar AP. Investigating the Role of Transformative Technologies and Smart Processes on Sustainable Business. *Building Smart and Sustainable Businesses With Transformative Technologies: IGI Global*; 2024. p. 38-51.
- Emami A, Hazrati R, Delshad MM, Pouri K, Khasraghi AS, Chobar AP. A novel mathematical model for emergency transfer point and facility location. *Journal of Engineering Research*. 2023. <https://doi.org/10.1016/j.jer.2023.11.004>
- Pan H, Bayanati M, Vaseei M, Chobar AP. Empowering solar photovoltaic logistic operations through cloud-enabled blockchain technology: a sustainable approach. *Frontiers in Energy Research*. 2023;11:1293449. <https://doi.org/10.3389/fenrg.2023.1293449>
- Welch D, Welch L, Wilkinson IF, Young LC. Network analysis of a new export grouping scheme: The role of economic and non-economic relations. *International Journal of Research in Marketing*. 1996;13(5):463-77. [https://doi.org/10.1016/S0167-8116\(96\)00023-7](https://doi.org/10.1016/S0167-8116(96)00023-7)
- Wilkinson A. Empowerment: theory and practice. *Personnel review*. 1998;27(1):40-56. <https://doi.org/10.1108/00483489810368549>
- Hoffmann WH, Schlosser R. Success factors of strategic alliances in small and medium-sized enterprises—An empirical survey. *Long range planning*. 2001;34(3):357-81. [https://doi.org/10.1016/S0024-6301\(01\)00041-3](https://doi.org/10.1016/S0024-6301(01)00041-3)
- Pfeffer J. Renaissance and renewal in management studies: Relevance regained. *European Management Review*. 2009;6(3):141-8. <https://doi.org/10.1057/emr.2009.13>
- Shoham A. Export performance: A conceptualization and empirical assessment. *Journal of international marketing*. 1998;6(3):59-81. <https://doi.org/10.1177/1069031X9800600308>
- Barney J. Firm resources and sustained competitive advantage. *Journal of management*. 1991;17(1):99-120. <https://doi.org/10.1177/014920639101700108>
- Bilkey WJ, Tesar G. The export behavior of smaller-sized Wisconsin manufacturing firms. *Journal of international business studies*. 1977;8:93-8. <https://doi.org/10.1057/palgrave.jibs.8490783>
- Ortega NRS, Sallum PC, Massad E. Fuzzy dynamical systems in epidemic modelling. *Kybernetes*. 2000;29(2):201-18. <https://doi.org/10.1108/03684920010312768>
- Chang P-T, Pai P-F, Lin K-P, Wu M-S. Applying fuzzy arithmetic to the system dynamics for the customer–producer–employment model. *International journal of systems science*. 2006;37(10):673-98. <https://doi.org/10.1080/00207720600774222>
- Ghazanfari M, Jafari M, Alizadeh S, editors. An approach to solve fuzzy system dynamics problems. *Proceedings of 21st International Conference of the System Dynamics Society*, paper; 2003.
- Xu J, Li X. Using system dynamics for simulation and optimization of one coal industry system under fuzzy environment. *Expert Systems with Applications*. 2011;38(9):11552-9. <https://doi.org/10.1016/j.eswa.2011.03.033>
- Carvalho JP, Tomè JA, editors. Rule based fuzzy cognitive maps-qualitative systems dynamics. *PeachFuzz 2000 19th International Conference of the North American Fuzzy Information Processing Society-NAFIPS (Cat No 00TH8500)*; 2000: IEEE.
- Herrera M, Becerra M, Romero O, Orjuela J, editors. Using system dynamics and fuzzy logic to assess the implementation rfid technology. *Proceedings of the 32st International Conference of System Dynamics Society Delft, Netherlands*; 2014.
- Bahrami S, Bayanati M, Nasiri Janagha MR, Malekian S, Abolghasemian M, Pourghader Chobar A. A Reliable Approach in Solving Multi-Attribute Decision-Making Problems through Fuzzy Rule-Based System and Z-numbers. *Fuzzy Optimization and Modeling Journal*. 2023. <https://doi.org/10.30495/fomj.2023.1998695.1124>
- Mousavi Neqabi SMAA, Morteza, . analysis of the role of cluster development approach in the sustainable development of small and medium enterprises. *National Conference on Future Studies, ; Management and Sustainable Development, Tehran*2018.
- Mousavi Neqabi SM. Challenges and solutions for sustainable development of micro, small and medium enterprises in Tarshiz region based on networking strategy, . *National sustainable development conference (with the approach of investment opportunities and challenges) in Tarshiz region)*; Kashmir, Kashmir Higher Education Center2018.
- Yan MK, Amir and Masoumi, Hamidreza, . Investigating the role of international trade and product export strategy in entering foreign markets and foreign investment performance. *The fourth national conference on community empowerment in the field of humanities and psychological studies, ; Tehran, Center Empowering the Cultural and Social Skills of Society*2019.

26. Mousavi Niaraki SAK, Amir and Masoumi, Hamidreza, . . Identification and prioritization of effective factors in increasing the export of products of knowledge-based companies, . Fourth National Conference on Community Empowerment in the Field of Humanities and Psychological Studies; Tehran, Cultural Skills Empowerment Center and Social Society2019.
27. Ali Ahmadi AaSN, Vahid. A comprehensive description of research methods (paradigms, strategies, quantitative, qualitative and hybrid designs and approaches). Tehran, Knowledge Production Publications2017.
28. Du N, Fathollahi-Fard AM, Wong KY. Wildlife resource conservation and utilization for achieving sustainable development in China: main barriers and problem identification. *Environmental Science and Pollution Research*. 2023;1-20. <https://doi.org/10.1007/s11356-023-26982-7>
29. Nozari H, Najafi E, Fallah M, Hosseinzadeh Lotfi F. Quantitative analysis of key performance indicators of green supply chain in FMCG industries using non-linear fuzzy method. *Mathematics*. 2019;7(11):1020. <https://doi.org/10.3390/math7111020>
30. Moosavi J, Fathollahi-Fard AM, Dulebenets MA. Supply chain disruption during the COVID-19 pandemic: Recognizing potential disruption management strategies. *International Journal of Disaster Risk Reduction*. 2022;75:102983. <https://doi.org/10.1016/j.ijdrr.2022.102983>
31. Gillespie K, Riddle L. Export promotion organization emergence and development: a call to research. *International Marketing Review*. 2004;21(4/5):462-73. <https://doi.org/10.1108/02651330410547144>
32. Antoldi F, Cerrato D, Depperu D. Export consortia in developing countries: successful management of cooperation among SMEs: Springer Science & Business Media; 2012.
33. Bekerman M, Wiñazky M, Moncaut N. Productive policies to strengthen the international insertion of SMEs: Export consortia in Argentina. *Revista Integración y Comercio (Integration and Trade Journal)*. 2013;37(17):55-70. 10.15678/IER.2019.0504.01
34. Chobar AP, Adibi MA, Kazemi A. Multi-objective hub-spoke network design of perishable tourism products using combination machine learning and meta-heuristic algorithms. *Environment, Development and Sustainability*. 2022;1-28. <https://doi.org/10.1007/s10668-022-02350-2>
35. Mohammadpour A. Quality assessment in qualitative research: principles and strategies for validation and generalizability. *Journal of Social Sciences*. 2010;48:73-107.
36. Kosko B. Fuzzy cognitive maps. *International journal of man-machine studies*. 1986;24(1):65-75. [https://doi.org/10.1016/S0020-7373\(86\)80040-2](https://doi.org/10.1016/S0020-7373(86)80040-2)
37. Case Denise M, Blackburn, T., Stylios, C. Modelling Construction Management Problems with Fuzzy Cognitive Maps. In F. Aminah Robinson (Ed.). *Fuzzy Hybrid Computing in Construction Engineering and Management*. 2018:413-49. <https://doi.org/10.1108/978-1-78743-868-220181015>
38. Biloslavo R, Dolinsek S, editors. Scenario planning for climate strategies development by integrating group Delphi, AHP and dynamic fuzzy cognitive maps. PICMET'08-2008 Portland International Conference on Management of Engineering & Technology; 2008: IEEE.
39. Biloslavo R, Grebenc A. Integrating group Delphi, analytic hierarchy process and dynamic fuzzy cognitive maps for a climate warning scenario. *Kybernetes*. 2012;41(3/4):414-28.
40. Poomagal S, Sujatha R, Kumar PS, Vo D-VN. A fuzzy cognitive map approach to predict the hazardous effects of malathion to environment (air, water and soil). *Chemosphere*. 2021;263:127926. <https://doi.org/10.1016/j.chemosphere.2020.127926>
41. Sweeney LB, Sterman JD. Bathtub dynamics: initial results of a systems thinking inventory. *System Dynamics Review: The Journal of the System Dynamics Society*. 2000;16(4):249-86. <https://doi.org/10.1002/sdr.198>
42. Daneshvar A, Radfar R, Ghasemi P, Bayanati M, Pourghader Chobar A. Design of an optimal robust possibilistic model in the distribution chain network of agricultural products with high perishability under uncertainty. *Sustainability*. 2023;15(15):11669. <https://doi.org/10.3390/su151511669>
43. Khorram-Manesh A, Dulebenets MA, Goniewicz K. Implementing public health strategies—the need for educational initiatives: a systematic review. *International Journal of Environmental Research and Public Health*. 2021;18(11):5888. <https://doi.org/10.3390/ijerph18115888>
44. Yuan J, Zou H, Xie K, Dulebenets MA. An assessment of social distancing obedience behavior during the COVID-19 post-epidemic period in China: A cross-sectional survey. *Sustainability*. 2021;13(14):8091. <https://doi.org/10.3390/su13148091>
45. Siegrist M, Luchsinger L, Bearth A. The impact of trust and risk perception on the acceptance of measures to reduce COVID-19 cases. *Risk Analysis*. 2021;41(5):787-800. <https://doi.org/10.1111/risa.13675>
46. Singh P, Pasha J, Moses R, Sobanjo J, Ozguven EE, Dulebenets MA. Development of exact and heuristic optimization methods for safety improvement projects at level crossings under conflicting objectives. *Reliability Engineering & System Safety*. 2022;220:108296. <https://doi.org/10.1016/j.res.2021.108296>

COPYRIGHTS

©2024 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

نبود کنسرسیوم های صادراتی فعال و نبود برنامه فنی، جدی و مدون برای توسعه آنها از مهم ترین دلایل باقی ماندن بنگاه های کوچک و متوسط ایران در مختصات صادراتی کشور است. در این پژوهش، داده ها با رویکرد ترکیبی (کیفی- کمی) که یک پارادایم انتقادی است، جمع آوری و تجزیه و تحلیل می شود. داده ها با استفاده از روش تحقیق کتابخانه ای و میدانی گردآوری شده است. در بخش میدانی از مصاحبه های ساخت یافته، اکتشافی و مشارکتی در مرحله کیفی و از پرسشنامه های محقق ساخته در مرحله کمی استفاده می شود. داده ها با استفاده از نظریه زمینه ای، جلسات طوفان فکری، نقشه شناختی فازی (FCM)، سیستم استنتاج فازی (FIS) و مدل سازی دینامیک سیستم (SDM) تجزیه و تحلیل می شوند. بر اساس نتایج، «ویژگی های اعضای کنسرسیوم»، «برنامه عملیاتی صادرات»، «عامل تقویت کنسرسیوم»، «شناخت حمایت از صادرات»، «عوامل فراملی»، «عوامل دولتی» و «ویژگی محصول» هفت گانه اصلی هستند. عوامل موفقیت کنسرسیوم های صادراتی بخش خصوصی در صنایع ایران. همچنین شناسایی مروج مناسب، شناسایی اعضای بالقوه، انجام مطالعه مورد نظر و تماس با شرکت های ذینفع، تعیین نمایندگان، برگزاری جلسات بین اعضای احتمالی، انجام مطالعات امکان سنجی و تهیه طرح تجاری، تشکیل رسمی کنسرسیوم و پیگیری امور کنسرسیوم. هشت گام برای ایجاد کنسرسیوم های صادراتی بخش خصوصی در صنایع ایران است.



Energy-Conscious Common Operation Scheduling in an Identical Parallel Machine Environment

H. Ataei, F. Ahmadizar*, J. Arkat

Department of Industrial Engineering, University of Kurdistan, Sanandaj, Iran

PAPER INFO

Paper history:

Received 14 October 2023
Received 30 November 2023
Accepted 12 December 2023

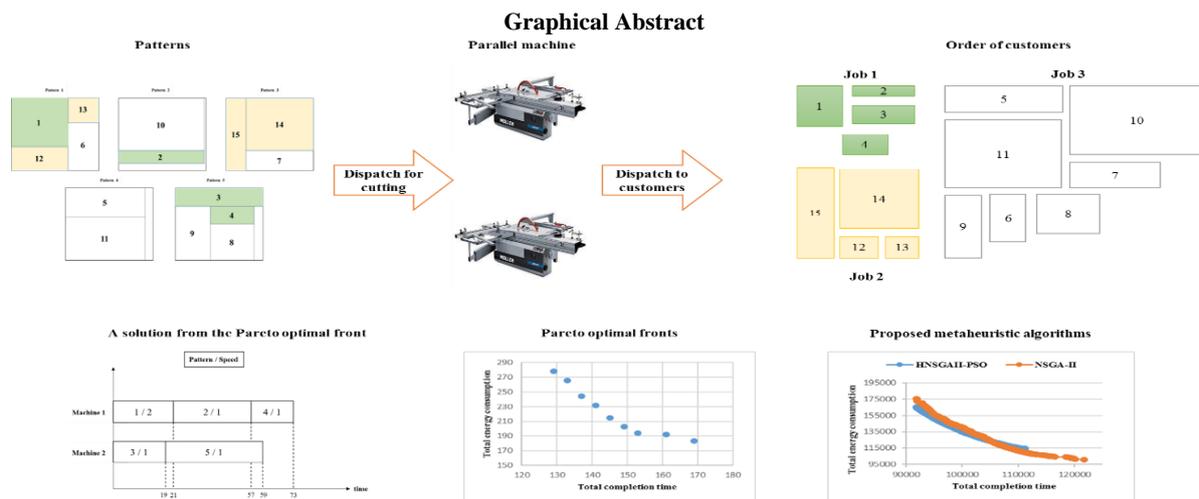
Keywords:

Bi-objective Mixed Integer Linear Programming
Identical Parallel Machine Scheduling
Common Operation
Total Energy Consumption
Total Completion Time

ABSTRACT

The relentless growth of global energy consumption poses a multitude of complex challenges, including the depletion of finite energy resources and the exacerbation of greenhouse gas emissions, which contribute to climate change. In the face of these pressing environmental concerns, the manufacturing sector, a significant energy consumer, is under immense pressure to adopt sustainable practices. The critical intersection of energy consumption management and production operation scheduling emerges as a pivotal domain for addressing these challenges. The scheduling of common operations, exemplified by the cutting stock problem in industries like furniture and apparel, represents a prevalent challenge in production environments. For the first time, this paper pioneers an investigation into an identical parallel machine scheduling problem, taking into account common operations to minimize total energy consumption and total completion time concurrently. For this purpose, two bi-objective mixed integer linear programming models are presented, and an augmented ϵ -constraint method is used to obtain the Pareto optimal front for small-scale instances. Considering the NP-hardness of this problem, a non-dominated sorting genetic algorithm (NSGA-II) and a hybrid non-dominated sorting genetic algorithm with particle swarm optimization (HNSGAI-PSO) are developed to solve medium- and large-scale instances to achieve good approximate Pareto fronts. The performance of the proposed algorithms is assessed by conducting computational experiments on test problems. The results demonstrate that the proposed HNSGAI-PSO performs better than the suggested NSGA-II in solving the test problems.

doi: 10.5829/ije.2024.37.07a.20



*Corresponding Author Institutional Email: f.ahmadizar@uok.ac.ir (F. Ahmadizar)

Please cite this article as: Ataei H, Ahmadizar F, Arkat J. Energy-Conscious Common Operation Scheduling in an Identical Parallel Machine Environment. International Journal of Engineering, Transactions A: Basics. 2024;37(07):1443-65.

1. INTRODUCTION

Energy plays a critical role in human society, and as nonrenewable energy resources continue to deplete, the need for effective and efficient use of energy becomes increasingly vital. The management of energy consumption across diverse processes and sectors has garnered significant attention among researchers owing to its demonstrated ability to mitigate energy expenditures, alleviate environmental repercussions, and enhance energy security (1-4). The production and consumption of energy are major sources of greenhouse gas emissions. The manufacturing sector, which accounts for roughly half of the world's total energy consumption, is a major contributor to greenhouse gas emissions (5-7). It is essential for manufacturing industries to prioritize improving energy efficiency and minimizing greenhouse gas emissions. To achieve this, production managers can choose from a variety of energy-saving strategies. Implementing energy-efficient machines is one potential strategy, although this approach may incur greater upfront costs, which can strain a company's finances. To reduce the impact of this issue, production managers may need to implement scheduling strategies that strike a balance between energy conservation and cost reduction (8). Scheduling problems have been classified into well-known categories based on the machine environment. The most common types of scheduling problems are single machine, parallel machine, flow shop, and job shop. Among these, parallel machine scheduling is particularly significant in scheduling problems because it is a generalization of single machine scheduling and a specific mode of flexible flow shop scheduling (9). Over the past few years, scheduling problems that involve simultaneously obtaining objectives for scheduling and energy have been the subject of extensive research. The related literature has introduced and accepted various electricity consumption strategies, such as power-down and speed-scaling. Furthermore, different policies for determining electricity consumption costs, including fixed, time-of-use (TOU), and tiered pricing, have also been formulated and employed (10-13). Parallel machine scheduling problems that consider energy consumption account for a significant portion of the studies mentioned.

Li et al. (14) investigated the unrelated parallel machine scheduling problem to minimize energy costs and tardiness. In this problem, energy consumption has been considered for various machine modes, including setup, idleness, and processing. For this problem, a mathematical model has been developed so that two objectives are regarded as one single objective. To solve the problem, ten heuristic methods based on the priority rules, combination rules, and energy consumption have been proposed. Che et al. (15) investigated an energy-efficient unrelated parallel machine scheduling problem to minimize total electricity consumption costs. In this

problem, the cost of electricity consumption is calculated using the TOU tariffs, and the makespan is limited. These researchers proposed a two-stage algorithm to solve the problem. Wang et al. (16) studied the parallel machine scheduling problem with a bounded power demand peak. In this problem, the jobs can be processed at different speeds and therefore have various processing times and power demands. The goal of this problem is to minimize the makespan, and to solve it, the researchers proposed a genetic algorithm based on a two-stage heuristic method.

Zeng et al. (17) conducted a comprehensive investigation of a bi-objective optimization problem pertaining to uniform parallel machine scheduling. The aim of their study was to minimize both the number of machines employed and the total electricity cost within the framework of TOU tariffs. To accomplish this, the researchers devised an iterative search framework, facilitating the attainment of the Pareto front. Wu & Che (18) investigated an energy-efficient bi-objective unrelated parallel machine scheduling problem with the goal of minimizing both total energy consumption and makespan. They suggested a memetic differential evolution (MDE) algorithm to solve the problem, and they developed a local search approach to improve the proposed algorithm. Cota et al. (19) conducted an investigation into the unrelated parallel machine scheduling problem, incorporating sequence-dependent setup times to minimize both total electrical energy consumption and makespan. Furthermore, to find solutions close to the Pareto optimal front, they developed and tested a novel heuristic algorithm called Smart Pool. Safarzadeh & Niaki (20) investigated bi-objective green scheduling in uniform parallel machine environments. In this problem, various green cost rates for every machine were considered to model the impact of production resources on sustainability, such as energy consumption and carbon emissions. The researchers modeled the problem with the objective of minimizing the total green costs and makespan, and they used the ϵ -constraint method to identify Pareto optimal solutions. Wang et al. (21) investigated the identical parallel machine scheduling problem with the goal of minimizing both the makespan and total energy consumption. In this problem, the cost of electricity consumption is calculated using the TOU tariffs. The researchers used the augmented ϵ -constraint method to solve the problem in small-scale instances, and they developed a constructive heuristic (CH) method with a local search strategy based on the problem's characteristics to solve the problem in larger-scale instances. Anghinolfi et al. (22) developed an ad hoc heuristic method to solve the problem proposed by Wang et al. (21). This method is divided into two parts. The first section of the method is an improved and modified version of the constructive heuristic (CH) algorithm proposed by Wang et al. (21). The second section introduces a new local search method for

enhancing the efficacy of Pareto solutions. They conducted computational tests to assess the suggested method's efficiency and effectiveness. Zhang et al. (23) delved into a two-stage parallel machine scheduling problem, aiming to minimize total electricity costs under TOU tariffs. Their study considered a two-stage parallel machine system comprising identical parallel speed-scaling machines at stage 1 and unrelated parallel machines at stage 2. The researchers modeled the investigated problem as mixed-integer linear programming and developed a Tabu Search-Greedy Insertion Hybrid (TS-GIH) algorithm to solve it. Keshavarz et al. (24) studied the unrelated parallel machine scheduling problem with sequence-dependent setup times to minimize the energy consumption costs and makespan. They used the ϵ -constraint method to solve the problem in small instances, and they developed the multiple objective simulated annealing and multiple objective particle swarm optimization algorithms to solve the problem in medium and large instances. Zhou & Gu (25) examined the unrelated parallel machine scheduling problem by taking into account multiple resource constraints to minimize total energy consumption and total completion time. They developed a multi-objective artificial immune algorithm to solve the problem. M6dos et al. (26) examined the problem of parallel dedicated machines while keeping energy consumption constraints in mind. In this problem, the peak energy consumption at specific time intervals should not exceed a specified limit. The researchers studied four different variants of the problem and designed a heuristic algorithm for the general problem. Rego et al. (27) proposed a novel bi-objective unrelated parallel machine scheduling problem that considers TOU tariffs and sequence-dependent setup times. To tackle the problem, the researchers suggested a bi-objective mixed-integer linear programming formulation to minimize the total energy consumption and makespan. To solve small and large instances of the problem, they used the weighted sum method and the non-dominated sorting genetic algorithm (NSGA-II), respectively. Asadpour et al. (28) studied the identical parallel machine scheduling problem with the job-splitting property to minimize the total number of tardy jobs and total energy consumption. In this problem, the jobs can be further subdivided. An augmented ϵ -constraint method was used to solve small-scale problems, and a simulated annealing (SA) algorithm was designed to solve medium- and large-scale problems. Heydar et al. (28) proposed an approximate dynamic programming (ADP) approach to address an energy-efficient unrelated parallel machine scheduling problem characterized by random job arrivals. The objective of their work is to minimize a weighted combination of makespan and total energy costs. The energy costs encompass the energy consumption incurred during machine switching, job processing, and idle periods. At

each stage of the ADP, a binary program was formulated to optimize the scheduling problem. The energy efficiency strategy considered in their study is TOU electricity tariffs. Gaggero et al. (29) investigated the problem of bi-objective scheduling on parallel identical machines with TOU costs (BPMSTP). They introduced a new mathematical formulation for the BPMSTP, which allowed them to develop a more efficient exact algorithm for finding the optimal Pareto front. Additionally, they proposed an alternative heuristic approach called the Enhanced Heuristic Scheduler (EHS), which proved to outperform existing heuristics in experiments. Not only did EHS demonstrate superior performance, but it also enhanced the computational efficiency of the exact approach. Sanati et al. (28) conducted a comprehensive investigation into an unrelated parallel machine scheduling problem considering sequence-dependent setup times under TOU electricity tariffs. Their study meticulously examined setup times in two distinct modes: disjointed from and jointed to processing time. For each of these problem variations, two mixed-integer linear programming models were meticulously formulated. The presented models for the problem with setup time disjointed from processing time demonstrated the capability of solving instances involving up to 16 machines and 45 jobs. In contrast, this capability was extended to 20 machines and 40 jobs for the processing time jointed to the setup time problem. Furthermore, to address large-size instances effectively, a fix and relax heuristic algorithm was proposed. This algorithm exhibited the ability to solve instances of up to 20 machines and 100 jobs for each of the two considered problems.

In studies on parallel machine scheduling, after a job is assigned to a machine, that machine processes the job, and the job is ready to be delivered. In the real world, however, there are problems in which each job consists of several sub-jobs, and the job is ready to be delivered when all of its sub-units have been completed after processing one or more activities. Furthermore, processing an activity may have an impact on the completion of multiple jobs. These problems are introduced as "common operation scheduling" (COS) problems. Common operation scheduling problems are used to find the optimal arrangement of operations required by a set of jobs under the assumption that when an activity is completed, it is completed for all jobs that require it (30). These types of problems have different applications, including movie shooting (31), progressive network recovery (32), and pattern sequencing in cutting stock problems (33).

The cutting stock problem includes cutting a set of available pieces in stock (objects) to produce a specific set of smaller pieces (items) that optimizes an objective function such as minimizing total waste, maximizing profit, or minimizing production costs, in order to meet

customer demand for different items. The cutting stock problem occurs in numerous industrial processes where the objects can be sheets of wood, sheets of metal, steel bars, paper or aluminum rolls, printed circuit boards, sheets of glass, etc. In these industries, using appropriate cutting programs is frequently associated with reducing production costs and increasing productivity. A cutting program is a solution to the cutting stock problem, which is proposed by a set of cutting patterns and the frequency of their use. A cutting pattern specifies a subset of items to be obtained from cutting an object. If the items in the cutting problem are two-dimensional or multi-dimensional, then the cutting patterns also determine the arrangement of the items on each of the objects (34). For example, in a furniture factory, sheets of wood must be cut by one or more machines according to predetermined cutting patterns to produce smaller pieces. Each customer order (job) consists of several small pieces that may be placed in one or more different patterns. Therefore, an order is ready to be delivered to the customer when all of the patterns required by that order have been processed by cutting machines.

Several papers in the related literature have attempted to categorize cutting stock problems, which can be used as references to study and gain more information about these categories (35-37). Different criteria can be used to categorize cutting stock problems. The most common criterion is cutting dimensions, used to describe cutting patterns according to the type of problem. Based on this criterion, cutting stock problems are divided into one-dimensional, two-dimensional, three-dimensional, or multi-dimensional problems. According to the literature review, most cutting stock problems are focused on one- and two-dimensional cutting, which can be used in problems such as cutting paper, cables, pipes, sheet wood, etc. (37). In most related studies, the problem not only involves designing cutting patterns and selecting a number of them to produce items smaller than objects (customer orders), but also determining the sequence of patterns to achieve objectives related to completion times, due dates, production costs, etc. Arbib & Marinelli (38) studied the one-dimensional cutting stock problem by considering due dates to minimize the weighted tardiness of the jobs and raw material costs. In this problem, each job has a due date and consists of several pieces of the same size. Researchers developed and tested implicit enumeration, upper bounds, and heuristic methods to solve the problem. Cui et al. (39) studied the one-dimensional cutting stock problem, considering the setup costs. In their study, they developed an integer linear programming model to minimize the sum of setup and material costs over a given pattern set. To solve the problem, they introduced a heuristic algorithm based on sequential grouping to generate patterns. Wuttke & Heese (40) investigated the two-dimensional cutting stock problem with sequence-dependent setup times and

permissible tolerances in the textile industry. For the problem under study, they provided a mixed-integer program, and to solve it, they used a sequential heuristic with a feedback loop based on Gilmore and Gomory's approach.

While mentioning that the optimal cutting patterns for the production of items smaller than objects are designed by commercial software considering the minimum cutting waste, some papers have considered the pre-designed patterns as the input of the problem and have set the sequence of patterns to achieve the objectives regarding completion time and due date (30, 33). Arbib et al. (30) studied the pattern sequencing problem, which is introduced as a common operation scheduling problem, with the goal of minimizing the weighted number of tardy jobs in the single machine environment. They reformulated the problem as a stable set problem on a special graph and analyzed the graph structure. In this problem, the processing times of the patterns on the machine are considered the same and equal to one unit. Arbib et al. (33) investigated the problem of common operation scheduling in single machine and parallel machine environments. In this study, in a single machine environment, the processing times of the patterns on the machine are variable, and in a parallel machine environment, the processing times of the patterns on each of the machines are considered the same and equal to one unit. The researchers formulated the problem as a set-covering problem and solved it using the branch-and-cut algorithm.

Table 1 provides a concise summary of the literature review conducted on energy-efficient parallel machine scheduling. The reviewed articles have been categorized based on the energy consumption strategy employed. To facilitate a comparison of the problem addressed in this article with those investigated in the literature, four criteria have been employed: problem properties, objective function, solution algorithm, and energy consumption strategy. These criteria and their corresponding details are also presented in Table 1. In the realm of energy-efficient parallel machine scheduling problems, an underlying assumption is that the processing of each operation by a machine exclusively affects the completion of a single job. However, this assumption may not always hold true for real-world problems. In certain scenarios, several jobs to be completed may require the common operation to be performed on a shared resource simultaneously, introducing a new dimension of complexity to the scheduling problem that has not been discussed in the literature.

This study focuses on common operation scheduling in an environment of identical parallel machines while considering energy consumption. To this end, two mixed-integer linear programming models, a position-based model and a sequence-based model, have been

TABLE 1. Summary of the Literature Review on Energy-Efficient Parallel Machine Scheduling

Author	Other properties of the problem	Objective/ Solution method	Strategy of energy consumption
[14]	unrelated parallel machine, energy consumption has been considered for various machine modes	tardiness and energy consumption cost/ heuristic algorithms	
[20]	uniform parallel machine	total green costs and makespan/ the ϵ -constraint method	fixed
[25]	unrelated parallel machine, multiple resource constraint	total energy consumption and total completion time/ multi-objective artificial immune algorithm	
[28]	identical parallel machine, job-splitting property	total number of tardy jobs and total energy consumption/ augmented ϵ -constraint, simulated annealing algorithm	
[16]	bounded power demand peak	makespan/ genetic algorithm	
[18]	unrelated parallel machine	total energy consumption and makespan/ memetic differential evolution	speed-scaling
[19]	unrelated parallel machine, sequence-dependent setup times	total energy consumption and makespan/ heuristic algorithm	
[15]	unrelated parallel machine, makespan is limited	energy consumption cost/ heuristic algorithm	
[17]	uniform parallel machine	number of machines employed and the total electricity cost/ heuristic algorithm	
[21]	identical parallel machine	makespan and total energy consumption/ augmented ϵ -constraint method, constructive heuristic, NSGA-II	
[23]	two-stage parallel machine	total energy consumption/ Tabu Search-Greedy Insertion Hybrid algorithm	
[24]	unrelated parallel machine, sequence-dependent setup times	energy consumption costs and makespan/ the ϵ -constraint method, multiple objective simulated annealing algorithm and multiple objective particle swarm optimization algorithms	TOU
[26]	parallel dedicated machines, peak energy consumption at specific time intervals should not exceed a specified limit	makespan/ heuristic algorithm	
[27]	unrelated parallel machine, sequence-dependent setup times	total energy consumption and makespan/ weighted sum method, NSGA-II	
[29]	unrelated parallel machine, random job arrivals	weighted combination of makespan and total energy costs/ approximate dynamic programming	
[30]	identical parallel machine	total energy consumption and makespan/ Enhanced Heuristic Scheduler	
[31]	unrelated parallel machine, sequence-dependent setup times, bounded makespan	total electricity cost/fix and relax heuristic algorithm	
Current paper	identical parallel machine, common operation scheduling	total energy consumption and total completion time/ augmented ϵ -constraint method, HNSGAI-PSO, NSGA-II	Speed-scaling

proposed for the problem under study in order to simultaneously minimize the total energy consumption and the total completion time. To solve small-scale instances, the augmented ϵ -constraint method (AUGMECON) is used to obtain the Pareto optimal front. Since the problem is NP-hard, a non-dominated sorting genetic algorithm (NSGA-II) and a hybrid non-dominated sorting genetic algorithm with particle swarm optimization (HNSGAI-PSO) have been developed to solve medium- and large-scale instances. In the proposed NSGA-II algorithm, each chromosome represents a solution from the problem-solving space, and the quality

of each chromosome can be evaluated by calculating the objective functions of total completion time and total energy consumption. In the proposed HNSGAI-PSO algorithm, each chromosome represents a region of the problem-solving space where all solutions in this region have the same total energy consumption. To evaluate the quality of each chromosome, the best solution for the region covered by that chromosome is considered. In the investigated problem, the best solution in the region covered by each chromosome is the one that minimizes the total completion time. Therefore, to determine the best solution in each region, the PSO algorithm is

executed once for each chromosome. The total completion time for each chromosome of the HNSGAI-PSO algorithm is determined based on the output of the PSO algorithm. According to the mentioned explain, the contributions of the paper can be summarized as follows:

- Energy consumption is considered in the common operation scheduling problem.
- Two bi-objective mixed integer linear programming models, namely the position-based and sequence-based models, are proposed to investigate the trade-off between total completion time and total energy consumption.
- The NSGA-II and HNSGAI-PSO algorithms are developed to solve large-scale instances.
- The performance of the methods is evaluated using computational experiments.

The remaining sections of this paper are as follows: After the problem is described in section 2, mathematical models are proposed. In section 3, solution methods will be presented. Section 4 discusses the results of the computations. In section 5, the conclusion and future research are mentioned.

2. PROBLEM DESCRIPTION AND MATHEMATICAL MODELING

The common operation scheduling problem is one of the scheduling problems with numerous applications in the actual world. One of the applications of common operation scheduling in manufacturing environments is the cutting stock problem, which is posed in industries such as the furniture industry (cutting wooden panels) and the apparel industry (cutting fabric). This paper investigates common operation scheduling in an identical parallel machine environment, considering energy consumption. In the problem under study, each job includes several small pieces (items), and all the pieces of different jobs are placed on a number of cutting patterns to produce the required small pieces according to these patterns and by cutting larger pieces (objects). Therefore, each job is completed when all the small pieces related to it have been produced by cutting the required patterns. The relationship between jobs and cutting patterns is shown by the job-pattern matrix. In fact, this matrix indicates which patterns must be processed to complete each job. The optimal cutting patterns are predetermined using relevant software, taking the dimensions of small pieces into account, and will be available at time 0. Each pattern contains one or more pieces and contributes to the completion of one or more jobs. Energy consumption is investigated by considering the speed-scaling strategy. In accordance with this strategy, each machine possesses varying speed levels. Consequently, when the machine operates at a higher speed, the processing time is reduced while the

consumption of electrical energy increases. In this problem, the assignment of cutting patterns to machines, the sequence of patterns on each machine, and the appropriate speed of the machine to process each of the assigned patterns are determined in order to simultaneously minimize the total completion time and the total energy consumption. The underlying assumptions are listed below:

- Cutting patterns have already been designed and determined and will be available at time zero.
- Larger pieces (objects) will be available at time 0 to produce smaller parts (items), and their number is equal to the number of cutting patterns.
- Each machine processes a maximum of one pattern at a time.
- Each pattern can be processed by only one machine at a time.
- The processing time of each pattern at the varying speed levels of each machine is specified and definite.
- The amount of energy consumed by each machine at various speed levels for processing each pattern is specified and definite.
- There is no precedence relationship between different patterns.
- There is no preemption in the processing of patterns.

In the following, the notations and parameters are presented.

Sets and Indices

I	Set of jobs $\{i \in I\}$.
J	Set of cutting patterns $\{j, k \in J\}$.
$V = J \cup \{0\}$	Set of cutting patterns that includes the fictitious pattern 0 $\{j, k \in V\}$.
M	Set of machines $\{m \in M\}$.
Q	Set of positions on each machine $\{q \in Q\}$.
S	Set of speed levels of each machine $\{s \in S\}$.
N_i	Subset of cutting patterns set that must be processed to complete job i ($\cup_i N_i = J$).

Parameters

w_j	The required workload of pattern j .
v_s	Machine processing speed at speed level s (the workload processed per unit of time by the machine at speed level s).
p_{js}	The time required to process pattern j at speed level s of the machine ($p_{js} = \frac{w_j}{v_s}$).
π_j	Machine energy consumption rate for processing pattern j per unit of time.
π_{js}	Machine energy consumption rate at speed level s for processing pattern j per unit of time ($\pi_{js} = \pi_j v_s^\alpha$, $\alpha > 1$).
e_{js}	The energy required to process pattern j at speed level s of the machine ($e_{js} = \pi_{js} p_{js}$).
B	A large number.

The **sequence-based** mathematical model is presented here.

Decision variables	
x_{mkjs}	Equal to 1 if pattern j is processed immediately after pattern k on machine m with speed level s , and 0 otherwise.
r_j	Completion time of pattern j .
c_i	Completion time of job i .

Sequence-based model	
$\min Z_1 = \sum_{i \in I} c_i$	(1)
$\min Z_2 = \sum_{m \in M} \sum_{k \in V: j \neq k} \sum_{j \in J} \sum_{s \in S} e_{js} x_{mkjs}$	(2)
$\sum_{m \in M} \sum_{k \in V: j \neq k} \sum_{s \in S} x_{mkjs} = 1 \quad \forall j \in J$	(3)
$\sum_{m \in M} \sum_{j \in J: j \neq k} \sum_{s \in S} x_{mkjs} \leq 1 \quad \forall k \in J$	(4)
$\sum_{j \in J} \sum_{s \in S} x_{m0js} \leq 1 \quad \forall m \in M$	(5)
$\sum_{j \in V: k \neq j} \sum_{s \in S} x_{mkjs} - \sum_{h \in V: h \neq j} \sum_{s \in S} x_{mhks} = 0 \quad \forall k \in J, m \in M$	(6)
$r_j - r_k + B(1 - x_{mkjs}) \geq p_{js} \quad \forall k \in V: j \neq k, j \in J, m \in M, s \in S$	(7)
$c_0 = 0$	(8)
$c_i = \max_{j \in N_i} r_j \quad \forall i \in I$	(9)
$x_{mkjs} = \{0, 1\} \quad \forall j, k \in V, m \in M, s \in S$	(10)
$c_i, r_j \geq 0 \quad \forall j \in V, i \in I$	(11)

The objective functions 1 and 2 represent the minimization of total completion time and total energy consumption for the sequence-based model, respectively. Constraint 3 guarantees that each pattern is assigned to only one machine at a specified speed level and that only one other pattern is processed before it. Constraint 4 indicates that after each pattern, a maximum of one other pattern can be processed. Constraint 5 ensures that in each machine, after fictitious pattern 0, a maximum of one other pattern can be processed. Constraint 6 ensures the right order for allocating patterns in each machine: if pattern k is processed before pattern j , then another pattern must be processed before pattern k . Constraint 7 computes the completion time of each pattern. If $x_{mkjs} = 1$, then the completion time of pattern j is obtained from the sum of the completion time of pattern k and the processing time of pattern j with the speed level s of the machine, and if $x_{mkjs} = 0$, then the large number B ensures the relation. Equation 8 shows that the completion time of the fictitious pattern 0 is zero. Each job's completion time is calculated by constraint 9. The completion time of job i is equal to the maximum

completion time of patterns that include the pieces of job i . Constraints 10 and 11 show the range of decision variables.

Considering Equation 9, the proposed model is mixed integer nonlinear programming. The proposed sequence-based model is converted into a mixed integer linear programming model by substituting Equation 12 for Equation 9.

$$c_i \geq r_j \quad \forall i \in I, j \in N_i \quad (12)$$

The **position-based** mathematical model is presented here.

Decision variables	
x_{jmqs}	Equal to 1 if pattern j is processed in position q of machine m with speed level s , and 0 otherwise.
h_{qm}	Start time of position q of machine m .
f_{qm}	Finish time of position q of machine m .
r_{jm}	Completion time of pattern j on machine m .
c_i	Completion time of job i .

Position-based model	
$\min Z_1 = \sum_{i \in I} c_i$	(13)
$\min Z_2 = \sum_{j \in J} \sum_{m \in M} \sum_{q \in Q} \sum_{s \in S} e_{js} x_{jmqs}$	(14)
$\sum_{m \in M} \sum_{q \in Q} \sum_{s \in S} x_{jmqs} = 1 \quad \forall j \in J$	(15)
$\sum_{j \in J} \sum_{s \in S} x_{jmqs} \leq 1 \quad \forall m \in M, q \in Q$	(16)
$f_{qm} = h_{qm} + \sum_{j \in J} \sum_{s \in S} p_{js} x_{jmqs} \quad \forall m \in M, q \in Q$	(17)
$f_{qm} \leq h_{(q+1)m} \quad \forall m \in M, q \in Q$	(18)
$\sum_{j \in J} \sum_{s \in S} x_{jmqs} \leq \sum_{j \in J} \sum_{s \in S} x_{jm(q-1)s} \quad \forall m \in M, q \in Q: q > 1M, s \in S$	(19)
$f_{qm} \leq r_{jm} + B(1 - \sum_{s \in S} x_{jmqs}) \quad \forall j \in J, m \in M, q \in Q$	(20)
$c_i = \max_{j \in N_i, m \in M} r_{jm} \quad \forall i \in I$	(21)
$x_{jmqs} = \{0, 1\} \quad \forall j \in J, m \in M, q \in Q, s \in S$	(22)
$c_i, r_{jm}, f_{qm}, h_{qm} \geq 0 \quad \forall j \in J, m \in M, q \in Q, i \in I$	(23)

The position-based model's objectives are to minimize total completion time and total energy consumption, which are demonstrated by Equations 13 and 14, respectively. Equation 15 guarantees that each pattern is processed only at one position on one machine and at a specified speed level. Constraint 16 indicates that at each position of each machine, a maximum of one

pattern is processed at a specified speed level. Constraints 17 and 18 compute the start time and end time of each position of each machine. Constraint 19 ensures that the positions of each machine are assigned to patterns in numerical order from first to last. Each pattern's completion time is calculated by constraint 20. The completion time of each job is calculated using Equation 21. The completion time of job i is equal to the maximum completion time of patterns that include the pieces of job i . Constraints 22 and 23 show the range of decision variables.

Considering Equation 21, the proposed model is mixed-integer nonlinear programming. In order to linearize the mathematical model, Equation 24 is used instead of Equation 21. Due to the changes made, the proposed position-based model is a mixed integer linear programming model.

$$c_i \geq r_{jm} \quad \forall i \in I, j \in N_i, m \in M \quad (24)$$

Based on the triple notation provided by Wuttker and Heese (41), the problem studied in this paper is denoted as $Pm|cos|\sum TEC, \sum C_i$, where Pm indicates identical parallel machines, cos indicates common operation, and TEC and $\sum C_i$ represent the total energy consumption and the total completion time, respectively. In this problem, if the machines have only one level in terms of processing speed and the objective function of the total energy consumption is not considered, and if each pattern is only effective in completing one job and the completion of each job only requires the processing of one pattern, then the problem becomes an identical parallel machine problem in order to minimize the total completion time and is shown as $Pm||\sum C_i$. According to the previous studies, the problem $Pm||\sum C_i$ is the NP-hard (42). Therefore, it can be concluded that the problem proposed in this article is at least NP-hard.

3. NUMERICAL EXAMPLE

To further illustrate the problem described, consider a small numerical example involving three jobs and two identical machines. In this example, *job 1* includes pieces {1,2,3,4}, *job 2* includes pieces {5,6,7,8,9,10,11}, and *job 3* includes pieces {12,13,14,15}, and each machine has two levels of slow speed (level 1) and fast speed (level 2). Figure 1 shows the placement of small pieces (items) on objects using five cutting patterns. Table 2 lists the processing time and energy required to process patterns at different speed levels for each machine. Figure 2 shows the job-pattern matrix, which is formed based on the placement of small pieces of each job and their relationship with cutting patterns. In other words, this matrix determines the patterns associated with each job.

Figure 3 shows a solution from the Pareto optimal front for this example. In this solution, the first machine

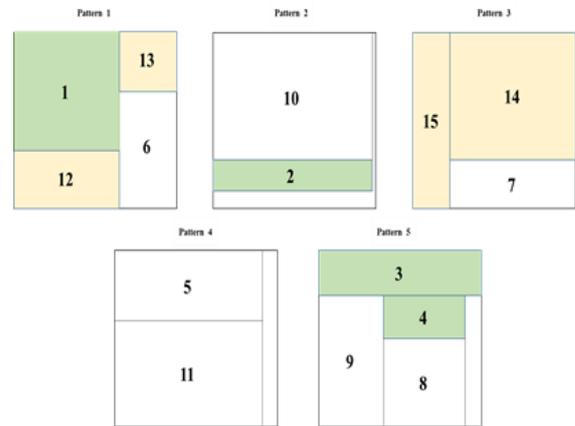


Figure 1. Cutting patterns for the numerical example

TABLE 2. Energy consumption and processing time to process patterns for the numerical example

		Pattern					
		1	2	3	4	5	
Machine	Level 1 (slow)	time	27	36	19	16	40
		energy	18	61	20	16	68
	Level 2 (fast)	time	21	28	15	12	32
		energy	29	95	32	25	97

		pattern				
		1	2	3	4	5
job	1	■	■	□	□	■
	2	■	■	■	■	■
	3	□	□	■	□	□

Figure 2. The job-pattern matrix for the numerical example

processes pattern 1 with speed level 2, then it processes, respectively, patterns 2 and 4 with speed level 1. Patterns 3 and 5 are assigned to the second machine and processed there at speed level 1 of the machine, respectively. In this solution, based on the completion times of the patterns and the job-pattern matrix, job 1 is completed at time 59, job 2 at time 73, and job 3 at time 21. Therefore, in this solution, the total completion time is 153 time units, and the energy consumption is 194 units. In Figure 4, the Pareto optimal front is provided for this example. The analysis demonstrates that the optimal speed of the machines depends on the objective function. If the objective function seeks to minimize the total completion time, then processing all patterns at speed level 2 yields the best result. However, this decision increases the objective function of total energy consumption. In

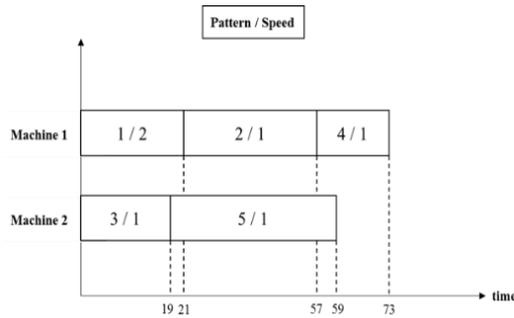


Figure 3. A solution from the Pareto optimal front for the numerical example

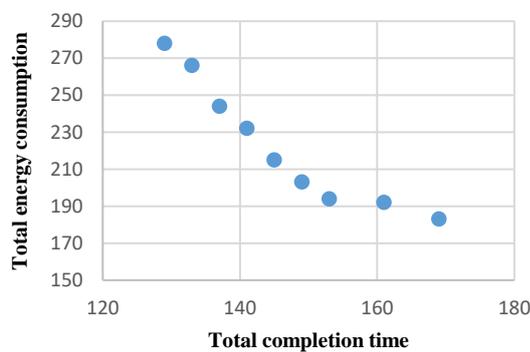


Figure 4. The Pareto optimal front for the numerical example

contrast, when all patterns are processed at speed level 1, the objective function of total energy consumption reaches its minimal value. Thus, decision-makers must consider the trade-off between the objectives and identify the best options for achieving the desired balance between the goals.

4. SOLUTION APPROACHES

In solving bi- and multi-objective problems, the primary goal is to identify a set of Pareto optimal solutions, considering trade-offs between objectives. Due to the NP-hard nature of the problem under investigation, a non-dominated sorting genetic algorithm (NSGA-II) and a hybrid non-dominated sorting genetic algorithm with particle swarm optimization (HNSGAI-PSO) are developed to tackle instances with medium and large scales and to obtain approximate Pareto fronts

4. 1. NSGA-II Algorithm The NSGA-II algorithm (43) is a popular evolutionary algorithm that is widely used in solving multi-objective optimization problems (44, 45). This algorithm can provide a suitable set of Pareto solutions for solving multi-objective problems by

using the elitism mechanism and taking the crowding distance of the solutions into account.

In the proposed NSGA-II algorithm, each chromosome represents a solution from the problem-solving space. In this algorithm, by decoding each chromosome, the cutting patterns assigned to each machine, the sequence of patterns on each machine, and the speed level of the machine for processing each pattern are determined, and using this information, the objective functions of total completion time and total energy consumption can be calculated for each chromosome. After calculating the value of the objective functions and determining the fitness for all chromosomes in the population, the sorting of chromosomes is done. This process begins with performing paired comparisons and calculating the number of times each solution is dominated in order to form Pareto fronts and determine the rank of each solution. Then the members placed in each Pareto front are sorted based on the crowding distance metric. Next, offspring are produced by selecting parents and using crossover and mutation operators. The offspring obtained from the mutation and crossover operators are added to the population, creating a larger population called R_t . After calculating the objective functions for the generated offspring, rank and crowding distance are determined for each member of the R_t population, and based on them, the R_t population is sorted. The sorting is done as follows: first, the members are sorted by rank and in ascending order so that the solutions with lower ranks are placed at the beginning of the list. Then, among the members with the same rank, another sorting is done based on the crowding distance and in a descending manner, so that the solution with the greatest crowding distance occupies a higher position among the members of the same rank. Finally, the elitist strategy is used to form the new generation, in which members with a higher position are selected from the R_t population in a number equal to the size of the population. This process continues until the termination condition of the algorithm is established. The proposed NSGA-II algorithm stops when a certain number of iterations is reached, and all the solutions in the first Pareto front are presented as the output of the algorithm.

4. 1. 1. Solution Representation In order for the solution algorithm to establish a logical relationship between the problem space and the search space, the solution properties should be represented by a string of symbols. Each chromosome of the NSGA-II algorithm is represented by a two-row matrix, where the length of each row is n (n : number of cutting patterns). The first row of the matrix contains the permutation of numbers 1 to n , so that the members of the set $\{1,2,3, \dots, n\}$ represent the cutting pattern number. The second row of the matrix shows the speed level of the machine for processing each of the cutting patterns, where each gene is coded

randomly with values from the set $\{1,2,3, \dots, s\}$. The cutting patterns assigned to each machine and the sequence of patterns on each machine are determined by decoding each chromosome. Extracting the required information is as follows: the pattern placed in the first gene of the first row of the matrix is assigned to the first machine and processed at the speed determined in the first gene of the second row of the matrix. This process continues until every machine is assigned a pattern. Then, the next pattern is assigned to the first machine that finishes processing the previously assigned pattern and is processed at the determined speed. This process continues until all patterns are assigned. For example, consider a problem with two machines and five cutting patterns. Each machine has two levels of slow speed (level 1) and fast speed (level 2) to process the assigned patterns. It is assumed that each pattern needs 2 and 1 units of time for processing at the machine speed levels of 1 and 2, respectively. Figure 5 shows a chromosome for the mentioned example. After decoding this solution, it becomes clear that pattern 3 is assigned to the first machine and is processed with a speed level of 2 in 1 time unit. Pattern 1 is assigned to the second machine and processed at speed level 1 in 2 time units. Since processing pattern 3 is finished faster by machine 1, pattern 2 is assigned to the first machine and is processed at speed level 1 in 2 time units. Next, patterns 5 and 4 are assigned to machines 2 and 1, and are processed at speed levels 1 and 2, respectively. In the proposed NSGA-II algorithm, members of the initial population are generated randomly and according to the presented representation for each solution.

4. 1. 2. Selection In the proposed NSGA-II algorithm, parent selection is done using the standard binary tournament selection strategy.

4. 1. 3. Crossover In the proposed NSGA-II algorithm, the double-point crossover operator is used. In view of the permutation of numbers 1 to n in the first row of the chromosome, the permutation of numbers may not be established in the first row of each of the generated offspring. Therefore, using the partially mapped crossover (PMX) approach, the required columns are moved, and the necessary corrections are made to establish the permutation of the numbers in the first row of the generated offspring. Figure 6 depicts the double-point crossover.

3	1	2	5	4
2	1	1	1	2

Figure 5. An example of solution representation in the NSGA-II algorithm

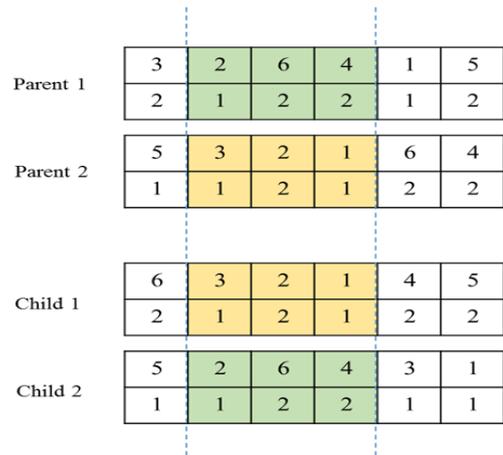


Figure 6. The crossover operator in the NSGA-II algorithm

4. 1. 4. Mutation In the proposed NSGA-II algorithm, Gaussian mutation and swap mutation operators are used.

The Gaussian mutation operator is applied to the second row of the candidate chromosome. The Gaussian mutation operates as follows: A random number between 0 and 1 is generated for each gene from the second row of the candidate chromosome. If this number is less than the gene's mutation rate, the mutation operator with $N(0,1)$ distribution is applied to that gene. Thus, the value of that gene may change. According to the chromosome's representation, the value of each gene in the second row of the chromosome can be one of the numbers in the set $S = \{1,2,3, \dots, s\}$. If, after applying the mutation operator, the value of the gene is greater than the largest value of the set S or less than the smallest value of that set, then we consider the value of the mentioned gene to be equal to the largest and smallest members of the set S , respectively. In Figure 7, the Gaussian mutation operator has been applied to three genes from the parent chromosome, and in the generated offspring, only the value of one of the selected genes has changed.

The swap mutation is executed as follows: Two columns are randomly chosen from the candidate chromosome, and they are then exchanged with one another Figure 8. Based on the results obtained from the numerical experiments, the Gaussian mutation and swap mutation operators are applied to the candidate chromosomes with probabilities of 0.8 and 0.2, respectively.

4. 2. HNSGAI-PSO Algorithm The proposed hybrid algorithm (HNSGAI-PSO) is a combination of the NSGA-II algorithm and the PSO algorithm. In the HNSGAI-PSO algorithm, each chromosome determines the speed level of the machine for processing each pattern, and this information can be used to calculate the

Parent	3	2	6	4	1	5
	2	1	2	2	1	2
Child	3	2	6	4	1	5
	1	2	2	2	1	2

Figure 7. The Gaussian mutation operator in the NSGA-II algorithm

Parent	3	2	6	4	1	5
	2	1	2	2	1	2
Child	3	5	6	4	1	2
	2	2	2	2	1	1

Figure 8. The swap mutation operator in the NSGA-II algorithm

objective function of the total energy consumption for that chromosome. In other words, each chromosome represents a region of the problem-solving space, so that all solutions in this region have the same total energy consumption. To evaluate the quality of each chromosome, the best solution from the region covered by that chromosome is considered. In the investigated problem, the best solution in the region covered by each chromosome is the one that minimizes the total completion time. Using the PSO algorithm, a global search is done to find the best solution in this region. Therefore, in each iteration of the hybrid algorithm, the PSO algorithm is executed once for each chromosome in the population. Each particle of the PSO algorithm specifies the assignment of patterns to machines as well as the sequence of patterns on each machine, and using this information and taking into account the speed level of the machine to process each pattern, the total completion time can be calculated for each particle of the PSO algorithm. Based on the output of the PSO algorithm, the assignment of patterns to machines, the sequence of patterns on each machine, and the objective function of the total completion time for each chromosome of the HNSGAI-PSO algorithm are determined. The main structure and the process of obtaining the approximate Pareto front in the HNSGAI-PSO algorithm are similar to the NSGA-II algorithm.

4. 2. 1. PSO Algorithm

The PSO algorithm is a population-based algorithm (46) and is widely used in solving scheduling problems (47, 48). Within this algorithm, each solution is conceptualized as a particle, each possessing its own position and velocity. The position of a particle facilitates the identification of a feasible solution to the problem under investigation. The

fitness value of each particle determines the quality of the corresponding solution. During each iteration of the algorithm, each particle endeavors to locate a new position based on its previous experiences and the position of the particle exhibiting the most favorable fitness value, thereby striving to enhance its own fitness value (49). In the HNSGAI-PSO algorithm proposed in this paper, for each chromosome, the PSO algorithm is executed once. To update the speed and position of each particle in the PSO algorithm, relations (25) and (26) are used, respectively.

$$v_j^{t+1} = \omega v_j^t + r_1 \times c_1 \times (pbest_j - x_j^t) + r_2 \times c_2 \times (gbest - x_j^t) \tag{25}$$

$$x_j^{t+1} = x_j^t + v_j^{t+1} \tag{26}$$

where:

- v_j Velocity of particle j .
- x_j Position of particle j .
- $pbest_j$ The best position ever visited by particle j .
- $gbest$ The best position ever visited by swarm.
- ω Inertia coefficient that controls the velocity of the particle.
- c_1 The learning coefficient considers the current particle's attraction to its previous best position.
- c_2 The learning coefficient considers the current particle's attraction to the previous best position of the swarm
- r_1, r_2 Random numbers are generated from the uniform interval [0,1].

4. 2. 2. Solution Representation

In the proposed HNSGAI-PSO algorithm, each chromosome contains a vector of length n , where n is the number of cutting patterns. The counter of each chromosome gene indicates the cutting pattern number, and the genes of each chromosome are randomly coded with values from the set $\{1, 2, 3, \dots, s\}$ that determine the speed level of the machine for processing the relevant cutting patterns, which can be used to calculate the objective function of the total energy consumption for each chromosome. In other words, each chromosome represents a region of the problem-solving space where all solutions in this region have the same total energy consumption. To evaluate the quality of each chromosome, the best solution from the region covered by that chromosome is considered. In the investigated problem, the best solution in the region covered by each chromosome is the one that minimizes the total completion time. The best solution in this region is determined using the PSO algorithm. Therefore, in each iteration of the hybrid algorithm, the PSO algorithm is executed once for each chromosome in the population.

Each particle in the PSO algorithm is represented as a matrix with one row and n columns, where n is the number of cutting patterns. A random real number between 0 and 1 is placed in each column of this matrix. For each particle, in order to determine the cutting patterns assigned to each machine and the sequence of patterns on each machine, first, through an appropriate mechanism, the real numbers must be converted into integers that represent the cutting patterns. To do this, we used the sorting-based method (50). Therefore, we sort the real numbers of the initial matrix in ascending order to form a new matrix. Then, in the new matrix, any real number is replaced by the counter of the corresponding column in the initial matrix. Accordingly, the new matrix contains a permutation of integers from 1 to n, and each number represents a cutting pattern. Now, by decoding the created matrix, the patterns assigned to each machine and the sequence of patterns on each machine can be determined. Using this information and taking into account the properties of the relevant chromosome (the speed level of the machines), the function of the total completion time for each particle of the PSO algorithm can be calculated. The output of the PSO algorithm determines the assignment of patterns to machines, the sequence of patterns on each machine, and the objective function of the total completion time for each chromosome of the HNSGAI-PSO algorithm.

Each particle of the PSO algorithm is decoded as follows: the pattern placed in the first column of the matrix is assigned to the first machine and processed at the corresponding speed. The pattern placed in the second column is assigned to the second machine and processed at the corresponding speed, and this process continues until each machine is assigned a pattern. Then, the next pattern is assigned to the first machine that finishes processing the previously assigned pattern, and this process continues until all patterns are assigned.

For example, consider a problem with two machines and five cutting patterns. Each machine has two levels of slow speed (level 1) and fast speed (level 2) to process the patterns assigned to it, so that the faster the machine processes, the more energy it consumes. It is assumed that each pattern needs 2 units and 1 unit of time to be processed at machine speed levels 1 and 2, respectively. Figure 9 is a chromosome for the mentioned example, which shows a region of the solution space. For all solutions in this region, cutting patterns numbers 1, 2, 3, 4, and 5 are processed at machine speed levels 2, 1, 1, 2, and 1, respectively. Figure 10 is a particle of the PSO algorithm and its decoding, which shows a solution of the region covered by Figure 9's chromosome. After decoding this particle, it becomes clear that pattern 3 is assigned to the first machine and is processed with a speed level 1 in 2 time units; pattern 1 is assigned to the second machine and processed at a speed level 2 in 1 time unit. Since the processing of pattern 1 is completed faster

by machine 2, pattern 2 is assigned to the second machine and is processed at speed level 1 in 2 time units. Next, patterns 5 and 4 are assigned to machines 1 and 2, and they are processed at speed levels 1 and 2, respectively.

In the proposed HNSGAI-PSO algorithm, each chromosome from the initial population is randomly generated according to the representation provided for each solution. The PSO algorithm is executed once for every member of the initial population. The initial population members of the PSO algorithm are also randomly generated.

4. 2. 3. Selection Parent selection in the proposed HNSGAI-PSO algorithm is done using the standard binary tournament selection strategy.

4. 2. 4. Crossover The double-point crossover operator is employed in the proposed HNSGAI-PSO algorithm. Figure 11 shows the double-point crossover.

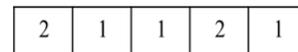


Figure 9. An example of solution representation in the HNSGAI-PSO algorithm

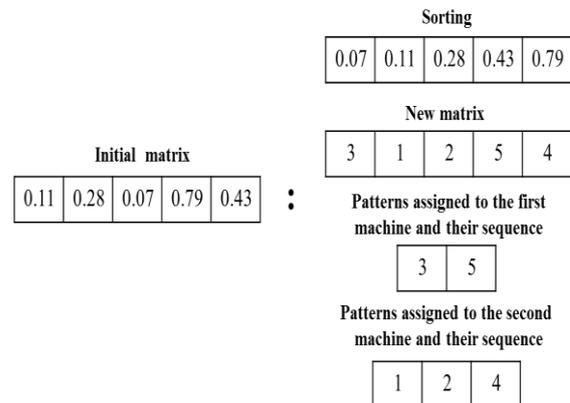


Figure 10. A solution of the region covered by the chromosome in Figure 9 and its decoding

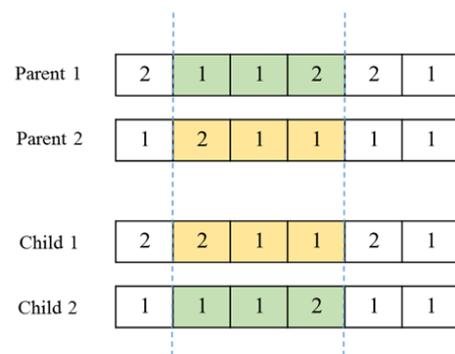


Figure 11. The crossover operator in the HNSGAI-PSO algorithm

4. 2. 5. Mutation

In the proposed hybrid algorithm, the Gaussian mutation operator is used. In this algorithm, according to the representation defined for the chromosome, all chromosomes in the population are candidates for the generation of offspring through the mutation operator. Next, for each gene on the candidate chromosome, a random number between 0 and 1 is generated. If this number is less than the mutation rate of the gene, the mutation operator with the standard normal distribution $N(0,1)$ is applied to that gene. Accordingly, the value of that gene may change in the generated offspring. Figure 12 shows how to produce offspring from the parent using the Gaussian mutation operator. In this figure, the mutation operator has been applied to three genes from the parent chromosome, and in the offspring generated, only the value of one of the selected genes has changed.

5. COMPUTATIONAL EXPERIMENT

In this section, in order to evaluate the performance of the mathematical models and solution methods presented, numerical experiments are performed using randomly generated test instances. The process of generating the necessary data for the instances is explained in detail. Given the bi-objective nature of the problem, performance measures used to evaluate the suggested solution methods are introduced. Moreover, the parameters of the proposed solution algorithms are tuned using the Taguchi method. Finally, the performance of the proposed mathematical models and solution methods is evaluated by solving small and large-scale instances. In this paper, to optimally solve small-scale instance problems using the AUGMECON method, GAMS 28.2.0 software is used, and the HNSGAI-PSO and NSGA-II algorithms are coded in the Visual C# environment. All experiments were performed on a personal computer with 4 GB of RAM and an Intel Core i5-2410M 2.30 GHz CPU.

5. 1. Data Generation

The size of instance problems is defined as $[I, J, M, \delta]$, where I represents the number of jobs, J denotes the number of cutting patterns, M shows the number of machines, and δ represents the density of the job-pattern matrix. In this paper, random test instances are generated as follows (18): To determine the required workload of each cutting pattern (w_j), a

number is generated randomly from a uniform distribution in the interval $[5, 50]$. The energy consumption rate of the machine for processing each cutting pattern per unit of time (π_j) is randomly selected from a uniform distribution in the interval $[4, 18]$. Therefore, the energy consumption rate of the machine at the speed level s to process pattern j per unit of time is determined from Equation 27. In this paper, $\alpha = 3$ and for machine processing speed levels, $S = 4$ and $v_s \in \{0.75, 1, 1.25, 1.5\}$ are considered. The time and energy required to process each cutting pattern at varying levels of machine speed are calculated using relations 28 and 29, respectively.

$$\pi_{js} = \pi_j v_s^\alpha, \quad \alpha > 1 \tag{27}$$

$$p_{js} = \frac{w_j}{v_s} \tag{28}$$

$$e_{js} = \pi_{js} p_{js} \tag{29}$$

To randomly generate the job-pattern matrix, first, the density δ of the matrix must be determined. Equation 30 calculates the value of δ for the job-pattern matrix. In this equation, $|N_i|$ is the number of cutting patterns that include small pieces (items) of job i . i and j also represent the number of jobs and the number of cutting patterns, respectively. In determining the minimum value of δ for test instances, a noteworthy point to consider is that each cutting pattern contributes to completing at least one job, and the pieces related to each job are placed in at least one cutting pattern. Therefore, the minimum possible value for δ is equal to $\max\left(\frac{1}{i}, \frac{1}{j}\right)$, where i is the number of jobs and j is the number of cutting patterns, and this relation should be considered to determine the value of δ in small-scale instances. In medium- and large-scale instances, we consider three values of 20%, 30%, and 40% for δ (33). After determining the value of δ , we generate the job-pattern matrix at random so that each cutting pattern contributes to the completion of at least one job and each job requires at least one cutting pattern to complete itself.

$$\delta = \frac{\sum_i |N_i|}{i \times j} \tag{30}$$

5. 2. Performance Measures

The Pareto solution set is the output of solving multi-objective problems, whose quality and diversity are evaluated. In this paper, to evaluate the results obtained from the solution algorithms, four performance measures are used as follows:

Number of Pareto solutions (NPS): This metric specifies the number of non-dominated solutions obtained from the solution algorithm. Based on this metric, the greater the number of these solutions, the more efficient the algorithm (21, 51).

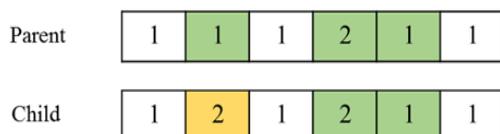


Figure 12. The mutation operator in the HNSGAI-PSO algorithm

Diversity of distribution (D_1): This performance measure shows that the existing solutions in the Pareto front are uniformly placed next to each other. This metric is calculated using Equation 31. In this equation, $|N|$ expresses the number of non-dominated solutions, d_i is the Euclidean distance between consecutive solutions, and \bar{d} is the average of d_i . For a solution algorithm, the lower the value of the D_1 metric, the more efficient that algorithm is (21, 51).

$$D_1 = \frac{\sum_{i=1}^{|N|-1} |d_i - \bar{d}|}{|N|-1} \tag{31}$$

Spacing (D_2): This measure is an extension of the previous metric and is obtained from Equation 32. In this equation, $|N|$ shows the number of non-dominated solutions. The values of d_i and \bar{d} are obtained via Equations 33 and 34, respectively. For a solution algorithm, the lower the value of the D_2 metric, the better the performance of that algorithm (22, 52).

$$D_2 = \left(\frac{1}{|N|} \sum_{i=1}^{|N|} (d_i - \bar{d})^2 \right)^{1/2} \tag{32}$$

$$d_i = \min_{k \in N, k \neq i} \sum_{m=1}^M |f_m^i - f_m^k| \tag{33}$$

$$\bar{d} = \frac{\sum_{i=1}^{|N|-1} d_i}{|N|-1} \tag{34}$$

Mean ideal distance (MID): This performance measure calculates the average Euclidean distance of the ideal solution from the Pareto front obtained by the solution algorithm. In this paper, the best possible value for each of the objective functions obtained by different algorithms is considered the ideal solution. Equation 35 is used to calculate MID . In this equation, $|N|$ is the number of non-dominated solutions, and C_i represents the Euclidean distance of each member of the Pareto front from the ideal point, which is calculated through Equation 36. An algorithm with a lower MID metric has better performance (51).

$$MID = \frac{1}{|N|} \sum_{i=1}^{|N|} C_i \tag{35}$$

$$C_i = \sqrt{(f_{1i} - f_1^*)^2 + \dots + (f_{mi} - f_m^*)^2} \tag{36}$$

5. 3. Parameter Tuning Metaheuristic algorithms possess inherent parameters, and assigning appropriate values to these parameters can substantially enhance the quality of the obtained results. This paper employs the Taguchi method to optimize the parameters of the developed algorithms. The Taguchi method utilizes orthogonal arrays, which are standardized arrays that enable the execution of a limited number of experiments while retaining comprehensive information on all factors influencing the performance of the algorithms (53). The HNSGAI-PSO factors are: number of population (N), number of generations (G), crossover rate (P_c), mutation

rate of the gene (P_{mg}), number of population for the PSO algorithm (N_{PSO}), number of generations for the PSO algorithm (G_{PSO}), inertia coefficient (ω), and learning coefficients (c_1 , c_2) and The NSGA-II factors are: number of population (N), number of generations (G), crossover rate (P_c), mutation rate (P_m) and mutation rate of the gene (P_{mg}). Tables 3 and 4 show the considered levels for the factors of the HNSGAI-PSO and NSGA-II algorithms for small-scale problems and medium- and large-scale problems, respectively.

TABLE 3. Algorithm parameters and their levels for small-scale problems

Algorithm	Parameters	Symbol	Level
HNSGAI-PSO	N	A	100 – 150 – 200
	G	B	50 – 80 – 100
	P_c	C	0.6 – 0.7 – 0.8
	P_{mg}	D	0.2 – 0.3 – 0.4
	N_{PSO}	E	2 – 3 – 4
	G_{PSO}	F	1 – 2 – 3
	ω	G	0.5 – 0.75 – 1
	c_1	H	1 – 1.5 – 2
	c_2	J	1 – 1.5 – 2
	NSGA-II	N	A
G		B	50 – 80 – 100
P_c		C	0.5 – 0.6 – 0.7
P_m		D	0.5 – 0.6 – 0.7
P_{mg}		E	0.5 – 0.6 – 0.7

TABLE 4. Algorithm parameters and their levels for medium- and large-scale problems

Algorithm	Parameters	Symbol	Level
HNSGAI-PSO	N	A	300 – 350 – 400
	G	B	100 – 150 – 200
	P_c	C	0.6 – 0.7 – 0.8
	P_{mg}	D	0.2 – 0.3 – 0.4
	N_{PSO}	E	2 – 3 – 4
	G_{PSO}	F	1 – 2 – 3
	ω	G	0.5 – 0.75 – 1
	c_1	H	1 – 1.5 – 2
	c_2	J	1 – 1.5 – 2
	NSGA-II	N	A
G		B	300 – 400 – 500
P_c		C	0.5 – 0.6 – 0.7
P_m		D	0.5 – 0.6 – 0.7
P_{mg}		E	0.5 – 0.6 – 0.7

Considering the number of parameters and determined levels, the right orthogonal array for both algorithms is L_{27} , which includes 27 experiments. Then, for each algorithm, all experiments are conducted on an instance problem considering various combinations of parameter levels, and the obtained results are recorded for performance measures. It's important to note that each experiment was repeated five times, and the average of the results was considered. Since the Taguchi method only accepts one value as a response for each experiment, the results obtained for the performance measures are unscaled using the relative percentage deviation (RPD) method. then for each experiment, the weighted average of the relevant unscaled performance measures is calculated as a combined function (CF) using equation 37 and considered as the result of that experiment (51).

$$CF = \frac{NPS + D_1 + D_2 + 2MID}{5} \quad (37)$$

Figures 13, 14, 15, and 16 show the results obtained from the Taguchi method for small-scale problems and medium- and large-scale problems, respectively. Based on these results, the parameter values of each of the NSGA-II and HNSGAI-PSO algorithms are presented in Tables 5 and 6, respectively.

5. 4. Evaluation of Solution Algorithms for Small-Scale Instances

In this sub-section, the performance of the AUGMECON method, the NSGA-II algorithm, and the HNSGAI-PSO algorithm are compared for solving small-scale instances. The results of solving 30 instances using the three methods mentioned are shown in Table 7. In this table, column 1 shows the instance number, and column 2 shows the instance size. The results of solving the instances with the AUGMECON method are presented for the position-based model and the sequence-based model in columns 3 and 4, respectively. The AUGMECON method for both presented models has the ability to obtain the optimal Pareto front for 26 instances; however, for 4 instances, this method is not able to solve the sequence-based model

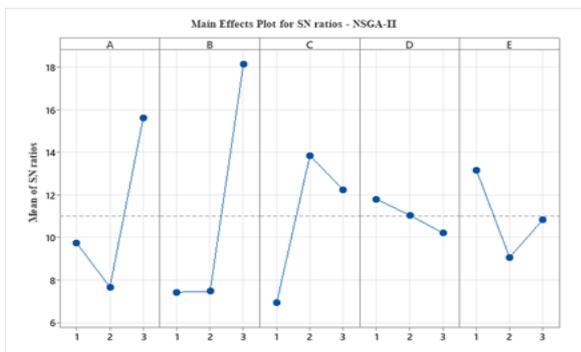


Figure 13. Main effects plot for S/N ratios for the NSGA-II algorithms: small-scale problems

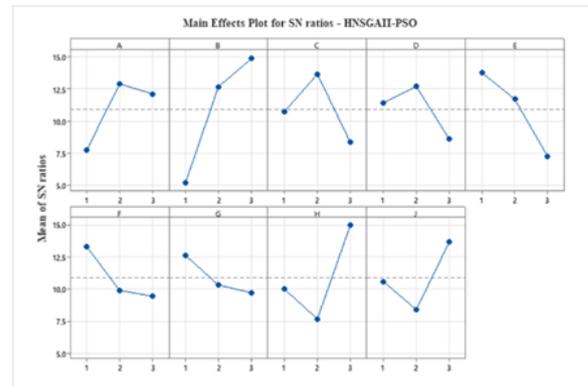


Figure 14. Main effects plot for S/N ratios for the HNSGAI-PSO algorithms: small-scale problems

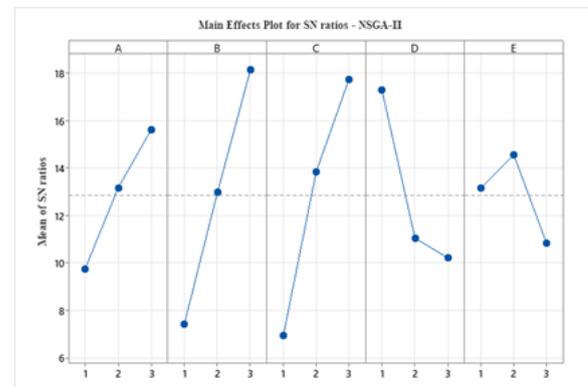


Figure 15. Main effects plot for S/N ratios for the NSGA-II algorithms: medium- and large-scale problems

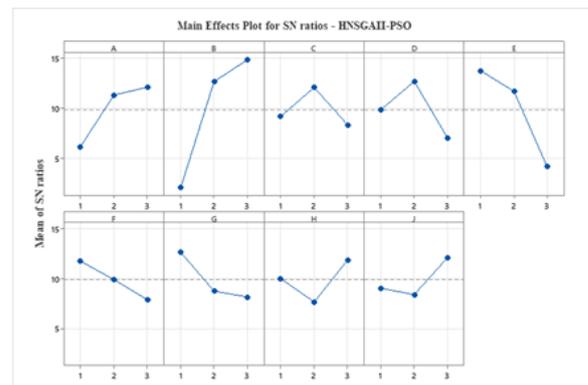


Figure 16. Main effects plot for S/N ratios for the HNSGAI-PSO algorithms: medium- and large-scale problems

TABLE 5. Parameter tuning results for the NSGA-II algorithm

	Parameters				
	N	G	P_c	P_m	P_{mg}
Small-scale	200	100	0.6	0.5	0.5
Medium and large-scale	800	500	0.7	0.5	0.6

TABLE 6. Parameter tuning results for the HNSGAI-PSO algorithm

	Parameters								
	N	G	P_c	P_{mg}	N_{PSO}	G_{PSO}	ω	c_1	c_2
Small-scale	150	100	0.7	0.3	2	1	0.5	2	2
Medium and large-scale	400	200	0.7	0.3	2	1	0.5	2	2

TABLE 7. Results of solving small-scale instances using the AUGMECON method, HNSGAI-PSO algorithm, and NSGA-II algorithm

No.	$[J, J, M, \delta]$	AUGMECON				HNSGAI-PSO					NSGA-II				
		NPS	MID	$t(s)$		NPS	D_1	D_2	MID	$t(s)$	NPS	D_1	D_2	MID	$t(s)$
				position-based model	sequence-based model										
1	2-4-2-50%	21	572.48	4.08	6.68	21	43.88	17.59	572.48	0.73	21	43.88	17.59	572.48	0.73
2	2-4-2-75%	24	947.25	5.54	6.78	24	65.61	60.42	947.25	0.87	24	65.61	60.42	947.25	0.87
3	2-5-2-50%	36	1581.3	62.36	183.26	36	69.01	48.09	1581.3	0.93	36	69.01	48.09	1581.3	0.93
4	2-5-2-60%	45	1635	101.35	290.27	45	53.94	52.92	1635	1.23	45	53.94	52.92	1635	1.23
5	3-4-2-40%	21	822.58	20.22	23.77	21	47.88	61.18	822.58	1.59	21	47.88	61.18	822.58	1.59
6	3-4-2-60%	25	786.76	18.65	21.55	25	38.19	44.54	786.76	1.6	25	38.19	44.54	786.76	1.6
7	3-4-3-50%	30	970.58	24.29	18.38	30	43.41	51.27	970.58	1.68	30	43.41	51.27	970.58	1.68
8	3-4-3-75%	23	959.61	13.58	11.54	23	47.31	58.02	959.61	1.69	23	47.31	58.02	959.61	1.69
9	3-5-2-40%	38	722.14	252.99	437.14	38	25.16	22.23	722.14	1.72	38	25.16	22.23	722.14	1.72
10	3-5-2-60%	38	614.85	93.64	179.86	38	35.73	41.67	614.85	1.76	38	35.73	41.67	614.85	1.76
11	3-5-3-40%	35	688.7	190.09	286.29	35	35.29	23.92	688.7	1.76	35	35.29	23.92	688.7	1.76
12	3-5-3-60%	40	732.17	146.49	211.24	40	26.01	19.71	732.17	1.78	40	26.01	19.71	732.17	1.78
13	4-4-2-25%	30	540.82	26.53	24.68	30	41.38	59.46	540.82	1.61	30	41.38	59.46	540.82	1.61
14	4-4-2-50%	39	756.27	37.95	39.74	39	22.89	24.71	756.27	1.65	39	22.89	24.71	756.27	1.65
15	4-4-2-75%	31	665.67	33.60	32.19	31	31.70	35.71	665.71	1.67	31	31.70	35.71	665.71	1.67
16	4-4-3-30%	34	915.67	25.56	28.75	34	26.58	30.79	915.67	1.73	34	26.58	30.79	915.67	1.73
17	4-4-3-50%	29	947.81	55.81	61.64	29	38.59	22.35	947.81	1.64	29	38.59	22.35	947.81	1.64
18	4-5-2-30%	48	868.82	118.48	124.83	48	37.85	31.46	868.82	1.78	48	37.85	31.46	868.82	1.78
19	4-5-2-40%	67	1011.1	468.33	789.60	64	18.97	16.04	1046	1.76	64	18.97	16.04	1046	1.76
20	4-5-3-30%	53	719.08	1743.38	1812.21	53	17.08	10.82	719.08	1.77	53	17.08	10.82	719.08	1.77
21	4-5-3-40%	66	1024.8	1927.79	1967.18	66	20.17	19.01	1024.8	1.78	66	20.17	19.01	1024.8	1.78
22	4-6-2-25%	55	614.16	5293.29	11035.73	55	12.04	14.58	614.16	1.79	55	12.04	14.58	614.16	1.79
23	4-6-2-50%	58	734.96	2083.45	29303.34	63	13.54	13.98	766.23	1.79	63	13.54	13.98	766.23	1.79
24	4-6-2-75%	65	698.25	2359.10	33179.49	65	12.17	14.71	699.64	1.81	65	12.17	14.71	699.64	1.81
25	4-6-3-30%	31	891.77	8876.68	11499.63	30	46.11	56.1	893.38	1.77	30	46.11	56.1	893.38	1.77
26	4-6-3-40%	53	763.95	13521.75	28961.49	57	35.82	29.54	792.65	1.78	57	35.82	29.54	792.65	1.78
27	4-6-3-50%	44	911.54	19492.31	-	42	43.13	45.92	923.75	1.8	42	43.13	45.92	923.75	1.8
28	5-7-2-20%	67	1174.23	12548.52	-	73	39.46	42.96	1227.62	1.78	73	39.46	42.96	1227.62	1.78
29	5-7-2-30%	61	938.47	21649.39	-	59	41.44	39.12	932.54	1.79	59	41.44	39.12	932.54	1.79
30	5-7-2-40%	71	1480.2	34394.65	-	64	33.01	27.84	1518.6	1.82	64	33.01	27.84	1518.6	1.82

The running time of the algorithm that obtained the optimal Pareto front is shown with a bold value

and find the optimal Pareto front in 36000 seconds. Based on the presented results, the position-based model has better performance in terms of optimal solution time using the AUGMECON method compared to the sequence-based model. To validate the approximate Pareto solutions obtained by the NSGA-II and HNSGAI-PSO algorithms, a comparison with the AUGMECON method was conducted. The proposed NSGA-II and HNSGAI-PSO algorithms were employed to solve small-scale instances. The results of solving small-scale instances using the HNSGAI-PSO and NSGA-II algorithms are presented in columns 5 and 6 of Table 7, respectively. An examination of Table 7 revealed that the NSGA-II and HNSGAI-PSO algorithms successfully identified the optimal Pareto front in 19 instances, demonstrating superior computational efficiency compared to the AUGMECON method. In the remaining instances, the solutions obtained by both algorithms exhibited close proximity to the optimal Pareto solutions identified by the AUGMECON method. Consequently, based on these findings, it is evident that the NSGA-II and HNSGAI-PSO algorithms constitute valid and effective approaches for solving instances in a reasonable time.

To comprehensively evaluate Pareto dominance and compare the solutions obtained by the AUGMECON method and the proposed algorithms, the MID metric was calculated for all three methods, and the results are presented in Figure 17. The analysis of small-scale instances revealed that the solutions generated by the proposed algorithms are comparable to the optimal Pareto fronts when compared to the AUGMECON method. Notably, the NSGA-II and HNSGAI-PSO algorithms provide acceptable results in significantly less time than the AUGMECON method. Figure 18 demonstrates the CPU time for solving small-scale instances by all three methods. As observed, with increasing instance sizes, the CPU time for the AUGMECON method escalates dramatically. Therefore, based on the comprehensive comparison of Pareto solutions, it can be concluded that NSGA-II and HNSGAI-PSO are viable and efficient algorithms for solving small-scale instances within a reasonable computational time frame.

5. 5. Evaluation of Solution Algorithms for Medium- and Large -Scale Instances

Based on the results presented in Table 7, by increasing the scale of the problem, the computation time increases. For example, in instance #30 with size [5,7,2,40%], the AUGMECON method needs 34395 seconds to obtain the optimal Pareto front. Therefore, HNSGAI-PSO and NSGA-II algorithms are used to solve medium- and large-scale instances, and their obtained results are compared. To ensure a fair comparison, we have set the

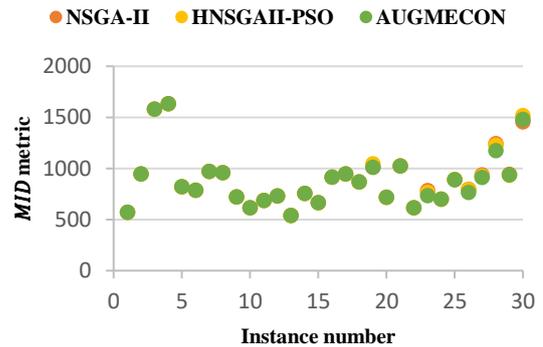


Figure 17. MID metric obtained through the NSGA-II, HNSGAI-PSO and AUGMECON methods for small-scale instances

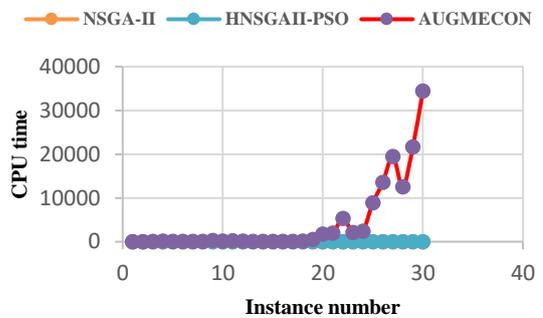


Figure 18. CPU time for solving small-scale instances by the NSGA-II, HNSGAI-PSO and AUGMECON methods

running times of both algorithms to be equal. This means that the time considered when comparing two algorithms is equal to the execution time of the algorithm that finishes faster in each instance. Each of the two algorithms is executed five times for each instance, and the average values derived for the performance measures are recorded. The outcomes of applying the NSGA-II and HNSGAI-PSO algorithms to solve 30 medium- and large-scale instances are presented in Table 8. Figure 19 shows the Pareto fronts generated by both algorithms for problem instance #55. Based on the obtained results, the HNSGAI-PSO algorithm performs better in the NPS, D_1 , and D_2 criteria, which means that this algorithm can generate a greater number of non-dominated solutions with more diversity. The evaluation of the MID criterion demonstrates that the HNSGAI-PSO algorithm performs better, indicating that it can produce non-dominated solutions with better convergence.

In order to further analyze the performance of the HNSGAI-PSO and NSGA-II algorithms in solving instance problems, paired samples t-test is performed for every performance measure. Statistical results for different measures are stated in Table 9. The mean and

TABLE 8. Results of solving medium- and large-scale instances using the HNSGAI-PSO algorithm and the NSGA-II algorithm

No.	$[I, J, M, \delta]$	Time (s)	NPS		D_2		D_1		MID	
			HNSGAI-PSO	NSGA-II	HNSGAI-PSO	NSGA-II	HNSGAI-PSO	NSGA-II	HNSGAI-PSO	NSGA-II
31	15-30-5-20%	25	59.60	47.03	59.60	47.03	59.60	47.03	4878.05	4801.21
32	30-30-5-30%	33	48.81	32.22	48.81	32.22	48.81	32.22	5065.85	5046.76
33	50-30-5-40%	42	22.38	42.54	22.38	42.54	22.38	42.54	7909.23	7894.54
34	30-50-7-20%	47	84.16	63.73	84.16	63.73	84.16	63.73	8521.25	8492.57
35	50-50-7-30%	56	28.01	52.37	28.01	52.37	28.01	52.37	9757.75	9737.86
36	80-50-7-40%	78	31.19	58.38	31.19	58.38	31.19	58.38	11026.79	11284.65
37	50-80-10-20%	81	38.53	77.22	38.53	77.22	38.53	77.22	12357.81	13701.54
38	80-80-10-30%	111	33.78	63.41	33.78	63.41	33.78	63.41	12549.56	12783.00
39	100-80-10-40%	128	40.47	73.42	40.47	73.42	40.47	73.42	13850.82	15555.38
40	80-100-12-20%	131	39.75	95.05	39.75	95.05	39.75	95.05	15473.07	16454.84
41	100-100-12-30%	152	41.16	71.94	41.16	71.94	41.16	71.94	13842.15	14764.67
42	150-100-12-40%	204	39.80	90.48	39.80	90.48	39.80	90.48	15940.74	16934.13
43	100-150-15-20%	223	54.25	103.11	54.25	103.11	54.25	103.11	17261.59	19042.25
44	150-150-15-30%	278	53.49	118.75	53.49	118.75	53.49	118.75	19587.96	21591.61
45	200-150-15-40%	336	58.67	111.19	58.67	111.19	58.67	111.19	20524.40	21737.06
46	150-200-18-20%	347	63.72	123.72	63.72	123.72	63.72	123.72	19943.05	23042.00
47	200-200-18-30%	419	66.89	153.43	66.89	153.43	66.89	153.43	23681.41	28178.78
48	250-200-18-40%	494	65.76	158.35	65.76	158.35	65.76	158.35	24610.11	28376.25
49	200-250-20-20%	504	86.13	145.93	86.13	145.93	86.13	145.93	24744.62	28052.78
50	250-250-20-30%	599	79.92	166.71	79.92	166.71	79.92	166.71	24539.97	29819.36
51	300-250-20-40%	698	78.02	159.17	78.02	159.17	78.02	159.17	25810.26	30246.06
52	250-300-25-20%	718	70.41	148.80	70.41	148.80	70.41	148.80	21604.30	28446.83
53	300-300-25-30%	825	95.05	177.20	95.05	177.20	95.05	177.20	26398.03	32207.54
54	400-300-25-40%	1050	107.09	212.11	107.09	212.11	107.09	212.11	29752.38	35398.91
55	300-400-30-20%	1100	98.86	195.93	98.86	195.93	98.86	195.93	27888.72	31671.08
56	400-400-30-30%	1290	124.71	220.91	124.71	220.91	124.71	220.91	32467.22	37604.07
57	500-400-30-40%	1690	119.76	232.55	119.76	232.55	119.76	232.55	33281.22	41192.72
58	400-500-35-20%	1760	136.67	198.92	136.67	198.92	136.67	198.92	33257.67	40642.41
59	450-500-35-30%	1984	138.73	245.32	138.73	245.32	138.73	245.32	35466.96	41484.99
60	500-500-35-40%	2145	139.89	321.04	139.89	321.04	139.89	321.04	38320.90	47772.89
Average value			350.53	302.67	90.59	156.25	71.52	132.03	20343.80	23465.29

The algorithm with better performance in each measure is shown with a bold value

SD columns show the mean value and standard deviation of the performance measures, respectively. Based on the results presented in Table 9, the p-value for all measures is smaller than 0.05. Considering the common significance level of 0.05, the results of the paired samples t-test showed that for all metrics, there is a significant difference between the performance of the

HNSGAI-PSO algorithm and that of the NSGA-II algorithm.

To meticulously evaluate the performance of the algorithms, comprehensive statistical analyses were conducted for each performance measure. The results obtained for the performance metrics in Table 8 were transformed using the relative percentage deviation

TABLE 9. Paired sample t-test results of the HNSGAI-PSO and NSGA-II algorithms

Metric	HNSGAI-PSO		NSGA-II		p-value	Significance
	Mean	SD	Mean	SD		
<i>NPS</i>	350.53	41.24	302.67	32.46	0.0021	Yes
<i>D</i> ₁	90.59	46.94	156.25	89.16	0.0003	Yes
<i>D</i> ₂	71.52	35.02	132.03	71.47	0.0002	Yes
<i>MID</i>	20343.8	9360.77	23465.29	11957.24	0.0008	Yes

(RPD) method to facilitate an unbiased comparison. Interval plots with a 95% confidence level were constructed for each performance measure to assess the algorithm's accuracy and robustness (53). A smaller interval indicates superior accuracy, while lower interval values compared to other algorithms indicate greater robustness. These findings are illustrated in Figures 20 to 23.

An analysis of the NPS metric (Figure 20) revealed that the HNSGAI-PSO algorithm exhibited superior robustness compared to the NSGA-II algorithm, despite

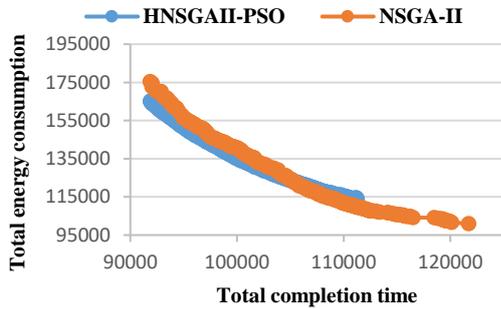


Figure 19. Pareto fronts obtained by various algorithms for problem instance #55

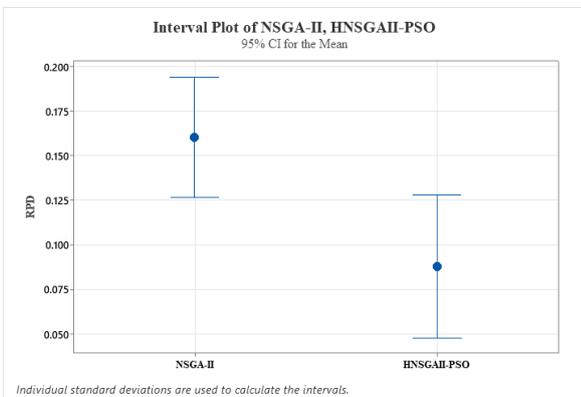


Figure 20. Interval plot based on a 95% confidence level for the comparison of the NSGA-II and HNSGAI-PSO algorithms using the NPS metric

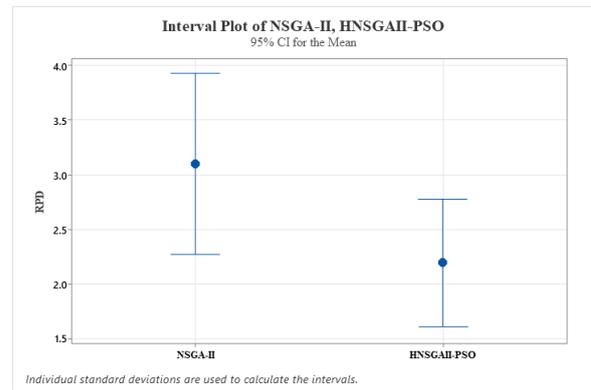


Figure 21. Interval plot based on a 95% confidence level for the comparison of the NSGA-II and HNSGAI-PSO algorithms using the *D*₁ metric

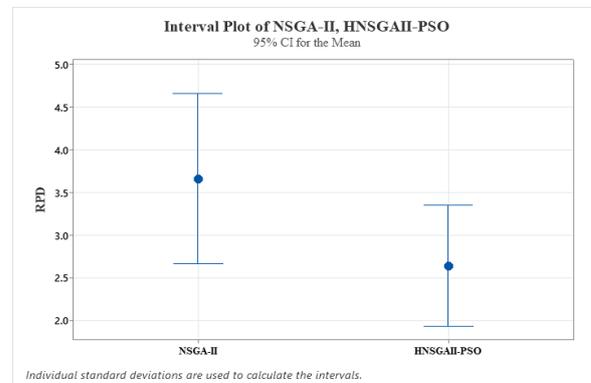


Figure 22. Interval plot based on a 95% confidence level for the comparison of the NSGA-II and HNSGAI-PSO algorithms using the *D*₂ metric

their equivalent accuracy. Regarding the *D*₁ indicator (Figure 21), the HNSGAI-PSO algorithm demonstrated exceptional performance, achieving both the highest level of robustness and accuracy. For both the *D*₂ and MID metrics, Figures 22 and 23, respectively, illustrate the HNSGAI-PSO algorithm's superior accuracy and robustness compared to the NSGA-II algorithm.

5. 6. Sensitivity Analysis

In this subsection, the effect of varying the parameters p_{js} , e_{js} , and m on the

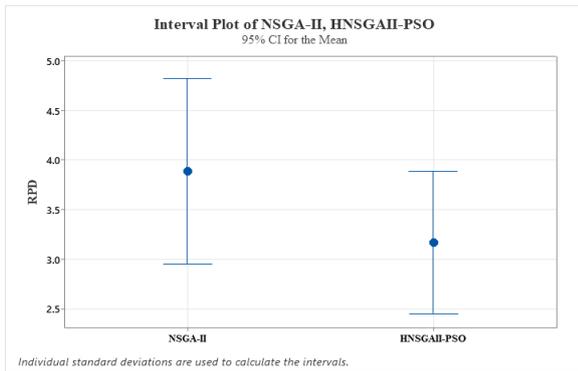


Figure 23. Interval plot based on a 95% confidence level for the comparison of the NSGA-II and HNSGAI-PSO algorithms using the MID metric

objective functions is analyzed. Problem instance #7 is considered for the sensitivity analysis. Figures 24, 25, and 26 depict the Pareto fronts obtained by solving novel problems based on varying the parameters p_{js} , e_{js} , and m , respectively. Figure 24 shows that by increasing the p_{js} parameter, the processing time of each pattern increased, and when the job-pattern matrix is considered, the objective function of the total completion time also increased. This could potentially lead to customer dissatisfaction. However, the utilization of advanced machines and trained personnel can lower the p_{js} parameter. To avoid an increase in the p_{js} parameter, decision-makers must allocate sufficient funds for the purchase and installation of new machines and provide regular training programs for employees to improve their skills. Figure 25 illustrates how varying the e_{js} parameter influences the total energy consumption objective function. The Pareto front shifts towards higher values of this objective function as the e_{js} parameter increases. Utilizing high-tech machinery and performing timely maintenance and repairs can be effective in preventing the increase in the e_{js} parameter and its associated expenses. Therefore, decision-makers should conduct the necessary comparisons between various alternatives to implement the optimal strategy for maximizing overall profit. Figure 26 demonstrates that decreasing the m parameter increases both the total completion time and total energy consumption. In fact, as the number of machines is reduced, more patterns are assigned to the remaining machines, resulting in an increase in the time required to complete all jobs. To tackle this issue and reduce customer dissatisfaction, it is necessary to select the machine's fast levels for pattern processing in order to reduce their completion time, resulting in an increase in energy consumption. This analysis enables decision-makers to strike a balance between the required budget for purchasing and deploying new machines and the costs resulting from an increase in total completion time and total energy consumption.

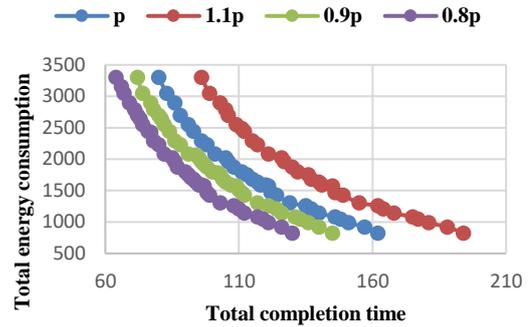


Figure 24. Sensitivity of the Pareto frontier to processing time

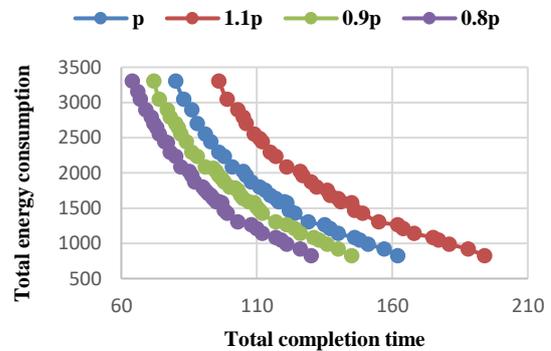


Figure 25. Sensitivity of the Pareto frontier to energy consumption

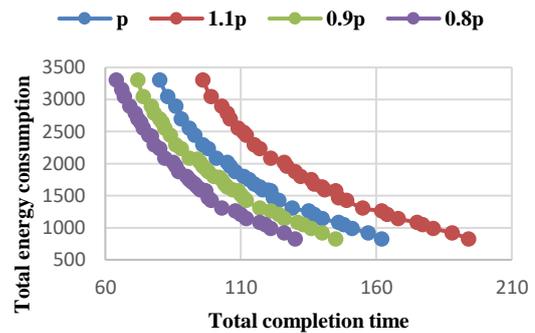


Figure 26. Sensitivity of the Pareto frontier to the number of machines

6. CONCLUSION AND FUTURE RESEARCH

In the realm of manufacturing, the cutting stock problem, frequently encountered in industries such as furniture and apparel, exemplifies the application of common operation scheduling. Conversely, effective management and energy consumption reduction have emerged as

pressing concerns within the industry. Addressing these issues is paramount, considering the manufacturing sector's substantial energy consumption. In this paper, the common operation scheduling in an environment of identical parallel machines was studied, considering the energy consumption. In the investigated problem, each job includes several pieces, and all the pieces of different jobs are placed on cutting patterns. Each cutting pattern can contribute to the completion of one or more jobs. Each job is completed when all pieces of that job have been produced by processing related patterns. Each machine in this problem possesses varying speed levels. Consequently, when the machine operates at a higher speed, the processing time is reduced while the consumption of electrical energy increases.

In the investigated problem, to simultaneously minimize the total completion time and the total electrical energy consumption, two position-based and sequence-based mixed integer linear programming models were presented, and to solve small-scale instances, the AUGMECON method was used to obtain the Pareto optimal front. To solve medium- and large-scale instances, the HNSGAI-PSO and NSGA-II algorithms were developed to achieve good approximate Pareto fronts. In the NSGA-II algorithm, each chromosome represents a solution from the problem-solving space, and the quality of each chromosome can be evaluated by decoding it. In the HNSGAI-PSO algorithm, each chromosome represents a region of the problem-solving space where all solutions in this region have the same total energy consumption. To assess the quality of each chromosome, the top solution within the chromosome's covered region is taken into consideration. This entails running the PSO algorithm once for every chromosome to determine the best solution in each region. The performance of the presented algorithms was evaluated by solving test instances of different sizes. The results of numerical experiments show that both presented algorithms perform well in solving small-scale instances and can obtain the optimal Pareto front in much less time than the AUGMECON method. Based on the results of solving medium- and large-scale instances, the HNSGAI-PSO algorithm has better performance compared to the NSGA-II algorithm and can obtain more diverse non-dominant solutions with better convergence. At last, the problem's sensitivity to the parameters of processing time, energy consumption, and the number of machines was analyzed, and the impact of varying each of these parameters on the two objective functions was demonstrated. The results indicate that decision-makers should compare the budgets, costs, and revenues associated with different changes to make well-informed decisions that maximize overall profit.

Considering sequence-dependent setup times for processing cutting patterns, studying the problem in an environment of unrelated parallel machines, and using

the TOU tariffs or the tiered price to calculate the cost of power consumption are all attractive fields for future research.

7. REFERENCES

1. Babae Tirkolaee E, Goli A, Mardani A. A novel two-echelon hierarchical location-allocation-routing optimization for green energy-efficient logistics systems. *Annals of Operations Research*. 2021. 10.1007/s10479-021-04363-y
2. Zhang X, Zhou H, Fu C, Mi M, Zhan C, Pham DT, et al. Application and planning of an energy-oriented stochastic disassembly line balancing problem. *Environmental Science and Pollution Research*. 2023;1-15. 10.1007/s11356-023-27288-4
3. Zandvakili A, Mansouri N, Javidi M. Energy-aware task scheduling in cloud computing based on discrete pathfinder algorithm. *International Journal of Engineering, Transactions C: Aspects*. 2021;34(9):2124-36. 10.5829/IJE.2021.34.09C.10
4. Tian G, Zhang C, Fathollahi-Fard AM, Li Z, Zhang C, Jiang Z. An enhanced social engineering optimizer for solving an energy-efficient disassembly line balancing problem based on bucket brigades and cloud theory. *IEEE Transactions on Industrial Informatics*. 2022. 10.1109/TII.2022.3193866
5. Fang K, Uhan N, Zhao F, Sutherland JW. A new approach to scheduling in manufacturing for power consumption and carbon footprint reduction. *Journal of Manufacturing Systems*. 2011;30(4):234-40. 10.1016/j.jmsy.2011.08.004
6. Lu C, Gao L, Li X, Pan Q, Wang Q. Energy-efficient permutation flow shop scheduling problem using a hybrid multi-objective backtracking search algorithm. *Journal of cleaner production*. 2017;144:228-38. 10.1016/j.jclepro.2017.01.011
7. Jovane F, Yoshikawa H, Altling L, Boer CR, Westkamper E, Williams D, et al. The incoming global technological and industrial revolution towards competitive sustainable manufacturing. *CIRP annals*. 2008;57(2):641-59. 10.1016/j.cirp.2008.09.010
8. Mori M, Fujishima M, Inamasu Y, Oda Y. A study on energy efficiency improvement for machine tools. *CIRP annals*. 2011;60(1):145-8. 10.1016/j.cirp.2011.03.099
9. Pinedo ML. *Scheduling*: Springer; 2012.
10. Ding J, Schulz S, Shen L, Buscher U, Lü Z. Energy aware scheduling in flexible flow shops with hybrid particle swarm optimization. *Computers & Operations Research*. 2021;125:105088. 10.1016/j.cor.2020.105088
11. Zhang M, Yan J, Zhang Y, Yan S. Optimization for energy-efficient flexible flow shop scheduling under time of use electricity tariffs. *Procedia CIRP*. 2019;80:251-6. 10.1016/j.procir.2019.01.062
12. Guo J, Lei D, Li M. Two-phase imperialist competitive algorithm for energy-efficient flexible job shop scheduling. *Journal of Intelligent & Fuzzy Systems*. 2021;40(6):12125-37. 10.3233/JIFS-210198
13. Fathollahi-Fard AM, Woodward L, Akhrif O. Sustainable distributed permutation flow-shop scheduling model based on a triple bottom line concept. *Journal of Industrial Information Integration*. 2021;24:100233. 10.1016/j.jii.2021.100233
14. Li Z, Yang H, Zhang S, Liu G. Unrelated parallel machine scheduling problem with energy and tardiness cost. *The International Journal of Advanced Manufacturing Technology*. 2016;84:213-26. 10.1007/s00170-015-7657-2
15. Che A, Zhang S, Wu X. Energy-conscious unrelated parallel machine scheduling under time-of-use electricity tariffs. *Journal*

- of cleaner production. 2017;156:688-97. 10.1016/j.jclepro.2017.04.018
16. Wang Y-C, Wang M-J, Lin S-C. Selection of cutting conditions for power constrained parallel machine scheduling. *Robotics and Computer-Integrated Manufacturing*. 2017;43:105-10. 10.1016/j.rcim.2015.10.010
 17. Zeng Y, Che A, Wu X. Bi-objective scheduling on uniform parallel machines considering electricity cost. *Engineering Optimization*. 2018;50(1):19-36. 10.1080/0305215X.2017.1296437
 18. Wu X, Che A. A memetic differential evolution algorithm for energy-efficient parallel machine scheduling. *Omega*. 2019;82:155-65. 10.1016/j.omega.2018.01.001
 19. Cota LP, Coelho VN, Guimarães FG, Souza MJ. Bi-criteria formulation for green scheduling with unrelated parallel machines with sequence-dependent setup times. *International Transactions in Operational Research*. 2021;28(2):996-1017. 10.1111/itor.12566
 20. Safarzadeh H, Niaki STA. Bi-objective green scheduling in uniform parallel machine environments. *Journal of cleaner production*. 2019;217:559-72. 10.1016/j.jclepro.2019.01.166
 21. Wang S, Wang X, Yu J, Ma S, Liu M. Bi-objective identical parallel machine scheduling to minimize total energy consumption and makespan. *Journal of cleaner production*. 2018;193:424-40. 10.1016/j.jclepro.2018.05.056
 22. Anghinolfi D, Paolucci M, Ronco R. A bi-objective heuristic approach for green identical parallel machine scheduling. *European Journal of Operational Research*. 2021;289(2):416-34. 10.1016/j.ejor.2020.07.020
 23. Zhang H, Wu Y, Pan R, Xu G. Two-stage parallel speed-scaling machine scheduling under time-of-use tariffs. *Journal of Intelligent Manufacturing*. 2021;32:91-112. 10.1007/s10845-020-01561-6
 24. Keshavarz T, Karimi E, Shakhshi-Niaei M. Unrelated parallel machines scheduling with sequence-dependent setup times to minimize makespan and tariff charged energy consumption. *Advances in Industrial Engineering*. 2021;55(1):91-113. 10.22059/aieng.2021.326682.1788
 25. Zhou B-H, Gu J. Energy-awareness scheduling of unrelated parallel machine scheduling problems with multiple resource constraints. *International Journal of Operational Research*. 2021;41(2):196-217. 10.1504/IJOR.2021.115623
 26. Módos I, Šucha P, Hanzálek Z. On parallel dedicated machines scheduling under energy consumption limit. *Computers & Industrial Engineering*. 2021;159:107209. 10.1016/j.cie.2021.107209
 27. Rego MF, Pinto JCE, Cota LP, Souza MJ. A mathematical formulation and an NSGA-II algorithm for minimizing the makespan and energy cost under time-of-use electricity price in an unrelated parallel machine scheduling. *PeerJ Computer Science*. 2022;8:e844. 10.7717/peerj-cs.844
 28. Asadpour M, Hodaei Z, Azami M, Kehtari E, Vesal N. A green model for identical parallel machines scheduling problem considering tardy jobs and job splitting property. *Sustainable Operations and Computers*. 2022;3:149-55.
 29. Gaggero M, Paolucci M, Ronco R. Exact and Heuristic Solution Approaches for Energy-Efficient Identical Parallel Machine Scheduling with Time-of-Use Costs. *European Journal of Operational Research*. 2023.
 30. Arbib C, Servilio M, Felici G, Servilio M. Sorting common operations to minimize the number of tardy jobs. *Networks*. 2014;64(4):306-20. 10.1002/net.21576
 31. Cheng T, Diamond J, Lin BM. Optimal scheduling in film production to minimize talent hold cost. *Journal of Optimization Theory and Applications*. 1993;79(3):479-92. 10.1007/BF00940554
 32. Wang J, Qiao C, Yu H, editors. On progressive network recovery after a major disruption. 2011 Proceedings IEEE INFOCOM; 2011: IEEE. 10.1109/INFCOM.2011.5934996
 33. Arbib C, Felici G, Servilio M. Common operation scheduling with general processing times: A branch-and-cut algorithm to minimize the weighted number of tardy jobs. *Omega*. 2019;84:18-30. 10.1016/j.omega.2018.04.002
 34. Cherri AC, Arenales MN, Yanasse HH, Poldi KC, Vianna ACG. The one-dimensional cutting stock problem with usable leftovers—A survey. *European Journal of Operational Research*. 2014;236(2):395-402. 10.1016/j.ejor.2013.11.026
 35. Hinxman A. The trim-loss and assortment problems: A survey. *European Journal of Operational Research*. 1980;5(1):8-18. 10.1016/0377-2217(80)90068-5
 36. Dyckhoff H. A typology of cutting and packing problems. *European journal of operational research*. 1990;44(2):145-59. 10.1016/0377-2217(90)90350-K
 37. Wäscher G, Haußner H, Schumann H. An improved typology of cutting and packing problems. *European journal of operational research*. 2007;183(3):1109-30. 10.1016/j.ejor.2005.12.047
 38. Arbib C, Marinelli F. On cutting stock with due dates. *Omega*. 2014;46:11-20. 10.1016/j.omega.2014.01.004
 39. Cui Y, Zhong C, Yao Y. Pattern-set generation algorithm for the one-dimensional cutting stock problem with setup cost. *European Journal of Operational Research*. 2015;243(2):540-6. 10.1016/j.ejor.2014.12.015
 40. Wuttke DA, Heese HS. Two-dimensional cutting stock problem with sequence dependent setup times. *European Journal of Operational Research*. 2018;265(1):303-15. 10.1016/j.ejor.2017.07.036
 41. Graham RL, Lawler EL, Lenstra JK, Kan AR. Optimization and approximation in deterministic sequencing and scheduling: a survey. *Annals of discrete mathematics*. 5: Elsevier; 1979. p. 287-326.
 42. Garey MR, Johnson DS. "strong"np-completeness results: Motivation, examples, and implications. *Journal of the ACM (JACM)*. 1978;25(3):499-508.
 43. Deb K, Agrawal S, Pratap A, Meyarivan T, editors. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. *International conference on parallel problem solving from nature*; 2000: Springer.
 44. Duan J, Wang J. Energy-efficient scheduling for a flexible job shop with machine breakdowns considering machine idle time arrangement and machine speed level selection. *Computers & Industrial Engineering*. 2021;161:107677.
 45. Xue L, Wang X. A multi-objective discrete differential evolution algorithm for energy-efficient two-stage flow shop scheduling under time-of-use electricity tariffs. *Applied Soft Computing*. 2023;133:109946.
 46. Eberhart R, Kennedy J, editors. Particle swarm optimization. *Proceedings of the IEEE international conference on neural networks*; 1995: Citeseer.
 47. Hulett M, Damodaran P, Amouie M. Scheduling non-identical parallel batch processing machines to minimize total weighted tardiness using particle swarm optimization. *Computers & Industrial Engineering*. 2017;113:425-36.
 48. Marichelvam M, Geetha M, Tosun Ö. An improved particle swarm optimization algorithm to solve hybrid flowshop scheduling problems with the effect of human factors—A case study. *Computers & Operations Research*. 2020;114:104812. 10.1016/j.cor.2019.104812

49. Damodaran P, Diyadawagamage DA, Ghrayeb O, Vélez-Gallego MC. A particle swarm optimization algorithm for minimizing makespan of nonidentical parallel batch processing machines. *The International Journal of Advanced Manufacturing Technology*. 2012;58(9):1131-40.
50. Lian Z. A united search particle swarm optimization algorithm for multiobjective scheduling problem. *Applied Mathematical Modelling*. 2010;34(11):3518-26. 10.1016/j.apm.2010.03.001
51. Afzalirad M, Rezaeian J. A realistic variant of bi-objective unrelated parallel machine scheduling problem: NSGA-II and MOACO approaches. *Applied Soft Computing*. 2017;50:109-23. 10.1016/j.asoc.2016.10.039
52. Zandi A, Ramezani R, Monplaisir L. Green parallel machines scheduling problem: A bi-objective model and a heuristic algorithm to obtain Pareto frontier. *Journal of the Operational Research Society*. 2020;71(6):967-78. 10.1080/01605682.2019.1595190
53. Taguchi G. *Introduction to quality engineering*, Asian productivity organization. Dearborn, Michigan: American Supplier Institute Inc. 1986.

COPYRIGHTS

©2024 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 بی‌امان مصرف انرژی در جهان، چالش‌های پیچیده زیادی از جمله کاهش منابع تجدیدناپذیر انرژی و تشدید انتشار گازهای گلخانه‌ای را به همراه دارد که به تغییرات آب و هوایی کمک می‌کند. در مواجهه با این نگرانی‌های شدید زیست‌محیطی، فشار زیادی به بخش تولید به عنوان مصرف‌کننده مهم انرژی برای اتخاذ شیوه‌های پایدار وارد می‌شود. بنابراین همزمان در نظر گرفتن مدیریت مصرف انرژی و زمان‌بندی عملیات تولید به عنوان یک حوزه محوری برای پرداختن به این چالش‌ها از اهمیت بالایی برخوردار است. زمان‌بندی عملیات مشترک که نمونه آن مسأله برش موجودی در صنایعی مانند مبل‌مان و پوشاک است، یک چالش رایج در محیط‌های تولید است. در این مقاله، برای اولین بار مسأله زمان‌بندی در محیط ماشین‌های موازی یکسان با در نظر گرفتن عملیات مشترک به منظور کمینه نمودن همزمان مجموع زمان‌های تکمیل و مجموع انرژی مصرفی مورد مطالعه قرار می‌گیرد. بدین منظور برای مسأله مورد بررسی، دو مدل برنامه‌ریزی خطی عدد صحیح آمیخته دو هدفه ارائه می‌گردد و برای حل مسائل با ابعاد کوچک از روش محدودیت اسپیلون تکامل‌یافته (AUGMECON) به منظور دستیابی به مجموعه نقاط پارتو بهینه استفاده می‌شود. با توجه به پیچیدگی محاسباتی مسأله، الگوریتم ژنتیک مرتب‌سازی نامغلوب (NSGA-II) و الگوریتم ژنتیک مرتب‌سازی نامغلوب ترکیب شده با بهینه‌سازی ازدحام ذرات (HNSGAI-PSO) برای حل مسائل با ابعاد متوسط و بزرگ و دستیابی به جبهه‌های پارتو تقریبی مناسب، توسعه داده می‌شوند. کارایی و عملکرد الگوریتم‌های پیشنهادی با انجام آزمایش‌های محاسباتی بر روی مسائل نمونه، مورد ارزیابی قرار می‌گیرد. نتایج به دست آمده نشان می‌دهند که در حل مسائل نمونه، الگوریتم HNSGAI-PSO در مقایسه با الگوریتم NSGA-II عملکرد بهتری دارد.

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.



THOMSON REUTERS
WEB OF SCIENCE



Scopus®

